Emotional Virtual Reality Stroop Task

1st Deniz Mevlevioğlu
School of Information Technology and Computer Science
University College Cork
Cork, Ireland
0000-0003-2776-1304

2nd Sabin Tabirca
School of Information Technology and Computer Science
University College Cork
Cork, Ireland
0000-0003-0743-8227

3rd David Murphy
School of Information Technology and Computer Science
University College Cork
Cork, Ireland
0000-0002-9685-8292

***Abstract*—Anxiety classification in Virtual Reality using biosensors is an ongoing challenge for the area. This paper presents the development and validation of a Virtual Reality version of the well-established cognitive task; the Emotional Stroop Colour-Word Task. This environment aims to be able to differentiate between varying states of anxiety of the user. The task is split into three "stages" to correspond with three different states; no anxiety, low anxiety and anxiety. To validate whether the environment can accurately separate between these states, we have used a repeated measures design comparing each stage based on the mean (1) average electrodermal activity, (2) average self-reported measure of anxiety, and (3) average performance of the users in the stage. Repeated measures ANOVA results show that there is a significant difference between stages in electrodermal activity and self-reported anxiety, but not in response time. This signifies that it might be possible to record different anxiety states using the application. Further research must be conducted to test this hypothesis.**

***Index Terms*—Stroop, VR, biosensors, anxiety, GSR, EEG, PPG**

I. Introduction

Recognition of a user's emotional state within Virtual Reality (VR) enables the possibility of receiving objective feedback from the user and it can be useful for various applications [1]. The collective term for research in computers and user emotion is called "affective computing" and it is used in many different areas of VR research such as gaming, medicine and psychology [2].

Anxiety is the feeling of discomfort that a person can feel regardless of the presence of a stressing stimulus [3]. Although it can be useful at times, persistent activation can be damaging to a person's health [4].

Anxiety recognition in Virtual Reality is generally done using physiological measures such as skin conductivity level (SCL), and heart rate (HR); and behavioural measures such as respiration and head movement [5], [6]. Generally, the anxiety within VR is measured by getting the arousal and valence of the user's response; high anxiety is defined as having low valence and high arousal [7]. As such, it is a complicated measure to derive from the user.

Science Foundation Ireland

The challenge of measuring anxiety accurately in VR comes from multiple reasons. Firstly, it is difficult to accurately define and assess what anxiety is, as this might differ from person to person [8]. Secondly, the most established methods of measuring anxiety using physiological sensors require the person being assessed to have limited mobility [5]. Thus, they can prove difficult to use within VR [7].

This paper aims to develop and investigate a new method for measuring ground truth in anxiety by using an adapted version of the well-established Stroop Colour-Word Task (SCWT).

A. Anxiety classification and Ground Truth

The ground truth refers to the reference point for the anxiety of the user. Establishing anxiety ground truth is necessary to be able to derive meaning from anxiety classification results. It is common to use self-report measures such as state-trait anxiety inventory (STAI) [9] or subjective units of distress scale (SUDS) [10] for the establishment of ground truth; however, self-report measures come with several biases caused by personal differences and the difficulty of being able to tell one's level of anxiety [11]. This is why other methods have become popular, such as using established cognitive tasks, some examples being card sort and SCWT [5].

B. Stroop Colour-Word Task

Traditional (SCWT) [12] is a task that was made for assessing how well a person's brain performs in the existence of conflicting information. The test involves identifying and matching the colour of the words shown where the meaning of the word either matches the colour (congruent or non-interference) or conflicts with the colour (incongruent or interference) [13]. Based on prior research, most people tend to perform worse in the existence of conflicting information [14]. The time difference in processing conflicting information and non-conflicting information is called the Stroop effect (SE).

Despite being made to assess cognitive ability, researchers discovered that it is a good way to measure anxiety and stress [15]. Being such a well-established cognitive task, a lot of research has been conducted around it, making different versions and researching different conditions [16].

There are many applications of this test, in-vivo, computerised and lately, in VR [16], [17]. Recent research shows that VR applications of the test show similar results compared to the in-vivo equivalent [18], [19].

