Evaluating a Brain-Computer Interface to Categorise Human Emotional Response

Katie Crowley, Aidan Sliney, Ian Pitt, Dave Murphy
Department of Computer Science
University College Cork
Cork, Ireland
ekc1@cs.ucc.ie, ajs2@cs.ucc.ie, i.pitt@cs.ucc.ie, d.murphy@cs.ucc.ie

***Abstract*—The aim of this study is to evaluate NeuroSky's Mindset headset as a minimally invasive method of measuring the attention and meditation levels of a subject. Two psychologically-based tests were conducted to assess the suitability of the headset to measure and categorise a user's level of attention and meditation.**

Keywords-Emotion; Assessment; NeuroSky; Brain-Computer Interface; Stroop; Stress; Attention; Serious Games; Computer Assisted Learning

I. INTRODUCTION

A problem arises in using serious games to assess the user's emotional response. A human assessor could answer questions that other forms of assessment cannot. Was the student nervous? Did the student look stressed? Did they show confidence? This information can be valuable for formative assessment. 'Even if the primary assessment purpose does not include emotional response, as secondary data it can be of considerable value [1].

One such emotion that could prove valuable for assessment is calmness. Our use of the term 'calm' describes how relaxed and composed a user is. Assessor's observation will always lead the way in emotion assessment. Our aim for this study is to assess the suitability of NeuroSky's Mindset to capture users' meditation and attention levels in an instance where it is not possible to have an observer present. The advantage and novelty of using NeuroSky's Mindset Brain-Computer Interface (BCI) to gain emotion information lies in the fact that learners can assess themselves at anytime without the need for assessors.

A. NeuroSky

A BCI is a communication system in which messages or commands that an individual sends to the external world do not pass through the brain's normal output pathways of peripheral nerves and muscles.

NeuroSky have developed a non-invasive, dry, bio sensor to read electrical activity in the brain to determine states of attention and relaxation. NeuroSky is a low-cost, easy to use Electro Encephalogram (EEC) developed for leisure. It captures neural activity using three dry electrodes (locations: beneath the ears and the forehead) and decodes them by applying algorithms. NeuroSky provides information on a user's Delta, Theta, Alpha, Beta, and Gamma brainwave band power levels [2].

B. Capturing Emotions for Assessment with NeuroSky

NeuroSky reads attention and meditation levels directly based on the user's brain activity, and outputs a number per second in a scale from 0 to 100 for each emotion captured. By using these figures we plan to group users into different attention categories. We looked at the possibility of time stamping any precise moments when the user makes a mistake. Our intention was to find these errors by recording any times of erratic output from the Mindset.

II. METHODS

Two psychological tests were completed to evaluate the effectiveness of using NeuroSky's Mindset for Assessment; a Stroop test and a test implementing the Towers of Hanoi. The tests were completed in a blacked-out room. The tests were conducted on a Dell T3400 with a dual monitor setup. Recordings were undertaken using a Shure SM-57 microphone. 20 users participated in the study (20 completed the Stroop and 17 undertook The Towers of Hanoi), with each test averaging 20 minutes. The BCI used to record the meditation and attention levels was NeuroSky's Mindset. A backup headset was used to check for consistent calibration.

A. The Stroop Test

The Stroop Colour-Word Interference Test (Stroop, 1935 [4]) is a well-known psychological test of selective attention, cognitive flexibility and processing speed, often utilized as a psychological or cognitive stressor [5]. The test exploits the fact that, for experienced readers, the reading of a word has become an automatism. Individuals can read words much faster than they can identify colours.

The assessment requires the subject to name the colour that is displayed and not the word i.e. to identify the colour stimulus and not the word stimulus. The automatic or most natural response is to determine the semantic meaning of the word i.e. reading the word 'red' the subject thinks of the colour 'red'. The cognitive mechanism involved in this task is directed attention, the subject must manage their attention by inhibiting one response in order to say or do something else. Rothkrantz et al. [6] have used an adaptation of the Stroop test to simulate stress in speech as part of their studies on Voice Stress Analysis.

The Rothkrantz variation is computer-based and incorporates a gradual increase of the level of difficulty in the test over a five-minute period. The difficulty of the task is increased as the time between the appearances of the colours is shortened every minute by half a second, decreasing from two and a half second intervals during minute one, to intervals of half a second in minute five.

978-0-7695-4055-9/10 $26.00 © 2010 IEEE
DOI 10.1109/ICALT.2010.81

Subjects are required to speak the name of the colour that is displayed to them, and this data is recorded for later analysis.

In this study, the first minute is considered to represent normal conditions. When the acoustic analysis of the data in minute one was compared to the data in minute five, an increase in fundamental frequency was observed, as the subject becomes increasingly stressed. Rothkrantz et al suggest that the acoustic changes measured during minute five of this test demonstrate induced stress caused by the Stroop test.

To examine the application of the NeuroSky headset as a means of measuring the meditation and attention level of a subject, we repeated Rothkrantz's adaptation of the Stroop test. Prior to the test, subjects were asked to assess their daily stress level on a scale of 1-10 (1 indicating very relaxed, and 10 indicating very stressed).

