Empathic AuRea: Exploring the Effects of an Augmented Reality Cue for Emotional Sharing Across Three Face-to-Face Tasks

Andreia Valente*
Instituto Superior Técnico
University of Lisbon

Daniel Simões Lopes†
INESC-ID
Instituto Superior Técnico
University of Lisbon

Nuno Nunes‡
ITI / LARSyS
Instituto Superior Técnico
University of Lisbon

Augusto Esteves§
ITI / LARSyS
Instituto Superior Técnico
University of Lisbon

Figure 1: The AuRea system being used by a pair of participants in what we describe as the 'Validation' task. Left: the participant with access to AuRea via an HMD (the *decoder*) observes its counterpart, wired to an ECG, watching a film clip (the *encoder*). Here, the goal of the decoder is to guess the emotional state of the encoder. Right: the feedback from AuRea as seen by the decoder.

ABSTRACT

The Empathy-Effective Communication hypothesis states the better a speaker can understand their listener's emotions, the better can they transmit information; and the better a listener can understand the speaker's emotions, the better can they apprehend the information. Previous emotional sharing systems have managed to create a space of emotional understanding between collaborators on remote locations using bio-sensing, but how a context of face-to-face communication can benefit from biofeedback is still to be studied. This study introduces a new Augmented Reality communication cue from an emotion recognition neural network model, trained using electrocardiogram physiological data (AuRea). The proposed design is meant to facilitate emotional state understanding, increasing cognitive empathy without compromising the existing verbal, nonverbal, and paraverbal communication cues. We conducted a study where pairs of participants (N=12) engaged in three tasks where AuRea was found to positively affect performance and emotional understanding, but negatively affect memorization.

Index Terms: Human-centered computing—Collaborative and social computing—Empirical studies in collaborative and social computing; Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Mixed / augmented reality

1 INTRODUCTION

Recent advances in bio-sensor research towards human-computer interaction have facilitate the use of these devices in everyday situations, beyond lab environments, as they shrink in size and increase in reliability. Much of this research has centered on developing continuous physiological measurements for healthcare monitoring via wearable and noninvasive devices. Commercial wearables such as Fitbit[1] already provide reliable information on sleep quality by measuring body acceleration, and on stress recognition by measuring electrodermal activity [4, 11, 12, 29]. This with the goal of informing users about their behavioral patterns and how real-world situations affect their emotional state so they can better manage them.

These modern sensors have enabled research on emotional sharing through visualizations of physiological data that has been shown to affect communication, interconnection, and collaboration. They have the potential to enable new communication cues that expand on the verbal, nonverbal, and paraverbal cues we tend to rely on to infer another's emotional state, and in turn to allow us to better adapt our own posture or speech to better match our counterpart's needs – in sum, to promote more effective communication [18, 35]. Further, emotional sharing might better enable mirrored emotions, known as emotional contagion, which has been shown to lead to interconnection between collaborators. These emotions promote actions towards a common goal and positively impacting task performance [2]. Both interpersonal emotional understanding and emotional contagion are processes that describe an empathic experience.

Projects on emotional sharing have explored various visualizations of one's emotional state: from sharing biosignals via a smartwatch [18] or ambient light [33], to augmenting video calls with stress indicators [35], to adding an animated plot of someone's heart rate to a transcripts [19]. But while these demonstrated higher emotional understanding and emotional contagion, they are not entirely suited for mobile settings and face-to-face communication as they can limit access to existing communication cues such as facial expressions, posture, voice tone, or gaze. A device with a more appropriate form factor are augmented reality (AR) smartglasses, but research in this area has primarily focused on proof-of-concept prototypes [28] or on assessing collocated task performance [7].

*e-mail: andreia.valente@tecnico.ulisboa.pt
†e-mail: daniel.lopes@inesc-id.pt
‡e-mail: nunojnunes@tecnico.ulisboa.pt
§e-mail: augusto.esteves@tecnico.ulisboa.pt

[1]Fitbit: https://www.fitbit.com/global/us/technology/stress

2642-5254/22/$31.00 ©2022 IEEE
DOI 10.1109/VR51125.2022.00034

As such, this paper starts to fill this gap by exploring emotional sharing in AR-supported face-to-face exchanges. We use bio-sensing to (i) enable a communication cue inferred from an emotion recognition system trained with electrocardiogram (ECG) data; and (ii) present a user study between pairs of participants across from each other and in which one of them has access the physiological data of the other via color-coded AR feedback (Fig. 1). We hypothesize that this will: (**H1**) increase emotional understanding; (**H2**) improve the transmitting of information; (**H3**) improve the apprehension of information; and (**H4**) increase interconnection.

2 Related Work

Researchers have explored the potential of adding bio-signals to the already existing communication cues that play a role in empathic connection (e.g. facial expressions, eye-gaze, and voice pitch) and studied their impact on empathy and collaborative and social tasks.

The MoodLight by Snyder *et al.* [33] and the GER Mood Sweater[2] mapped user's arousal levels to color hues and displayed this information via ambient lights or color-changing garments. The former described a two-user setting where both users' arousal levels contributed to the ambient light. This positively affected collaboration via self-revelation, but introduced a feedback loop where users would be too aware of their own emotional states.

Digital representations of physiological data is another popular research strand in this domain. Researchers have predominantly focused on sharing heart rate visualizations, which tend to be associated with underlying emotional and psychological states. Liu *et al.* [19] demonstrated that providing an animated graph with heart rate information on transcriptions increased emotional perspective-taking and empathic concern towards members of a stigmatized group. Liu *et al.* later developed the Animo [18] system, that allowed users to share their bio-signals voluntarily with each other via a smartwatch app. Users reported that the bio-signal information created new insights into their counterpart's context and prompted discussions about each other's emotional states.

The studies above offer important insights into the potential increase of empathy resulting from different types of emotional cue sharing, but some also can also lead to ambiguous interpretations. For example, by sharing unprocessed signals like heart rate [19], these systems ultimately task users with attributing meaning to information that does not always have a direct link to affective states. For example, the same heart rate reading can mean the person is angry or happy. Others focus on remote communication, and could not easily support face-to-face scenarios as they require the user to look away from the person they are talking to (e.g., to look at the visualization on their smartwatch [18]) and miss on typical communication cues to the same effect such as facial expressions .

Tan *el al.* [35] addressed some of these problems in the context of remote video-mediated assistance. Users had access to cues beyond pure physiological signals, such as video animations and sound, which they could access concurrently. The work found that displaying the heart rate information representation of the worker in the field lowered the workload and the stress of both the instructor and the worker, and increased task engagement for both parties. Instructors were observed engaging in fewer task-irrelevant cognitive interferences, suggesting that the biofeedback visualizations allowed them to spend more time focusing on the task at hand. Ultimately, this study presented key findings supporting the Empathy-Effective Communication hypothesis, and illustrated an approach that effectively preserved verbal, nonverbal, and paraverbal cues in remote communications. It demonstrated that a system that augments live communication cues instead of attempting to replace them can be beneficial to users.