C. Emotional Stroop Colour-Word Task

One of the variants of SCWT examines the effect of emotional valence and arousal of word use on reaction time [20]. The so-called emotional Stroop task, sometimes called eStroop, involves words from affective norms for English words (ANEW) [21]. As opposed to the traditional Stroop task, the emotional Stroop task examines the interference of emotionally loaded words on the speed of processing rather than the interference of conflicting information. As such, this effect is named the emotional Stroop effect (ESE). ESE typically is not as pronounced as SE, where there is a cognitive conflict [22]. However, it is effective in creating emotional arousal [23]. There is not yet any research on VR validity and interaction of eStroop.

D. Related Work

There are several VR applications that show SCWT performs similarly in VR as it does in vivo [17]. As well as testing the effect in a different environment, some of these applications add in elements such as distractions and stressful environments to test different scenarios and how this interacts with the SE [24].

One of the most well-known examples of VR applications of SCWT is that created by Parsons et al. in 2011 [24]. In this experience, a similar effect to SE was achieved by experiencing the task in VR [18]. Furthermore, a more complex model was developed in the experiment, involving higher and lower threat environments and seeing the interaction between threat level and the Stroop effect [24]. This application is quite different from traditional applications of Stroop due to the addition of environmental threats.

Another application of the Stroop Task within VR is by Gradl et al. in 2019 called the Stroop Room [19]. This application brings a different interaction by changing an entire room's walls to specific colours and letting the users point at them. They have observed a significant increase in heart rate as well as a significant effect on reaction time [19]. However, it is hard to directly compare the reaction time in this application due to the increased complexity of the interaction.

Some examples of previous studies that investigate the Stroop effect in different formats can be found in table I.

Our work aims to introduce eStroop into VR and examine its effect on the user, focusing on negative arousal.

TABLE I
PREVIOUS WORK ON STROOP AND ESTROOP TASKS

Reference	VR?	Stroop Type	SCL incr. ($\mu\%S$)	SE (ms)
Renaud et al. [25]	NO	Stroop	NS[a]	S[b]
Malhi et al. [23]	NO	eStroop	26.6	1.4
Strauss et al. [26]	NO	eStroop	n/a	-2
Eilola et al. [27]	NO	eStroop	10	38.5
Blanchette et al. [28]	NO	eStroop	5	S[b]
Gradl et al. [19]	YES	Stroop	12	910

[a]Not significant (rates unreported), [b]Significant (rates unreported)

Fig. 1. The waiting room scene, including the interaction platform

II. IMPLEMENTATION

The VR application was developed using the Unity 3D engine [29]. The scene is modelled to look like a waiting room of a clinic to provide a familiar environment to many users. The scene aims to be similar to real-life environments without trying to be photo-realistic to avoid the "uncanny valley" effect. The interaction uses virtual buttons in the environment that can be clicked by moving the controllers on top of them. There is a hovering platform in front of the user that includes a "start/next" button, red, yellow, green and blue colour buttons and an exit button (fig. 1). The words are displayed on a television screen placed several meters away from the user. The user has control of when to start the tutorial and the task itself. The exit button allows the user to leave the experience before its completion.

A. Tutorial

The tutorial serves two purposes. The first one is to familiarise the user with the interaction system and the task by going through every step that needs to be taken to successfully finish the task. The other one is to allow the user to get adjusted to the VR environment and get rid of the novelty effect of first being in an unfamiliar virtual environment. By doing so, the experience aims to minimise the extraneous effect of the environment on the user's emotional state.

When the user is first in the environment, they have time to adjust to the surroundings before the tutorial starts. They are in front of the screen, where the screen reads "Please click the start button to start the tutorial". This allows the user to pay full attention to the tutorial once it starts. When the user clicks the button and starts the tutorial, the system welcomes the user into the environment by artificially generated speech.

The tutorial is multi-modal, the users can read the text on the screen as well as hear it to enhance accessibility and increase understanding. The experience is briefed in the tutorial. This includes how long the experience will take and what the user will need to do. Then the user is invited to match words until they get four of them correct to ensure they understand the task. In this part of the tutorial, they receive feedback on whether they match the word correctly or not. Once the user successfully completes this part, they are given ten examples of colour matching, the exact way it is in the task to make sure they are prepared for it and will not expect feedback. After the no-feedback examples are finished, the users are asked to input their anxiety level on a SUDS scale. As they need to do a "grab" action for this (in this case, by pulling the trigger on a Vive controller), the controller button for the action lights up on the screen and the controller vibrates. After changing their anxiety level on the slider, the user is able to proceed once again by clicking the start/next button. When this step is completed, the tutorial is finished and the user is invited to start the task by clicking the start button.