The attention and meditation levels of the subject were recorded using the NeuroSky headset. The tester noted the accuracy of each response from the subject for post-test analysis.

On completion, subjects were asked to indicate how stressed they felt **during** the Stroop test, using a 1-10 scale. The majority of the subjects reported an increase in perceived stress level during the stress test (in comparison with the initial questionnaire).

The meditation scale measured by the NeuroSky headset reflects the level of relaxation of the subject, and consequently we expected this measure to decrease as the Stroop test induced stress for the subject. The level of the subject's attention (as measured by the NeuroSky headset) during the test was also a point of interest. According to the design of the Stroop test, the subject should be most relaxed in minute one of the test and least relaxed i.e. most stressed, during minute five of the test. Using the test response accuracy data collected by the tester, each incorrect response during the Stroop test was time-stamped. The attention and meditation timelines were then compared with the error timeline to examine any relationship between changes in meditation or attention and erroneous responses.

No overlap was found between the exact moment of error and the attention or meditation level of the subject. Consequently, we were unable to correlate precise moments of stress or lack of attention from the headset data with erroneous responses from the Stroop test. It is assumed that the stress response to making an error in the test is not instantaneous, i.e. the headset does not register the error at the precise moment the subject makes the mistake. While it was not possible to identify isolated instances of error the overall attention and meditation plots revealed a clear pattern for the majority of subjects, often highlighting an increase in stress level during minute five of the test.

Fig. 1 is a typical example of a plot from the headset data. The line plotted in this graph represents meditation (the lower the mark the more stress in the subject). The three significant minimums in the line plot roughly correspond to the onset of a new minute in the Stroop test, i.e. an increase in the speed of the test and consequently an initial increase in the level of stress of the subject. While we were unable to pinpoint precise moments of error using NeuroSky, the plot in Fig. 1 clearly demonstrates NeuroSky's ability to monitor a subject's level of meditation or stress over a given period of time.

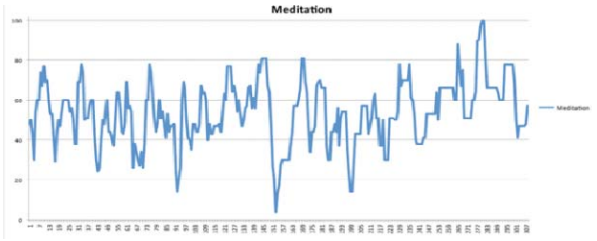

Figure 1: Plot representing meditation level of subject – clearly indicating speed changes of the test

B. The Towers of Hanoi

The Towers of Hanoi is a mathematical game or puzzle. It consists of three rods, and a number of disks of different sizes which can slide onto any rod. The Tower of Hanoi is frequently used in psychological research on problem solving [3].

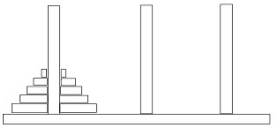

Figure 2: Five Disc Towers Of Hanoi

The objective of the puzzle is to move the entire stack to another rod, obeying basic rules: only move one disk per move, only move the top disk on a stack, with the restriction that you can't move a larger disk on top of a smaller. The participants used a computer version of The Towers of Hanoi written in JavaScript.

The Towers of Hanoi seems impossible to many novices, yet is solvable with a simple algorithm. Once the user realises how to break the puzzle into smaller steps it becomes a simple matter of repetition. It is for these characteristics the Towers of Hanoi was chosen. The puzzle seeming impossible at the start puts the users under stress. Once they have understood the puzzle and are simply repeating steps they should be calm. Each user attempts the puzzle 3 times.

Watching a participant, observers were confident that they could see and record if the subject was stressed when completing the puzzle. It is this record that we hoped would match the Mindset.

To further understand the figures returned by the Mindset we also recorded users undertaking the task of filling out a usability form. With this information we got the basic levels of attention and meditation for a routine task. When a user completed a routine stress-free task the measurements returned by the Mindset never dropped below 40.

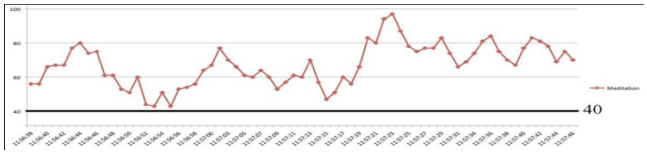

Figure 3: The red line shows the figures returned by the Mindset when a user completes a routine stressless task

Fig. 3 was used as a baseline to create the four categories. Using the figures returned from the headset while completing The Towers of Hanoi, we are able to place each subject into a category.

The participant's categorisation was dependant on the percentage of time the user's levels of meditation dropped below 40.

X = Meditation time below 40 / overall time * 100

If (X >= 25)	Poor Attention - Stressed
If (10<X < 25)	Normal Attention – Calm
If (X <=10)	High Attention – Very Calm

An example of using (1) would be if a participant's meditation level was below 40 for an accumulated time of 51 seconds on a 300 second task then their mediation level (X) would be at 17%. This would categorise the participant as 'Stressed'.