Inspired by these efforts, we expand on the previous approach by presenting the first emphatic AR system that provides an emotional state feedback visualization based on raw physiological data for in-situ face-to-face communication. We describe this as AuRea, and validate our approach via a user study that explores its effects on collaboration performance and interpersonal connection.

[2]GER Mood Sweater: https://sensoree.com/artifacts/ger-mood-sweater

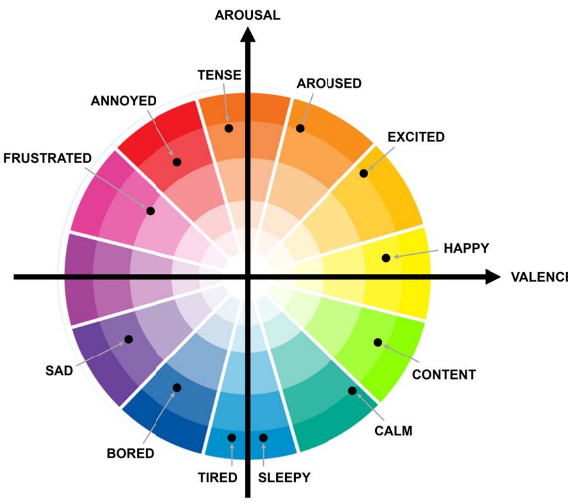

Figure 2: AuRea's color model with three emotions mapped to each quadrant of valence and arousal of Russell's model of affect [30].

3 The AuRea System

AuRea is an AR system for emotional sharing during face-to-face interactions. In order to be able to accurately classify various emotions from physiological data, we conducted a user study where we recorded electrocardiogram data from 21 subjects while watching emotional film clips. This data was then used to train a deep neural network (DNN) regression model, together with the participants' self-reported values of valence and arousal. The predicted emotion from the regression model was mapped to a hue and brightness according to a validated color model, and the resulting color effect was displayed in AR as ripple effect around the user's head (Fig. 1).

3.1 Emotion Recognition Model

3.1.1 Apparatus

Physiological data captured using a BITalino board [3] with an ECG sensor. This operated at 1000Hz, using three electrodes, following a Einthoven's triangle lead-II placement [10]. The physiological signals were transferred wirelessly via a Class II Bluetooth v2.0 module to a desktop computer running the OpenSignals software[4].

3.1.2 Participants

We recruited 21 participants (13 identified as women and eight as men), between 20 and 29 years of age (M = 24.90 SD = 2.20). Participants described having no cardiac problems and no diagnosed daltonism. Because this experiment took place during the COVID-19 pandemic, only participants who had received both doses of the vaccine at least two weeks prior to the start of the study were invited to participate. Finally, A list containing the American Motion Picture Association film ratings for the clips depicting mature content and their official content warnings was provided as part of our study recruitment invitation. Participation was voluntary and no compensation was provided.

[3]https://bitalino.com/products/plugged-kit-dual-mode-ble-bt
[4]https://biosignalsplux.com/products/software/opensignals.html

3.1.3 Procedure

Upon arrival, participants were informed that they would watch various short clips lasting approximately two minutes and that, after each clip, they would report on their emotional state by filling out a questionnaire. They were encouraged to report on their actual feelings instead of what they believe they should have felt while watching the clip. Before starting, participants filled in both the Questionnaire of Cognitive and Affective Empathy (QCAE) [27], which assesses cognitive and affective empathy, and a custom emotion-to-color mapping questionnaire (described in detail below). Next, ECG electrodes were placed by participants on their bodies using a reference figure, and the signal was checked for any interference or incorrect electrode placement. Over-ear headphones were provided, the lights were dimmed, and videos played on a 21.5” HD monitor. Due to the COVID-19 pandemic, the researcher wore a mask during the study sessions. All equipment was thoroughly disinfected in-between sessions, and the experiments took place in an otherwise empty and well-ventilated room.

3.1.4 Metrics (Participant Self-Assessment)

After watching each film clip, participants provided a self-assessment of their emotional state via a 9-point version of Self Assessment Manikin (SAM) scales measuring arousal (calm to aroused) and valence (negative to positive) [3]. They also mapped the clip to an emotional label, and were asked by the researcher if they had seen the clip in question before. The list of emotional labels contained three labels from each quadrant. The first quadrant (positive valence, high arousal) or (+V, +A) contained the labels ‘Happy’, ‘Excited’, and ‘Aroused’. The second quadrant (-V, +A) contained ‘Frustrated’, ‘Annoyed’, and ‘Tense’. The third quadrant (-V, -A) contained ‘Sad’, ‘Bored’, and ‘Tired’. And the fourth quadrant contained the labels ‘Sleepy’, ‘Calm’, and ‘Content’. These labels are part of the Russell’s circumplex model of affect [30] and were chosen as they can cover a wide range of theoretical levels of arousal and valence, and because they can be used describe various face-to-face contexts.

3.1.5 Task (Emotional Film Clips)

We used 19 clips in total that mapped to quadrants of valence and arousal from Russell’s circumplex model of affect. Nine were chosen from the FilmStim database by Schaefer *et al.* [32] representing emotions relevant to collaboration and social interaction such as sadness or anger: three clips mapped to neutral valence and neutral arousal (V, A), two to (-V, +A), two to (-V, -A), and two clips mapped to (+V, +A). Ten additional clips were chosen by us to represent relevant emotions not covered in the FilmStim database such as excitement, stress, or calmness. The latter mapped to an underrepresented valence and arousal quadrant in the FilmStim database: (+V, -A). Our ten custom clips mapped to these quadrants as follows: four clips mapped to (+V, -A), two to (-V, +A), two to (-V, -A), and two clips were mapped to (+V, +V). We followed Schaefer et al.’s protocol regarding how these clips should be edited, and we used scenes from movies and TV shows such as Marley & Me, Mystic River, or Brooklyn Nine-Nine.

Participants watched these clips following Schaefer *et al.*’s protocol: they never watched two clips targeting the same valence consecutively; the order of the clips in each quadrant was randomized; the valance of the first clip was counterbalanced between participants; and before each clip participants completed a 20 s breathing exercise to minimize stimuli effects across clips.