B. Task

The task is set up to be intuitive and fun. Words change every two seconds, or once an answer is recorded. When the user clicks one of the colour buttons, their reaction time and whether the answer was correct are logged into an external .csv file. The task involves three stages that include words from ANEW, and each stage uses words that receive higher arousal and lower (negative) valence responses. The task aims to measure emotional response and not mental load, therefore, the task does not get increasingly difficult. All the variables stay the same except the words.

C. SUDS

Before the start of the task and after each stage, the user is asked to log their anxiety on a subjective scale. They do so by precisely choosing a number from zero to a hundred on a VR slider, zero being no anxiety and a hundred being the most anxiety possible. Once they do so, their anxiety is logged in the external .csv file.

III. Methods

A. Design

This study examines the differences in means of electrodermal activity, self-report anxiety and reaction time in different stages of the developed task. By examining these differences, we aim to find out whether the emotional Virtual Reality Stroop Task (eVRST) is suitable for differentiating between different levels of anxiety. For physiological measures, SCL is used as a marker for high arousal levels. Higher arousal is interpreted as negative based on self-reported anxiety levels.

For the eStroop task, a blocked design was used to focus on the general effect over a time period [30]. Additionally, a fixed order instead of a mixed order was used, because mixed order does not allow to record anxiety changes due to sustained effects of the emotional words [30].

This study received ethical approval from University College Cork Social Research Ethics Committee with Log number 2021-222.

B. Participants

Twenty-nine volunteers participated in the study. The age distribution of the participants is as follows: eleven participants between the ages eighteen and twenty-four (38%), fifteen participants between the ages twenty-five and thirty-four (51%), two participants between the ages forty-five and fifty-five (7%) and one participant older than 55(3%). Eleven participants were female (38%) and eighteen participants were male (62%). There were seven participants who never used VR (24%), fifteen participants who used it a couple of times (51%), five participants who has one to two years of experience in VR (17%) and two participants with more two to five years experience (7%). Seven participants stated they never experience motion sickness (24%), fourteen participants experienced it rarely (48%) and eight participants sometimes (28%). Four participants were diagnosed with colour blindness (14%) and three participants were diagnosed with dyslexia (10%).

C. Materials

Two questionnaires were administered; one before the experiment and one after. The first questionnaire included demographics and a short version of the State-Trait Anxiety Inventory (STAI-Y1) scale. State-Trait Anxiety Inventory Y1 was developed by Spielberger et al. [9] and modified to be shorter by Morteau and Bekker [31]. In the VR experience, subjective units of distress scale (SUDS) was used to log anxiety as a number between 0-100. The second questionnaire included subjective classification of each stage in eVRST and Likert-style subjective questions.

A shimmer3 GSR+ device was used to measure SCL, and an optical pulse ear-clip add-on was used to measure PPG which was then used to estimate heart rate [32]. Shimmer API was used to process these signals [33]. MyndPlay MyndBand was used to collect electrical brain activity [34]. NeuroSky API was used to process raw EEG signals [35].

The application was developed and tested using the HTC Vive head-mounted display (HMD) and using a personal computer with GeForce GTX Titan X as the graphics card and an Intel i7-5820k processor. The audio was delivered via the HTC Vive earphones.

D. Procedure

Participants first were given an information sheet and consent sheet. Once they were satisfied with the information and signed the consent form, they were given the first questionnaire. These questionnaires were assigned a unique ID to match with their physiological information and second questionnaire. After this point, they were moved to the "play area" for the VR part of the experiment. The experimenter placed the Shimmer GSR+ device on their left wrist, connected

