5-2018

Stress Levels of CS1 Students During Programming- Measurement and a Cause and Effect Analysis

Gopi Satya Sainadh Raju Sarikonda

STRESS LEVELS OF CSU STUDENTS DURING PROGRAMMING- MEASURMENT AND A CAUSE AND EFFECT ANALYSIS.

Gopi Satya Sainadh Raju Sarikonda

2018

Columbus State University

TSYS School of Computer Science

The Graduate Program in Applied Computer Science

Stress Levels of CS1 Students During Programming- Measurement and a Cause and Effect Analysis.

A Thesis in

Applied Computer Science

By

Gopi Satya Sainadh Raju Sarikonda

Submitted in Partial Fulfillment
Of the Requirements
For the Degree of

Master of Science

May 2018

I have submitted this thesis in partial fulfillment of the requirements for the degree of Master of Science.

Date	Gopi Satya Sainadh Raju Sarikonda

We approve of the thesis of Gopi Satya Sainadh Raju Sarikonda as presented here.

Date	Lydia Ray, Ph.D. Associate Professor, Thesis Advisor
Date	Alfredo Perez, Ph. D. Assistant Professor
Date	Shamim Khan, Ph.D. Professor
Date	Yesem Peker, Ph.D. Associate Professor
Date	Wayne Summers, Ph.D. Distinguished Chairperson Professor of Computer Science

ABSTRACT:

Stress is a kind of feeling we experience when we are under pressure. Stress is the word that we use when we feel that we are overloaded mentally in our thoughts and wonder whether we can cope with those placed upon us. The effects of stress are different for different people when we take their age, profession, gender and other aspects into consideration. Many studies show that stress in a learning environment impacts learning negatively.

In this thesis, the role of stress on students in an introductory programming course (CS1) at CSU has been explored. Introductory programming course has a high attrition rate nationwide. A project was developed to investigate whether students feel stress during programming . This thesis also gives whether that stress is correlated with gender, prior exposure to programming, and math background.

A device called Neurosky is used to perform a low-cost EEG (electroencephalogram) by using inexpensive dry sensors to record the brain wave values of the participants in two different tasks. The test results have shown that the students feel stress during programming. A basic metric is used to determine the stress levels of the participants. Later four hypotheses are proposed by observing the comparisons and tested using statistical hypothesis testing to correlate stress with the participant's gender, major, prior programming experience and with their ACT/SAT scores. The test results proved that all these hypotheses are valid with the collected data.

Table of Contents:

List of Figures:

List of Tables:

Acknowledgment:

My advisor and committee members: Dr. Lydia Ray, Dr. Alfredo Perez, Dr. Shamim Khan, Dr. Yesem Peker.

My family.

Chapter. 1 Introduction

1.1Problem Statement

Stress is one of the major problems for college students throughout the world. Experiencing stress leads to many issues related to emotional and physical health. One of the most frightening consequences of college students stress is committing suicide. According to the survey of American Psychological Association, the percentage of the college students seeking counseling who reported depression was 49 percent, who felt stress was 45 percent, family issues-31 percent, academic performance-28 percent and relationship problems-27 percent. In fact, the 2015 American College Health Association National College Health Assessment found that 85.6 percent of students felt overthrown/overwhelmed in the past year. According to the 2010 American College Health Association National College Health Assessment, more than 25% of students told that the reason for their lower grades or inability to finish a course was stress. Stress decreases the sleep-quality, causes anxiety, anger, frustration, etc., [4]

Stress causes symptoms by affecting our mental behavior, Some of them are listed below,

- Agitation, frustration, depression.
- Feeling overthrown.
- Difficulty in becoming calm, not being able to relax.
- Low self-esteem, lonely and avoiding others.

These problems are very likely to impact a person's cognitive ability.

Stress in computer science students: In 2005, around 55,000 students in the US completed their under-graduate studies in computer sciences or related fields. That figure became lower year by year, and the percentage of the students earning their degrees in computer science decreased from 3.67 (of all majors) to 3.14 in the past decade.

Introductory programming course students suffer from a wide range of difficulties and lack of confidence. Programming courses are generally regarded as difficult and often have the highest dropout rates.

Many case studies show that the students feel depressed, lack of confidence, sleeplessness, angry (which are the most common symptoms of stress) while studying computer science courses.

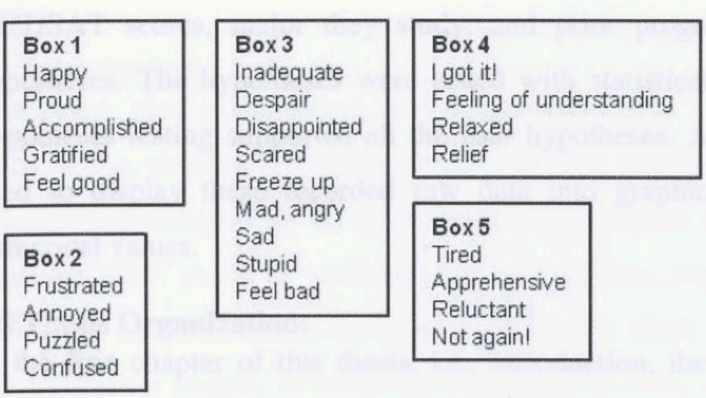

Figure 1: Students emotions while programming in Kinnunen, Beth Simon study.