3.2 Color Model

At the start of the study, participants performed an emotion-to-color mapping between the 12 emotional labels (three labels per quadrant) and 12 colors. These were the primary, secondary and tertiary colors of the Red-Yellow-Blue (RYB) color model, seen in Fig. 2. Different emotions could be mapped to the same color, and the order in which these emotional labels were presented to participants was randomized. Emotion-to-color models have been widely used in human emotion research, mostly based on The Plutchik’s Wheel of Emotion [25]. This considers eight primary opposing emotions (e.g., joy and sadness) and maps these to a primary or secondary color. More recent works have continued to explore the relationship between emotions and color (e.g., [23]). We used the RYB color instead of the more commonly used RGB because the former features more prominently the yellow and red colors. We presumed participant’s emotions would range from (+V, A) to (-V, +A), which mapped to yellow and red colors in the theoretical models, respectively.

Based on the results from participant’s responses, warm colors were used to represent high arousal emotional states from the labels in the first (87% of responses) and second quadrants (85%). Cool colors were used to represent low arousal emotional states from the third (79%) and fourth quadrants (86%). Overall, these results are aligned with Plutchik’s mapping. As such, we adopt the latter’s use of brightness – our final color model can be seen in Fig. 2.

Table 1: All 29 features extracted to build AuRea’s emotion recognition model. 10 features were selected for predicting *angle* (bold), and 10 for *distance* (underlined) across the valence-arousal axis.

Domain	ECG Feats	Statistical Feats
Time	Signal	<u>min</u>, <u>max</u>, <u>mean</u>, **var**
	HRV	**<u>RMSSD</u>**, <u>MeanNN</u>, **SDNN**, **SDSD**, CVNN, **CVSD**, <u>MedianNN</u>, MadNN, IQRNN, <u>pNN50</u>, **pNN20**
	HR	<u>BPM</u>
Frequency	HRV	**<u>VLF</u>**, **LF**, **HF**, VHF, LF/HF
Non-Linear	Poincaré plot	SD1, SD2, SD1/SD2
	EDR	<u>min</u>, max, **var**, mean, RSP_{rate}

Table 2: The error measured for angle and distance prediction per quadrant (valence-arousal) of AuRea’s emotion recognition model.

	Angle		Distance	
Quadrant	**MAE**	**RMSE**	**MAE**	**RMSE**
1	36.12	50.80	0.61	0.80
2	20.85	28.42	0.62	0.93
3	24.98	30.81	0.87	1.60
4	23.24	31.93	1.07	2.51

3.3 Regression Model

We relied on the NeuroKit 2.0 Python packages [21] and the BioSPPy [6] for all ECG processing. Regarding the self-reported answers for valence and arousal, these were translated into polar coordinates following the emotion-recognition model proposed by Han et al. [13]. Each self-reported pair of valence and arousal was converted into (φ, Θ), $\{\varphi \varepsilon \mathbb{Q} \mid 0 \leq \varphi \leq 3\}$ and $\{\Theta \varepsilon \mathbb{N} \mid 0 \leq \Theta \leq 360\}$.

3.3.1 Pre-processing

Even though participants were instructed to stay still, the ECG signal still suffered from common noise. The ECG signal pre-processing procedure consisted in applying a IIR 3rd-order Butterworth bandpass filter between the 2 to 45Hz frequencies [5]. This was done to mitigate the effects of baseline wander caused by the participants’s breathing or body movements, and power-line interferences. For QRS complex detection, we used the Pan–Tompkins algorithm which squares the signal to delineate the QRS signal contribution and apply adaptive thresholds for the each peak [24].

3.3.2 Windowing

As the system is intended for real-time emotion detection, we opted to employ the Ultra-short-term protocol by Salahuddin *et al.* [31]

and a window of 60 s for the analysis of time and frequency domain features such as heart rate (HR), RMSSD, and heart rate variability (HRV). If the latest 60 s included body movement or electrode misplacement artifacts, this was discarded.

3.3.3 Feature Extraction, Normalization, and Selection

A total of 29 features, listed in Table 1, were extracted from 60 samples. The feature normalization process was performed per participant, in order to normalize participant data to the rate of change in physiological reaction. All samples relating to self-reported states of emotional neutrality were used as the personalized baseline for each participant (i.e., clips that participants mapped to a neutral emotional label). Every feature was then normalized to the rate of change from the baseline:

$$u_{ij} = ((x_{raw})_{ij} - (w_{baseline})_{ij})/(w_{baseline})_{ij} \quad (1)$$

where u_{ij} represents the normalized ith feature value for the jth participant, $(x_{raw})_{ij}$ the input feature value, and $(w_{baseline})_{ij}$ the feature value measured on the baseline samples.

3.3.4 DNN Architecture

With the dataset ready for training, two supervised DNNs were created to predict the two polar coordinates dimensions: one for angle prediction and other for distance to center prediction. The DNNs had two hidden layers, constructed by repeated experimental training: the input layer had 10 neurons (the size of the feature vector), the first hidden layer had six neurons, and the second hidden layer had three neurons. The output layer had the one neuron needed for a regression model. The activation function used was the Rectified Linear Unit with a Root Mean Squared Propagation optimizer. We trained the model using a learning rate of 0.001 and a validation split of 0.2, and using a batch size of 32 for prediction.

3.3.5 Results

The robustness of the system was evaluated using 10-fold cross-validation. First we calculated the Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) for each polar coordinates dimension on the overall predictions. Afterwards, the predictions were divided by the four quadrants and each quadrant's MAE and RMSE was calculated. The predictions were then converted to labels of their predicted quadrant and an accuracy score per quadrant was calculated. Converting the angle and distance predictions into a two-dimensional point on the valence-arousal axis, the model reached a MAE of *1.22* and a RMSE of *2.04*. Angle prediction had an overall MAE of *23.24* and RMSE of *31.92*, and distance prediction had an overall MAE of *0.77* and RMSE of *1.320*. The highest overall quadrant accuracy was 70.79%, with 0.72 of precision, 0.71 of recall, and a f-score of 0.70. The best results for the quadrant specific evaluation metrics can be seen on table Table 2, and the confusion matrix for quadrant classification can be seen in Fig. 3.

The results of the developed emotion recognition model did not reach the state-of-the-art standards of similar models using ECG data. Self-supervised learning models have achieved 82.78% [17] of accuracy in a four-quadrant classification problem. Chen *et al.* [8] achieved 82.63% and 74.88% accuracy for valence and arousal, respectively, using fusion of long short-term memory networks and, in Zhang *et al.* [36], an accuracy of 92% for a four-quadrant classification problem was achieved, using a combination of K-Nearest Neighbors algorithm with a Max-Min Ant System feature selection. The 70.79% of accuracy achieved on the developed model was deemed appropriate to its experimental purpose.