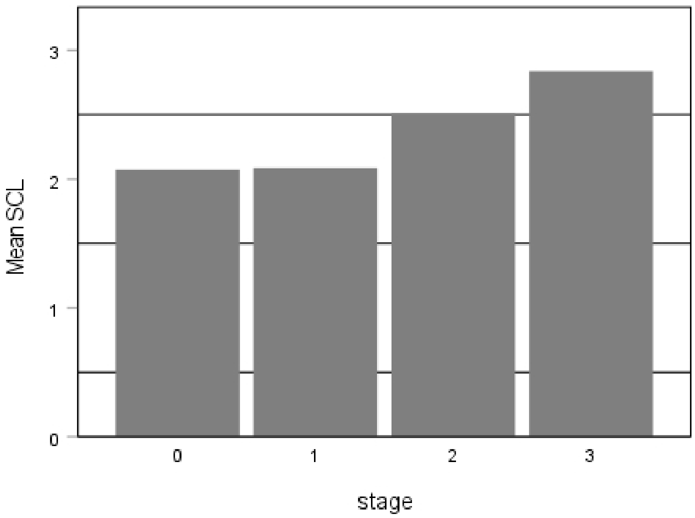

Fig. 2. Mean SCL for baseline (0), stages 1, 2 and 3

Measure	(I) stage	(J) stage	Mean Difference (I-J)	SE	Sig.
SCL	1	2	-.579	.154	.002
		3	-1.000	.214	<.001
	2	1	.579	.154	.002
		3	-.421	.074	<.001
	3	1	1.000	.214	<.001
		3	-.421	.074	<.001

the electrodes to the middle and ring fingers, and connected the optical pulse to their earlobe. After, the HMD was placed on their head until they felt comfortable and didn't have any blur in their vision. Then, the earphones were attached and they were handed the controllers. They were given instructions on how to leave if they feel uncomfortable during the experiment. Then they started the tutorial when they felt ready. They experienced the tutorial, the three stages and breaks in between, in the specified order. After they have completed the experiment, the equipment was removed from them by the experimenter. Once they felt ready, they were given the second questionnaire. After it was all done, they had the opportunity to ask the experimenter questions.

IV. Results

Repeated Measures ANOVA were run to compare the means of user SCL (measured in μS), self-reported anxiety (measured by SUDS) and reaction time (measured in ms) between the three stages of the eVRST. The data was not spherical, so Greenhouse-Geisser correction was applied.

The mean levels of SCL between different stages can be viewed in fig. 2. On average, a user's anxiety increased 31% (SD = .30) from stages 1 to 3. Based on the results, there is a significant difference in SCL between the three levels ($F = 20.182$, $p < .001$). Bonferroni posthoc results show that there is a significant difference between each level (table IV.

The anxiety scores were measured using SUDS. The difference in means of SUDS between stages 3 and 1 was four milliseconds. There was a significant difference in self-reported anxiety between the three levels ($F = 3.877$, $p < .05$). Bonferroni's post hoc results did not reveal any pair-wise differences between the stages.

TABLE II
Questionnaire responses

Question	Mean	N	Std. Deviation
I found the equipment worn during the experiment cumbersome and uncomfortable.	2.72	29	.960
I was able to understand the instructions clearly.	4.59	29	1.053
I was conscious of the devices on me at all times.	3.17	29	1.104
Amount of equipment I was wearing reduced my immersion in the virtual world.	2.45	29	.870
I felt bored during the experiment.	2.07	29	.961
I was comfortable with the devices on me.	3.62	29	.820
I felt motion sick during the experiment.	1.21	28	.418
I found the platform with the buttons easy to use.	4.10	29	.860
I found the task confusing.	1.48	29	.634
The experiment had me engaged at all times.	4.14	28	.756
I wish the tutorial was more thorough.	1.72	29	.649
I found it difficult to understand and carry on the task given to me.	1.28	29	.455

The difference between each stage in reaction time was investigated using the instructions from Ben-Haim et al. published in 2016 [30]. The average reaction time in stage 1 is 980.44ms, the average reaction time in stage 2 is 973.85ms, and the average reaction time in stage 3 is 951.96ms. The performance of the users was enhanced in the emotional stage, with an increase of 29.8 milliseconds in performance. Repeated measures ANOVA results show no significant difference in reaction time between the three levels ($F = 1.996$, $p > .05$).

There were also numerous subjective questions studying the usability of the system on a Likert scale of 1-5. The responses to these questions can be observed in table II.

STAI had a weak linear positive correlation with the difference in SCL between the first and third stages ($r = .457$, $p < .05$), meaning the more anxious the user was initially, the more their SCL increased between the first and third stages. There was no correlation between STAI and either reaction time increase or self-report anxiety increase.