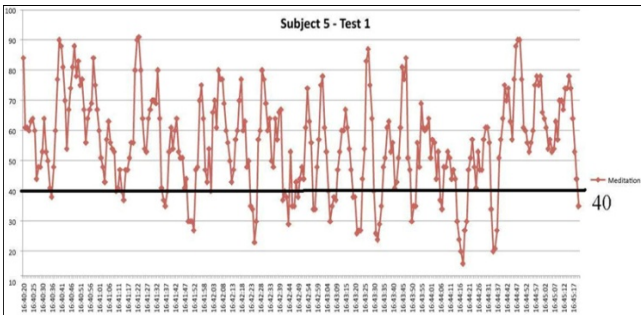

Figure 4: figures are returned by the Mindset when participant 5 completes The Towers of Hanoi. The figures were under 40 for 51 seconds

The Tower of Hanoi was completed 41 times (m = 41). 10 tests were not completed as the user was unable to finish the puzzle. When comparing the Mindset categorisation of m to the observer categorisation 78.04% (32/41) are placed in the same category.

TABLE I. TOWER OF HANOI RESULTS

	First Hanoi		Second Hanoi		Third Hanoi	
#	*Visual*	*Headset*	*Visual*	*Headset*	*Visual*	*Headset*
2	Stressed	23.00%	Stressed	16.40%	Calm	3.80%
4	V.Stressed	28.00%	Stressed	37.00%	Calm	24.00%
5	V.Stressed	17.00%	Stressed	11.00%	Calm	5.70%
6	Stressed	5.00%	Calm	3.00%	Calm	0.00%
9	Stressed	6.70%	Calm	3.20%	Calm	0.00%
10	Calm	4.20%	Calm	2.70%	Calm	4.10%
11	~~DNF~~	-	~~DNF~~	-	~~DNF~~	-
12	~~DNF~~	-	~~DNF~~	-	~~DNF~~	-
13	Stressed	24.50%	Stressed	20.00%	Calm	0.00%
15	Stressed	6.90%	Calm	0.00%	Calm	0.00%
16	V.Stressed	33.20%	Stressed	23.80%	Calm	0.00%
17	V.Stessed	5.00%	Stressed	11.60%	Stressed	5.70%
19	~~DNF~~	-	~~DNF~~	-	~~DNF~~	-
20	Stressed	20.00%	Calm	8.00%	Calm	16.00%
21	~~DNF~~	-	Calm	4.00%	Calm	3.00%
22	Calm	6.00%	Calm	0.00%	Calm	0.00%
24	Calm	0.00%	Calm	3.70%	Calm	0.00%

In table 1 are the 17 participants of The Towers of Hanoi test. The 10 results where the observer categorization is different to that of the Mindset are shown with a border.

The difference in stress levels between the first test to third test is clearly notable. In the first test the users are typically stressed or very stressed whereas in the third test they are typically calm. This is a trait we anticipated to observe with The Towers of Hanoi. Two users (Subject 22 and Subject 23) had prior experience of the Towers of Hanoi before the test. Both users understood the puzzle and knew a solution. With this in mind it would be expected that the stress levels would be low even on their first attempts. Both methods of categorisations demonstrate this fact by classifying these users as calm in each of the three tests.

III. RESULTS /DISCUSSION

The data collected from the headset during the Stroop and Towers of Hanoi tests clearly demonstrate NeuroSky's suitability as a minimally invasive means of measuring the attention and meditation level of a subject. Both tests in this study show that the attention and meditation datasets outputted by the headset clearly indicate when a subject undergoes a change in these emotions.

While it was not possible to correlate isolated moments of human error with a precise change in either attention or meditation, the overall trend of the subject's emotional pattern was clearly visible.

IV. CONCLUSION / FUTURE WORK

Using a well-recognised psychological examination such as the Stroop test, which is often used to induce stress, we were able to assess the suitability of the NeuroSky headset for measuring the meditation and attention level of an individual. A successful result in this study allowed us to apply the same principles to a new method of testing, The Towers of Hanoi.

While this experiment has affirmed the suitability of the Stroop test to vary a subject's stress level we plan to run a further study using a test design that will increase stress in the majority of subjects. The headset can now be used as a monitor for emotional response (attention/mediation) in a test environment.

V. REFERENCES

[1] A. Sliney, D. Murphy, and J. O'Mullane, "Secondary Assessment Data within Serious Games," Serious Games on the Move, 2009, pp. 225-233.

[2] Serious Games Institute - White Paper - Attention.pdf.

[3] A. Hinz and A. Hinz, "The Tower of Hanoi," Enseignement Mathematique, vol. 35, 1989, pp. 239–321.

[4] J. R Stroop, "Studies of Interference in Serial Verbal Reactions," Journal of Experimental Psychology, vol. 28, 1935., pp 643-662.

[5] S. F. Taylor, S Kornblum, E Lauber, S Minoshima, R Koeppe, "Isolation of Specific Interference Processing in the Stroop Task: PET Activation Studies," NeuroImage Journal, vol. 6, 1997, pp 81-92.

[6] L.J.M Rothkrantz, P Wiggers, J.W.A van Wees, R.J van Vark, "Voice Stress Analysis"