3.4 The AuRea System Architecture

In order to enhance empathic experiences in face-to-face communication, we propose the Empathic AR system or AuRea: an AR system that presents a colorful representation of physiological data

Figure 3: The confusion matrix for the arousal-valence quadrant classification of AuRea's emotion recognition model.

acquired from a single-lead ECG sensor. AuRea aims to add an emotional cue to the verbal, nonverbal, and paraverbal cues that already play a role in the encoding and decoding of emotional states between people. It was developed for a two-user setting, where one user is connected to ECG sensors that send physiological data to a processing pipeline that predicts a point on a two-dimensional model of valence and arousal, as described so far. The user's counterpart is equipped with an HTC VIVE Pro Eye headset[5] and a ZED Mini camera[6] that enable an AR video see-through experience. This is displayed with a resolution of 720p in order to achieve a performance of 60 frames-per-second (FPS) with a Field-of-View (FOV) of 90º (H) $\times$ 60º (V). This enables the wearer to naturally observe the person in front of them and important communication cues such as body posture or facial expressions, and for AuRea to display its emotional feedback inferred from the ECG data: a color-coded ripple effect that surrounds the body of the person in view.

AuRea identifies each pair of users as follows: the **encoder** is the one that expresses or is associated with an emotional cue (i.e., it is the one connected to the ECG sensors); the **decoder** is the one that attemps to interpret the emotional cue displayed (i.e., the one equipped with the AR headset). For stability during the user study we will describe next, AuRea's ripple effect uses an ArUco marker[7] placed on the wall behind the encoder for positioning in 3D space, but is designed as a 2D billboard with camera position alignment. To preserve the communication cues of the encoder, it was used a stencil buffer placed on the same plane as the billboard, with the shape of the human upper-body (Fig. 1).

The data from the encoder's ECG sensor is sent via Bluetooth using the OpenSignals' Lab Streaming Layer to a desktop computer. This processes the data with the regression pipeline detailed earlier, using the same 60 s segment but sliding the window every 2 s, discarding the initial 2 s of the window, and adding the signal points from the most recent 2 at its end. 2 s were used because the intelligent model takes around 1.8 s to complete its pipeline. The model outputs the polar coordinates predicted (angle and distance) and sends the two values to a second computer by a crossover LAN wired connection, using the a TCP/IPv4 protocol. The second computer receives the values on the Unity programming environment[8] and converts the polar coordinates into hue and brightness, according to the color model we described above. The predicted angle and distance to the center of the valence-arousal axis were also coded as

[5]HTC Vive Pro Eye: https://www.vive.com/us/product/vive-pro-eye/
[6]ZED Mini camera: https://www.stereolabs.com/zed-mini/
[7]ArUco marker: https://docs.opencv.org/4.5.3/d9/d6a/group__aruco/
[8]Unity programming enviroment: https://unity.com/

the velocity of the ripple affect: higher arousal is represented by a faster ripple effect. This second computer was used solely for the sake of performance during the user study we describe next.

4 User Study

The evaluation of AuRea took place across three tasks in a face-to-face setup between pairs of participants. Beyond system validation of the AR ripple effect, this study explored its effects on self-reported metrics of interconnection, cognitive load, engagement, worry, and distress, and on quantitative metrics of performance.

4.1 Participants

We recruited 12 participants (six identified as women and six as men), between 20 and 29 years of age (M = 24.33 SD = 1.70). All took part in the first study and their data was part of the training of the regression model. This was also done to mitigate a confounding factor on self-perceived empathy: only participants with cumulative QCAE scores inside the interquartile range were considered (captured in the first study). The 12 participants were divided in six pairs. To mitigate further confounding factors, pairs were equally characterized by gender (female-female, male-male, and female-male pairs) – to minimize known gender differences in empathic disposition [9] – and by familiarity (three pairs of participants had never met before the experiment). The latter was further defined via the Inclusion of Other in the Self (IOS) scale, a tool commonly used to measure perceived inter-personal closeness [1].

4.2 Procedure and Metrics

Upon arrival participants placed the ECG sensors on their bodies as they had done before in the first study. Both participants watched a 2-minute video of a television weather forecast while their physiological signals were recorded and pre-processed [34], and the features for our model's baseline were extracted – effectively establishing an emotionally neutral physiological baseline. We followed the same COVID-19 precautions as in the previous study, thoroughly cleaned the equipment when participants swapped roles at the end of each task, and added plexiglass between the pair of participants.

4.2.1 Task 1: 'Validation'

As illustrated in Fig. 1, the encoder was given headphones and was asked to watch a film clip from the FilmStim database [32]. Afterwards, both the encoder and the decoder were asked to provide an assessment of emotional state using SAM scales of valence and arousal as they had done in the previous study. The encoder reported on its own emotional state, while the decoder reported on the emotional state of the encoder as they had perceived it. The encoder was instructed to watch the video naturally, without over-reacting. As this served as the habituation task, the decoder was allowed to consult a printed figure of our color model at any point. While providing their assessment of the encoder's emotional state, the decoder was also asked to rate (from 0 to 100) the effect of the AuRea system in their assessment (in comparison to other available cues, such as facial expressions). Each encoder watched two clips in a random order: one with a negative and another with a positive valence.

4.2.2 Task 2: 'Pattern Blocks'

As illustrated in Fig. 4, at the start of the task the decoder was given one figure out of four (selected at random). Each figure represented a tangram of equal difficulty (i.e., same number of pieces and representing an animal that was easy to recognize). The decoder was told they could not show the figure to the encoder, but instead had to instruct them on how to build the tangram. 24 wooden pieces were place in front of the encoder, inside a box, that hid them from the decoder. There was no time limit to this task, and participants were instructed to speak freely to one another. We assessed participant performance (task duration and tangram correctness), interconnection (ratio for pre- and post-task IOS questionnaires), perceived cognitive load (the NASA-TLX [14]), and engagement, distress, and worry (ratio for pre- and post-task Short Stress State Questionnaires, SSSQ [15]).

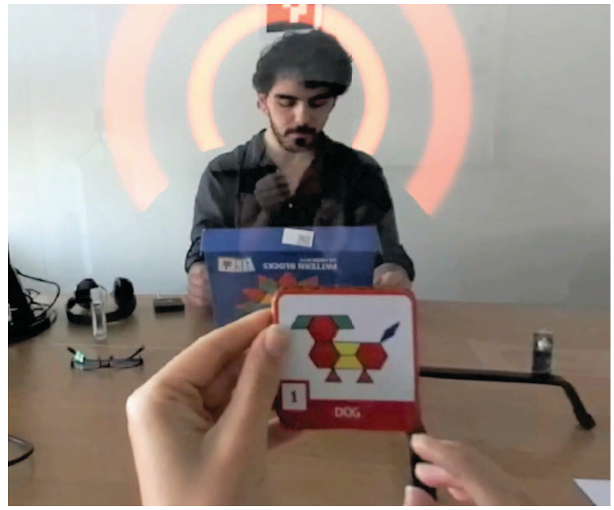

Figure 4: A screenshot from the perspective of a decoder in our second task ('Pattern Blocks'). In this task, the decoders were handed an image of a tangram they had to convey to the encoders so that they could recreate it using wooden blocks. The decoders could not see the encoders' progress, only their natural and augmented cues.