Upon examining the graphs, there were no correlations found between the change in SCL and the change in self-reported anxiety.

ANOVA tests were run to find the main effects of age group and VR exposure on overall reaction time. A main effect of age group was found ($F = 4.70$, $p < .05$), and no effect of VR exposure was found ($F = 1.03$, $p > 0.5$).

Open-ended comments from the participants were grouped into some common categories. In terms of the experience, two participants commented that the platform was too low for them to reach, two participants commented that it was easy to misclick due to the sensitivity of the sensors and the interactions system, and one participant commented that more feedback on their performance and progress was necessary, one participant commented that it was challenging to quantify one's own anxiety and one user commented that they had a lot of fun during the experiment. In terms of the equipment used, two participants commented that the HMD was very heavy, two participants commented that the HMD was not tight enough around their head, two participants commented on not being able to use their prescription glasses with the HMD, two participants commented that the earphones were too loose, one participant commented that the volume was too high and one participant commented that the equipment needs to be reduced for comfort.

V. Discussion

Our paper presented a virtual reality application of the emotional Stroop colour-word task, as well as incorporating a third level into the experience. Due to these differences, it is challenging to make direct comparisons between in-vivo and virtual reality versions. Furthermore, this study focused on the changed arousal throughout stages instead of solely focusing on the difference in performance. In traditional emotional SCWT, performance is a lot more important to measure the impact of a "Stroop effect". However, in our experiment, we aimed to record the trends of arousal change between the stages.

Although not the main objective of the study, ESE was examined. Interestingly, the reaction time difference between the neutral (stage 1) and high emotional stages (stage 3) was the opposite of the literature, with a decrease in reaction time in the third stage. As with block design, emotional stages followed neutral stages to enhance the effect of ESE. Despite possible fatigue effects, the performance of the participants increased in the last stage. This was possibly due to the learning effect and the difficulty of getting used to a task in Virtual Reality as compared to in-vivo and computer applications. This means that in future applications for finding ESE, it is important to prolong the tutorial to eliminate learning effects.

It can be seen from the results that our study is effective in increasing arousal similar to studies cited in the literature, including non-VR studies [19], [23], [25], [27], [28]. Furthermore, this 31% increase in SCL is quite high compared to traditional Stroop results [19], [25]. There is a clear trend of increased arousal through stages, evidenced by the SCL of the users. It is possible to interpret this change of arousal as negatively valenced, based on the self-report measures. The emotional Stroop effect was not observed during the experiment, furthermore, there was a negative Stroop effect observed. This is most likely due to VR applications of Stroop needing a lot longer time for learning effects, and the tutorial within this experiment was not enough to account for the full learning effect.

The results show that this application is effective at showing different levels of anxiety, however, based on self-reported responses, the difference is mild.

Based on the responses to the usability questionnaire, most users found it easy to understand the instructions and carry on the given task using the tutorial. Despite the efforts to decrease the amount of equipment that is used, it still had a slight effect on the immersion in the environment and comfort. Based on some participant feedback, this could partly be attributed to the weight of the HMD used. The experience caused minimal motion sickness. Most users were highly engaged during the experiment and not very bored.

A. Limitations

This study has some limitations based on the sample and the difficulty of dealing with subjective human emotions. First of all, the study was conducted in a university, which possibly biased the sample to be more educated and technically able. 76% of the participants had tried VR before, which is higher than the population average [36]. It is possible that the experience of the sample compared to the general population might have biased the results and prevented the discovery of unique challenges that might be caused by being new to VR.

Another limitation is that to amplify the user's arousal in the stage, the different groups were presented in the same order to every participant. This makes it easier to gather data in different states of emotion, however, makes it harder to derive information about the specific reason for the increased arousal and Stroop effect. Fatigue and prolonged VR exposure or VR sickness might have contributed to these emotions.

B. Future Work

The information received from this pilot will be used to enhance the VR experience. The delay between the questions will be increased while tutorial and break times will be increased. The order of the stages will be randomised to be able to observe the Stroop effect in a less biased matter. Auxiliary experiments will be run to see the effect of ordering on the reaction time.