4.2.3 Task 3: 'Storyteller'

This task is based on an experiment by Ramsberger *et al.* [26] that measures transactional success in conversations. The encoder was asked to watch a 2-minute video and retell its events to the decoder with as much detail as possible. The decoder had to be able to confidently re-convey this story, and was free to engage in a semi-structured conversation with the encoder to extract as much information as needed (no time limit was imposed). This task is at the core of social empathic interactions, where participants have to co-construct meaning out of verbal, nonverbal, and paraverbal cues that can positively or negatively affect these interactions (e.g., how the voice tone or body posture affects the meaning of a word or idea). We assessed participant performance during these interactions, interconnection (ratio for pre- and post-task IOS questionnaires), and the perceived cognitive load (the NASA-TLX).

4.3 Experimental Design

The user study primarily followed a within-subject design with a single independent variable: the use or not of the AuRea system (counterbalanced across sessions). The first task was only executed in the AuRea condition, and tasks were always completed in the order they are presented above (one to three).

5 Results

In the following sections, the results from the participants' responses to the various questionnaires and open-interviews, and the tasks' quantitative metrics are detailed. Statistical tests were performed using the IBM SPSS v26.

5.1 System Use and Accuracy

We asked participants in the first habituation task to report on a scale from 0 to 100 the importance or use of AuRea while assessing the encoder's emotions, compared to relying on traditional face-to-face communications cues. Participants reported a mean value of 59.58 (*SD = 29.85*). We then compared decoders' accuracy by comparing the cumulative MAE of their valence and arousal reports and encoders' self-assessments. Two groups were created, those who reported a predominant use of AuRea (≥ 50) and those we relied on more natural cues (< 50).

A paired samples t-test found significant differences in accuracy between these two groups use (t(5) = -2.712, p = .042): trials with a predominant use of AuRea achieved a smaller MAE (M = 1.67, SD = 0.82) than trials with predominant use of natural communication cues (M = 3.33, SD = 1.03). Further, no significant differences were found in accuracy for familiarity (t(11) = -.233, p = .820) or for groups with a predominant use of AuRea across familiarly levels (t(11) = -.991, p = .343).

5.2 Effects on Performance

Paired samples t-test were performed across the 'Pattern Blocks' and 'Storyteller' tasks, comparing the performance of pairs of users with AuRea and without (baseline) – Fig. 5.

5.2.1 Task 2: 'Pattern Blocks'

Participant performance in the 'Pattern Blocks' task was measured as the ratio of correctly placed pieces / total number of pieces, AuRea had a positive effect on performance (M = .95, SD = .10) when compared to the baseline (M = .79, SD = .19): t(11) = -3.80, p = .003. No significant differences were found for task duration between the AuRea condition (M = 258 s, SD = 127) and the baseline (M = 215 s, SD = 124): t(11) = -.851, p = .413.

5.2.2 Task 3: 'Storyteller'

Participant performance in the 'Storyteller' task was measured as the ratio of the correct number of key ideas the decoder was able to retell by the number of key ideas the encoder retold from the video. A video had 30 key ideas compiled by the researcher, who then match them to a transcription of these interactions. We found that AuRea had a negative effect on performance (M = .61, SD = .15) when compared to the baseline (M = .70, SD = .12): t(11) = 3.303, p = .007. Task duration was not measured for this task.

5.3 Effects on Perceived Workload

The participants answers to the NASA-TLX questionnaire were analyzed using the Wilcoxon Signed-Rank test and the raw NASA-TLX scores. We report exact and not asymptotic p-values due to our relatively small sample size [22].

5.3.1 Task 2: 'Pattern Blocks'

- **Decoder:** significant differences were found for the combined task load index (Z = -2.984, p = .006) where the AuRea (M = 78.42, SD = 14.13) increased the perceived task load when compared to the baseline (M = 53.00, SD = 15.11). No significant differences were found for mental demand (Z = -1.561, p = .549), temporal demand (Z = -1.964, p = .065), overall performance (Z =-.268b, p = 1.000), effort (Z = -1.140b, p = .774), or frustration levels (Z = -1.729, p = .146). Significant differences were found for physical demand (Z = -3.065, p $< .001$) where decoders found interacting with AuRea (M = 14.17, SD = 3.41) more physically demanding than without (M = .83, SD = 1.27).

- **Encoder:** no significant differences were found for the combined task load index (Z = -.178, p = 1.000), mental demand (Z = -.446, p = 1.000), physical demand (Z = -.155, p = .754), temporal demand (Z = -1.592, p = .774), overall performance (Z = -.670, p = 1.000), effort (Z = -.625, p = .549) or frustration levels (Z = -.579, p = 1.000).

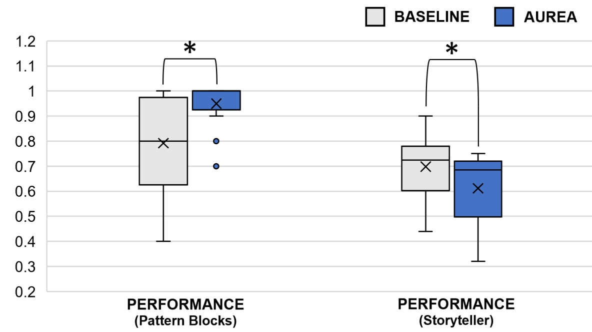

Figure 5: Performance results during the 'Pattern Blocks' (correctly placed pieces / total number of pieces) and 'Storyteller' tasks (key ideas retold by the decoder / key ideas retold by the encoder). * denotes $p < .05$

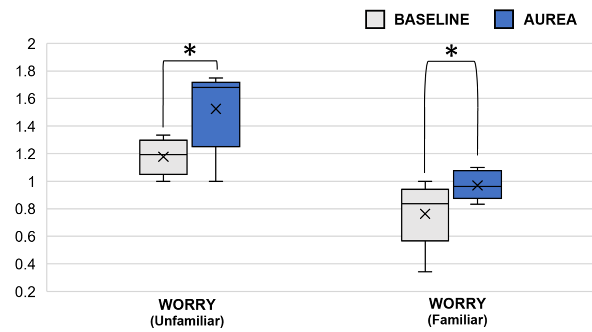

Figure 6: Worry results (ratio between pre- and post-experiment scores) from the encoder in the 'Pattern Blocks' task, across *familiar* and *unfamiliar* pairs of participants (i.e. participants who did not know each other before the start of the study). * denotes $p < .05$

5.3.2 Task 3: 'Storyteller'

- **Decoder:** the AuRea negatively affected decoders' combined task load index (M = 55.33, SD = 11.35) when compared to the baseline (M = 36.25, SD = 7.49): Z = -3.063, p < .001. No significant differences were found for mental demand (Z = -2.591, p = .065), temporal demand (Z = -1.543, p = .125), overall performance (Z =-1.130, p = 1.000), effort (Z = -2.242, p = .146), or frustration levels (Z = -1.940, p = .109). Significant differences were found for physical demand (Z = -3.064, $p < .001$), where the AuRea condition (M = 12.58, SD = 5.32) was reported by the decoders to be more physically demanding than the baseline (M = .83, SD = 1.99).