The data collected from this experiment will be used to develop a machine learning classification model to determine how accurately the stage of the experiment can be determined based on the physiological information recorded such as HR, SCL and electrical brain activity.

VI. Conclusion

The eVRST application demonstrates three levels of user state of high arousal and low valence, as shown by the SCL. It also shows that it acts in similar ways to demonstrate a Stroop effect caused by emotions, however, more research is needed to gather more meaningful data.

Acknowledgment

This publication has emanated from research supported in part by a Grant from Science Foundation Ireland under Grant number 18/CRT/6222.

REFERENCES

[1] G. Riva, F. Mantovani, C. S. Capideville, A. Preziosa, F. Morganti, D. Villani, A. Gaggioli, C. Botella, and M. Alcañiz, "Affective interactions using virtual reality: the link between presence and emotions," *Cyberpsychology & behavior*, vol. 10, no. 1, pp. 45–56, 2007.

[2] J. Diemer, G. W. Alpers, H. M. Peperkorn, Y. Shiban, and A. Mühlberger, "The impact of perception and presence on emotional reactions: a review of research in virtual reality," *Frontiers in psychology*, vol. 6, p. 26, 2015.

[3] C. D. Spielberger, *Anxiety and behavior*. Academic press, 2013.

[4] E. E. Levitt, *The psychology of anxiety*. Routledge, 2015.

[5] G. Giannakakis, D. Grigoriadis, K. Giannakaki, O. Simantiraki, A. Roniotis, and M. Tsiknakis, "Review on psychological stress detection using biosignals," *IEEE Transactions on Affective Computing*, 2019.

[6] D. Kamińska, K. Smółka, and G. Zwoliński, "Detection of mental stress through eeg signal in virtual reality environment," *Electronics (Switzerland)*, vol. 10, no. 22, Nov 2021.

[7] M. Cao, T. Xie, and Z. Chen, "Wearable sensors and equipment in vr games: a review," *Transactions on Edutainment XV*, pp. 3–12, 2019.

[8] C. Costello and A. L. Comrey, "Scales for measuring depression and anxiety," *The Journal of psychology*, vol. 66, no. 2, pp. 303–313, 1967.

[9] C. Spielberger, R. Gorsuch, R. Lushene, P. Vagg, and G. Jacobs, *Manual for the State-Trait Anxiety Inventory (Form Y1 – Y2)*. Consulting Psychologists Press, 01 1983, vol. IV.

[10] J. Wolpe, *The practice of behavior therapy*. Pergamon press, 1990.

[11] I. H. Gotlib and D. B. Cane, "Self-report assessment of depression and anxiety." *Anxiety and depression: Distinctive and overlapping features*, 1989.

[12] J. Stroop, "Stroop color word test," *J Exp Physiol*, vol. 18, pp. 643–62, 1935.

[13] A. R. Jensen and W. D. Rohwer Jr, "The stroop color-word test: a review," *Acta psychologica*, vol. 25, pp. 36–93, 1966.

[14] C. M. MacLeod, "Half a century of research on the stroop effect: an integrative review." *Psychological bulletin*, vol. 109, no. 2, p. 163, 1991.

[15] P. Karthikeyan, M. Murugappan, and S. Yaacob, "Analysis of stroop color word test-based human stress detection using electrocardiography and heart rate variability signals," *Arabian Journal for Science and Engineering*, vol. 39, no. 3, pp. 1835–1847, 2014.

[16] F. Scarpina and S. Tagini, "The stroop color and word test," *Frontiers in psychology*, vol. 8, p. 557, 2017.

[17] D. Wu, C. G. Courtney, B. J. Lance, S. S. Narayanan, M. E. Dawson, K. S. Oie, and T. D. Parsons, "Optimal arousal identification and classification for affective computing using physiological signals: Virtual reality stroop task," *IEEE Transactions on Affective Computing*, vol. 1, no. 2, pp. 109–118, 2010.

[18] C. M. Armstrong, G. M. Reger, J. Edwards, A. A. Rizzo, C. G. Courtney, and T. D. Parsons, "Validity of the virtual reality stroop task (vrst) in active duty military," *Journal of Clinical and Experimental Neuropsychology*, vol. 35, no. 2, pp. 113–123, 2013.

[19] S. Gradl, M. Wirth, N. Mächtlinger, R. Poguntke, A. Wonner, N. Rohleder, and B. M. Eskofier, "The stroop room: A virtual reality-enhanced stroop test," in *25th ACM Symposium on Virtual Reality Software and Technology*, 2019, pp. 1–12.

[20] J. M. G. Williams, A. Mathews, and C. MacLeod, "The emotional stroop task and psychopathology." *Psychological bulletin*, vol. 120, no. 1, p. 3, 1996.

[21] M. M. Bradley and P. J. Lang, "Affective norms for english words (anew): Instruction manual and affective ratings," Technical report C-1, the center for research in psychophysiology . . . , Tech. Rep., 1999.

[22] M. Kindt, D. Bierman, and J. F. Brosschot, "Stroop versus stroop: Comparison of a card format and a single-trial format of the standard color-word stroop task and the emotional stroop task," *Personality and individual Differences*, vol. 21, no. 5, pp. 653–661, 1996.

[23] G. S. Malhi, J. Lagopoulos, P. S. Sachdev, B. Ivanovski, and R. Shnier, "An emotional stroop functional mri study of euthymic bipolar disorder," *Bipolar disorders*, vol. 7, pp. 58–69, 2005.

[24] T. D. Parsons, C. G. Courtney, B. Arizmendi, and M. Dawson, "Virtual reality stroop task for neurocognitive assessment," in *Medicine Meets Virtual Reality 18*. IOS Press, 2011, pp. 433–439.

[25] P. Renaud and J.-P. Blondin, "The stress of stroop performance: Physiological and emotional responses to color–word interference, task pacing, and pacing speed," *International Journal of Psychophysiology*, vol. 27, no. 2, pp. 87–97, 1997.

[26] G. P. Strauss, D. N. Allen, M. L. Jorgensen, and S. L. Cramer, "Test-retest reliability of standard and emotional stroop tasks: An investigation of color-word and picture-word versions," *Assessment*, vol. 12, no. 3, pp. 330–337, 2005, pMID: 16123253. [Online]. Available: https://doi.org/10.1177/1073191105276375

[27] T. M. Eilola and J. Havelka, "Behavioural and physiological responses to the emotional and taboo stroop tasks in native and non-native speakers of english," *International Journal of Bilingualism*, vol. 15, no. 3, pp. 353–369, 2011.

[28] I. Blanchette and A. Richards, "Is emotional stroop interference linked to affective responses? evidence from skin conductance and facial electromyography." *Emotion*, vol. 13, no. 1, p. 129, 2013.

[29] Unity. (2022) Unity real-time development platform. https://unity.com/. Accessed: 2022-12-06.

[30] M. S. Ben-Haim, P. Williams, Z. Howard, Y. Mama, A. Eidels, and D. Algom, "The emotional stroop task: assessing cognitive performance under exposure to emotional content," *JoVE (Journal of Visualized Experiments)*, no. 112, p. e53720, 2016.

[31] T. M. Marteau and H. Bekker, "The development of a six-item short-form of the state scale of the spielberger state—trait anxiety inventory (stai)," *British Journal of Clinical Psychology*, vol. 31, no. 3, pp. 301–306, 1992.

[32] S. Sensing. (2022) Shimmer3 gsr+ unit - shimmer wearable sensor technology. https://shimmersensing.com/product/shimmer3-gsr-unit/. Accessed: 2022-12-06.

[33] A. Burns, E. P. Doheny, B. R. Greene, T. Foran, D. Leahy, K. O'Donovan, and M. J. McGrath, "Shimmer™: an extensible platform for physiological signal capture," in *2010 annual international conference of the IEEE engineering in medicine and biology*. IEEE, 2010, pp. 3759–3762.

[34] Neurosky. (2022) Myndband eeg brainwave headset. https://store.myndplay.com/products.php?prod=48. Accessed: 2022-12-06.

[35] W. Sałabun, "Processing and spectral analysis of the raw eeg signal from the mindwave," *Przeglad Elektrotechniczny*, vol. 90, no. 2, pp. 169–174, 2014.

[36] A. Insider. (2020) New report: 19% of u.s. adults have tried vr. https://arinsider.co/2020/05/20/new-report-19-of-u-s-adults-have-tried-vr/. Accessed: 2022-20-12.