- **Encoder:** no significant differences were found for the task load index (Z = -.132, p = 1.000), mental demand (Z = -.633, p = 1.000), physical demand (Z = -.281, p = 1.000), temporal demand (Z = -0.323, p = 1.000), overall performance (Z = -.115, p = 1.000), effort (Z = -.403, p = .741), or frustration levels (Z = -.604, p = 1.000).

5.4 Effects on Engagement, Distress, and Worry

We used the Wilcoxon Signed-Rank test to analyze participants' ratio between pre- and post-experiment SSSQ scores during the 'Pattern Blocks' task (Fig. 6). As before, we report exact and not asymptotic p-values.

- **Decoder:** no significant differences were found on components of engagement (Z = -.157, p = 1.000), distress (Z = .000, p = .549), or worry (Z = -.356, p = .774).
- **Encoder:** no significant differences were found for components of engagement (Z = -.711, p = .477) or distress (Z = -.133, p = .894), but were found for worry (Z = -2.312, p = .021). Further, a Friedman tests revealed significant difference for worry between familiarity groups ($\chi^2(3)$ = 15.362, p = .002). Encoders partnered with a participant they were unfamiliar with (IOS = 1) experienced overall higher levels of worry in both the baseline (M = 1.18, SD = .13) and AuRea conditions (M = 1.53, SD = .30), when compared to encoders paired with a familiar participant (IOS $>$ 1): baseline (M = .76, SD = .24) and AuRea (M = .969, SD = .104). In sum, all encoders experienced higher levels of worry using AuRea.

5.5 Effects on Interpersonal Connection

We used the Wilcoxon Signed-Rank test to analyze participants' ratio between pre- and post-experiment IOS scores during the 'Pattern Blocks' task. We report exact and not asymptotic p-values.

5.5.1 Task 2: 'Pattern Blocks'

- **Decoder:** we found significant differences for interpersonal connection (Z = -1.958, p = .050), where decoders reported a higher increase in connection with AuRea (M = 1.66, SD = .83) than without (M = 1.13, SD = .33).
- **Encoder:** we found significant differences for interpersonal connection (Z = -2.113, p = .035), where encoders reported a decrease in connection when the decoder was AuRea (M = .91, SD = .22) compared to the baseline (M = 1.44, SD = .62).

5.5.2 Task 3: 'Storyteller'

- **Decoder:** no significant differences for interpersonal connection were found (Z = -1.054, p = .292).
- **Encoder:** no significant difference for interpersonal connection were found (Z = -.272, p = .785).

6 Discussion

6.1 System Use and Accuracy

All participants were quick to understand and memorize the color model used on the system and rarely did they consult the printed reference image with the color model. The high perceived accuracy of the system was first seen on the 'System Validation' task, where participants relied strongly on AuRea information to infer their partner's emotional state. The significant higher accuracy on this inference with higher use of AuRea might be related to the Hawthorne effect, which led encoders to suppress visible emotional reactions to the emotional stimuli and made it more difficult to read the emotional cues in non-extreme emotional states or with encoders that are usually less expressive. With these suppressed emotional reactions, which are common in formal settings or between unfamiliar people, for example, AuRea provides a way to see beyond the controllable layer of emotions. Nonetheless, the higher accuracy of emotional understanding with the use of AuRea validates the first research hypothesis (**H1**).

The level of familiarity was expected to impact the use of the system, as unfamiliar people had less experience on decoding their partner's emotional state and were expected to rely more on the system. This assumption was proven to be wrong on the sampled group and might be linked with the quicker entrustment on the system by the decoders that were familiar with their partners. The familiar groups seemed to validate the AuRea visualization quicker because they were able to better decipher their partner's emotional state without the system, thus, observing quicker and trusting more strongly the accuracy of the system.

6.2 Effects on Performance

The effects of the proposed design on metrics of performance seem to be task-dependent. On the 'Pattern Blocks' task, the increase of performance with AuRea seem to validate the hypothesis presented by Hogan et al. [16] that the more accurately a speaker can decipher the emotional state of their listener, the more effectively can they transmit the message, which in turn supports this study's second hypothesis (**H2**). Since trials with AuRea did not have a significant higher duration, the increase in performance might be attributed to the expected change in instructional style by the decoder. The participants were asked if they felt any change in how they decided to perform the task with the system, to which some participants declared, substantiating the mentioned hypothesis, that AuRea helped them understand behavioral patterns in their partner. Two participants noted that accessing the partner's emotional state helped them define the pace of the task, as they adapted to the partner's nervousness. Three participants reported that AuRea was not relevant for their behavior during the task. One of them said they tried to abstract the system from the task, as the considerable use of color was too distracting. The other two declared that they believed the task would not be affected by acknowledging the partner's emotional state.

On the other hand, the decrease in performance on the 'Storyteller' task shows that the proposed design is not appropriate for tasks involving the retaining of information. The part of the *Empathy-Effective Communication hypothesis* that states that better perspective-taking by the listener leads to better message apprehension was not observed on this study and, thus, our third hypothesis was rejected (**H3**). From the participants' statements, it was clear that the system and the task setup promoted more complex perspective-taking experiences, which made it a viable emotional sharing tool for a pure social interaction, but its design was too obtrusive for tasks that demand memorization. This insight should lead to different design approaches that can be accessed without a diversion of attention, such as opting for a slower color transition and the removal of ripple velocity.

6.3 Effects on Perceived Workload

The increase of physical demand when using AuRea was expected, as the video see-through setup provided a much lower resolution (720p) of the physical environment and of the encoder. Some participants found difficult the use of the VIVE headset and reported eyestrain, as well as motion sickness, on prolonged tasks. This limitation was then reflected on the answers for system usability where 4 participants reported they had difficulty using the system. The higher physical demand on the 'Pattern Blocks' than on the 'Storyteller' task might be related to the requirement for more dynamic vision target. On the 'Storyteller', the decoders had to focus only on the upper body of the encoder, without needing to move, while on the 'Pattern Blocks' task, the decoders had to diverge their focus point from the encoder to the reference image often. With the low FOV, 90°(H) x 60°(V), and the depth perception distorted, since the Zed camera was about 6 cm away from the eyes, the higher physical demand was correctly expected to become more evident on the 'Pattern Blocks' task.

Tan *et al.* [35] reported an increase of perceived performance, by both worker and instructor, when biofeedback was introduced to the task, which was not observed on this study, even with metrics of performance increasing with the biofeedback. This might be related to task characteristics, since the study by Tan *et al.* allowed instructors to have access to the worker's progression, thus allowing for exchanges of instructor reassuring correct placement, which might be correlated to final perceived performance. In this study, perceived performance was presumed to be correlated, for the encoder, to discernment of final pattern image and, for the decoder, to the encoder's confirmation of comprehensibility, characteristic that were not presumed to be impacted significantly by AuRea.

6.4 Effects on Engagement, Distress, and Worry

The AuRea condition showed no effect of the biofeedback to metrics of engagement or distress for the encoder and the decoder, but an increase of worry for the encoder. On the SSSQ, the worry component is connect to statements of awareness of oneself as separate from others (e.g. "I feel concerned about the impression I am making") and overall personal insecurity (e.g. "I feel self-conscious"). With the unnatural over-exposure of an emotional sharing system, it was expected higher levels of worry, as concealing increased nervousness became a more difficult task. Moreover, the higher levels of worry relating to groups with low IOS score is consonant with the assumption that emotional exposure to unfamiliar partners will naturally make participants more self-conscious of their own emotional state. On the final questionnaire, when asked to rate the sentence "I felt comfortable sharing my emotional state with the other person through the AuRea system", both participants from group 3, a low IOS score group, stated they did not feel as much discomfort as they would on real-world applications.

6.5 Effects on Interpersonal Connection

In our fourth hypothesis (**H4**), we excepted the interconnection between participants to increase, but this hypothesis was only partially supported by the results, since decoders reported an increase in interconnection while the encoders reported a decrease in interconnection. AuRea was able to achieve higher cognitive empathy from the decoder to the encoder, as revealed on the validation task, and this increase of emotional understanding is presumed to have been translated into perceived interpersonal connection, as decoders reported higher connection with the encoders while instructing them on the 'Pattern Blocks' task. As mentioned previously, decoders were able to better understand the needs of the encoders while performing the assembly task and to adapt their instructional style, better synchronizing the two parties on the task. However, encoders reported lower interpersonal connection on the same task, which, using the same hypothesis, could mean that encoders lost emotional understanding when the decoder was using AuRea, which shows a one-sided increase of empathy that results on the other side's lost of empathy. The adaptation of the decoder's behavior to the encoder's needs did not overpower, in terms of interconnection, the consequences of removing some of the decoder's important communication cues, like gaze and facial expressions. This disadvantage could be managed by replacing the video see-through with an optical see-through system that would preserved the decoder's communication cues.

6.6 Limitations and Future Directions

The main limitation of our experiments is their small sample size, both for the emotion recognition training and for the final user study. The main reason for this was of course the COVID-19 pandemic and our strict recruitment policy (e.g., only participants who had received both doses of the COVID-19 vaccine at least two weeks prior to the start of the study were invited to participate). Collecting more data to train would increase model robustness and allow for the use of the system as a generalized emotion recognition model, instead of the personalized model for the sample group constrained by, e.g., their age range. Moreover, a larger sample size would increase the reliability of the self-reported metrics. One limitation observed during the data analysis was how reliable the metrics of engagement, worry and distress were on self-reported metrics. One aspect not explored in our work was the emotional contagion phenomenon – a normal consequence of cognitive empathy. In order to explore this, future work would need to collect the decoder's physiological data as well in order to study physiological changes during the tasks and emotional state synchrony between the pairs of participants.

Finally, with the validation of the system as a cognitive empathy support tool, it would be important to study how the system could facilitate social interactions for people on the autism spectrum or with similar conditions that affect the capacity for cognitive empathy, as these conditions make the reading of communication cues more difficult [20]. For settings with an instructor-student dynamic, as in a classroom scenario, the AuRea system, as well as previous systems of emotional sharing, seem to have evident benefits when the instructor can access the student emotional state. It would be of value to expand the work to a scenario closer to a classroom setting with multiple encoders/students and to understand the design implications of an increased space of physiological information and, more importantly, what technique of self-focus would need to be implemented to reduce the level of distraction of an augmented space of such dimensions (e.g., the gaze-assisted tool presented in [28]).

7 Conclusion

On this project, we developed a emotional recognition DNN model, which served to infer emotional states from ECG data and created an AR system to showcase emotional state through color and movement. We evaluated, in a face-to-face setting, the effect of this emotional sharing system in task performance, interpersonal connection, cognitive load, engagement, distress and worry. A user study was conducted where we validated the system as an effective tool for the increase of emotional understanding and then compared a baseline variant with a system variant during two collaborative task relating to instruction-giving and memorization. Following a within-subjects design with 6 pairs of participants, we found that the instruction-giving task saw an increase on task performance while the memorization task saw a detriment to performance. The system introduced increased worry for the people disclosing physiological data, where participants paired with participants there were not familiar with reported higher levels of worry than participants who were paired with friends or acquaintances. The system increased interpersonal connection for the participant accessing the other's emotional visualization but decreased interpersonal connection for the participant showcasing the emotional state. We discussed the implication of these finding and suggested future directions for emotional sharing system in face-to-face settings.

Acknowledgments

This research was supported by the Fundação para a Ciência e a Tecnologia through grants UIDB/50009/2020 and UIDB/50021/2020, and by the Agência Regional para o Desenvolvimento da Investigação, Tecnologia e Inovação through grant M1420-01-0145-FEDER-00000.

References

[1] A. Aron, E. N. Aron, and D. Smollan. Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of personality and social psychology*, 63(4):596, 1992.

[2] S. G. Barsade. The ripple effect: Emotional contagion and its influence on group behavior. *Administrative science quarterly*, 47(4):644–675, 2002.

[3] M. M. Bradley and P. J. Lang. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1):49–59, 1994. doi: 10.1016/0005-7916(94)90063-9

[4] A.-M. Brouwer, M. A. Neerincx, V. Kallen, L. van der Leer, and M. ten Brinke. Eeg alpha asymmetry, heart rate variability and cortisol in response to virtual reality induced stress. *J. Cyberther. Rehabil*, 4:21–34, 2011.

[5] S. Butterworth et al. On the theory of filter amplifiers. *Wireless Engineer*, 7(6):536–541, 1930.

[6] C. Carreiras, A. P. Alves, A. Lourenço, F. Canento, H. Silva, A. Fred, et al. Biosppy: Biosignal processing in python. *Accessed on*, 3(28):2018, 2015.

[7] L. Chen, Y. Liu, Y. Li, L. Yu, B. Gao, M. Caon, Y. Yue, and H.-N. Liang. Effect of visual cues on pointing tasks in co-located augmented reality collaboration. *arXiv preprint arXiv:2110.04045*, 2021.

[8] T. Chen, H. Yin, X. Yuan, Y. Gu, F. Ren, and X. Sun. Emotion recognition based on fusion of long short-term memory networks and svms. *Digital Signal Processing*, 117:103153, 2021.

[9] L. Christov-Moore, E. A. Simpson, G. Coudé, K. Grigaityte, M. Iacoboni, and P. F. Ferrari. Empathy: Gender effects in brain and behavior. *Neuroscience & biobehavioral reviews*, 46:604–627, 2014.

[10] M. B. Conover. *Understanding electrocardiography*. Elsevier Health Sciences, 2002.

[11] D. Egan, S. Brennan, J. Barrett, Y. Qiao, C. Timmerer, and N. Murray. An evaluation of heart rate and electrodermal activity as an objective qoe evaluation method for immersive virtual reality environments. In *2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX)*, pp. 1–6, 2016. doi: 10.1109/QoMEX.2016.7498964

[12] L. M. Feehan, J. Geldman, E. C. Sayre, C. Park, A. M. Ezzat, J. Y. Yoo, C. B. Hamilton, and L. C. Li. Accuracy of fitbit devices: Systematic review and narrative syntheses of quantitative data. *JMIR Mhealth Uhealth*, 6(8):e10527, Aug 2018. doi: 10.2196/10527

[13] B.-j. Han, S. Rho, R. B. Dannenberg, and E. Hwang. Smers: Music emotion recognition using support vector regression. In *ISMIR*, pp. 651–656. Citeseer, 2009.

[14] S. G. Hart. Nasa-task load index (nasa-tlx); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9):904–908, 2006. doi: 10.1177/154193120605000909

[15] W. S. Helton and K. Näswall. Short stress state questionnaire. *European Journal of Psychological Assessment*, 31(1):20–30, 2015. doi: 10.1027/1015-5759/a000200

[16] R. Hogan and N. Henley. A test of the empathy-effective communication hypothesis (research report 84). *Baltimore: Johns Hopkins University, Center forSocial Organization of Schools*, 1970.

[17] Y.-L. Hsu, J.-S. Wang, W.-C. Chiang, and C.-H. Hung. Automatic ecg-based emotion recognition in music listening. *IEEE Transactions on Affective Computing*, 11(1):85–99, 2017.

[18] F. Liu, M. Esparza, M. Pavlovskaia, G. Kaufman, L. Dabbish, and A. Monroy-Hernández. Animo: Sharing biosignals on a smartwatch for lightweight social connection. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 3(1):1–19, 2019.

[19] F. Liu, G. Kaufman, and L. Dabbish. The effect of expressive biosignals on empathy and closeness for a stigmatized group member. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–17, 2019.

[20] M. V. Lombardo, J. L. Barnes, S. J. Wheelwright, and S. Baron-Cohen. Self-referential cognition and empathy in autism. *PloS one*, 2(9):e883, 2007.

[21] D. Makowski, T. Pham, Z. J. Lau, J. C. Brammer, F. Lespinasse, H. Pham, C. Schölzel, and S. A. Chen. Neurokit2: A python toolbox for neurophysiological signal processing. *Behavior Research Methods*, pp. 1–8, 2021.

[22] C. R. Mehta and N. R. Patel. Ibm spss exact tests. *Armonk, NY: IBM Corporation*, 2011.

[23] K. NAz and H. Epps. Relationship between color and emotion: A study of college students. *College Student J*, 38(3):396, 2004.

[24] J. Pan and W. J. Tompkins. A real-time qrs detection algorithm. *IEEE Transactions on Biomedical Engineering*, BME-32(3):230–236, 1985. doi: 10.1109/TBME.1985.325532

[25] R. Plutchik. *The emotions*. University Press of America, 1991.

[26] G. Ramsberger and B. Rende. Measuring transactional success in the conversation of people with aphasia. *Aphasiology*, 16(3):337–353, 2002.

[27] R. L. E. P. Reniers, R. Corcoran, R. Drake, N. M. Shryane, and B. A. Völlm. The qcae: A questionnaire of cognitive and affective empathy. *Journal of Personality Assessment*, 93(1):84–95, 2011. PMID: 21184334. doi: 10.1080/00223891.2010.528484

[28] R. Rivu, Y. Abdrabou, K. Pfeuffer, A. Esteves, S. Meitner, and F. Alt. Stare: Gaze-assisted face-to-face communication in augmented reality. In *ACM Symposium on Eye Tracking Research and Applications*, pp. 1–5, 2020.

[29] W. Romine, T. Banerjee, and G. Goodman. Toward sensor-based sleep monitoring with electrodermal activity measures. *Sensors*, 19(6), 2019. doi: 10.3390/s19061417

[30] J. A. Russell. A circumplex model of affect. *Journal of personality and social psychology*, 39(6):1161, 1980.

[31] L. Salahuddin, J. Cho, M. G. Jeong, and D. Kim. Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings. In *2007 29th annual international conference of the ieee engineering in medicine and biology society*, pp. 4656–4659. IEEE, 2007.

[32] A. Schaefer, F. Nils, X. Sanchez, and P. Philippot. Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers. *Cognition and emotion*, 24(7):1153–1172, 2010.

[33] J. Snyder, M. Matthews, J. Chien, P. F. Chang, E. Sun, S. Abdullah, and G. Gay. Moodlight: Exploring personal and social implications of ambient display of biosensor data. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pp. 143–153, 2015.

[34] M. Soleymani, M. Pantic, and T. Pun. Multimodal emotion recognition in response to videos. *IEEE Transactions on Affective Computing*, 3(2):211–223, 2012. doi: 10.1109/T-AFFC.2011.37

[35] C. S. S. Tan, J. Schöning, K. Luyten, and K. Coninx. Investigating the effects of using biofeedback as visual stress indicator during video-mediated collaboration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 71–80, 2014.

[36] Z. Zhang, X. Wang, P. Li, X. Chen, and L. Shao. Research on emotion recognition based on ecg signal. In *Journal of Physics: Conference Series*, vol. 1678, p. 012091. IOP Publishing, 2020.