In Kinnunen and Simon study authors separated the different types of emotions shown by the students in a computer science course [23]. Figure 1 shows that most of them were the symptoms of stress. However, to the best of our knowledge, no formal study has been done to measure and investigate stress levels of CS1 students.

We designed a project to explore stress in the minds of CS1 students while they write a program using a device called neurosky. The objective of this thesis is to collect data on the stress level of students of introductory computer programming course (CS1) while they are coding. The data collected was used to evaluate whether stress level varies significantly when a CS1 student writes a computer program. A significant rise in the stress level of a CS1 student during coding was found. Further, the data was analyzed to see if there is any correlation between stress level during coding and the gender, major field of student, prior exposure to programming and math background of an average student. The stress level of a student was measured by using a device called Neurosky which measures brain waves and calculates the stress level from that data.

1.4 Our contribution:

An email was sent to every student requesting them to participate in this experiment. Once the students were available, they were asked to wear a device called neurosky. The device is used to detect and measure the brain wave forms of the participants. Our contribution is to measure and analyze the brain waves of the participants. The beta waves increase when the person feels stress whereas the alpha and theta waves decrease [26]. Observing this fact from the recorded brain wave values, we proposed a metric beta/(alpha*theta) to find the stress levels of the participants from the recorded brain wave form values, which increases with the presence of stress in the person. Later these stress values were used to find if stress has any correlation with their gender, ACT/SAT scores, major they study, and prior programming experience by assuming four hypotheses. The hypotheses were tested with statistical hypothesis testing with z values. The hypotheses testing supported all the four hypotheses. An open source android application was used to display these recorded raw data into graphical forms which further save them as numerical values.

1.5 Thesis Organization:

In the first chapter of this thesis, i.e., introduction, there are problem statement, the stress of computer science students, the goal of the project, and methodology. The second part, i.e., the second chapter in this thesis provides the background information, affective computing and its history, related work and also provides an overview on some of the developed affective computing systems and stress monitoring systems and the relationship between stress and affective computing. Chapter three explains about the device used to perform EEG and the mobile application used to display these recorded values and EEG and experiment design, chapter four discusses the analyses used with the metric contribution and the results obtained from data collected through the system and the hypotheses assumed with their testing using z-test and offers recommendations for future work.

Chapter 2: Background

2.1 Introduction

Computers play a significant role in our daily life. Today, computers can perform complex calculations, weather prediction, communication, analyses etc. Even intelligence system such as robotics technology help us to do many daily routine tasks. But still today we don't consider them as "alive" entity. Even though today computers can do so many things simultaneously they still have some limitations. Emotion is the feature which makes humans different than computers [1]. Efficiency enables us to see the computers as tools and not as an obstacle in completing our tasks and avoids frustration at use [10]. Emotions affect learning and achievement, mediated by attention, self-regulation, and motivation. They direct the person toward or away from learning the subject. Positive emotions are considered as a pleasant state of emotion whereas negative emotions are considered as an unpleasant state of emotions [2]. Affective computing is a study and development of systems and devices which can identify, interpret, process, and simulate human effects. Affective Computing is a rapidly developing area within computer science. There is now a great transition to make technologies, such as robotic systems, game characters, etc. Affective computing builds an "affect model" based on the various sensors which capture information. Hence it is possible to build a customized computing system with the capability of perception; interpretation of human's emotion as well as giving us intelligent responses. There are varieties of goals in Affective Computing. One is to sense and respond to human emotion. For example, if a person is interacting with technology and he is frustrated or confused, the technology needs to be able to respond differently to that person. Every person will respond differently to the technology hence technology should also respond to it accordingly [3].

2.2 Affective computing:

Affective Computing deals with the devices and system design which can acknowledge, interpret and process emotions. Affective Computing concerns multi-skilled background knowledge, such as psychology and computer sciences. With the aim of improving the quality of human-computer interaction, and also to enhance the intelligence of the computer, affective computing builds an "affect model" based on the various sensors-captured information. Affective Computing can be seen as a computing system which can detect, respond and imitate human emotional states [5]. Humans interact through speech, and their facial expressions. Systems that can understand human emotions and display results through multimodal transmitting channels could be useful and valuable. The ultimate goal is to build systems that can respond to human emotions. Even though the machine does not feel emotion, it must be able to express and interpret those emotions to interact better with us humans [3].

2.2.1 History of Affective Computing

A computer can express emotion without having emotions, like the way actors can express emotions that they do not have. The basic requirement for a computer to express emotions is to have communication media such as voice, image, and an ability to communicate affection over those channels. Affective computing is a field of research dealing with emotions and computers [14]. The original version for Affective Computing was designed to identify what users are experiencing when they are interacting with the computer systems. The emotional responses from these interactions need to be processed, and the resulting model used to modify the interaction. These ideas came out from Artificial Intelligence with an aim to understand people and to bring emotions. The term "Affective Computing" was established only in 1995, with the book “Affective Computing” by Rosalind Picard. Affective Computing is a field of computing which can detect, respond, and simulate human emotional states [3]. The necessary background for affective computing is the knowledge on emotions and its role in human behavior and cognitive processes. eMuu is a robot developed by Bartneck to work as the interface between the end user and smart home [17]. The user can give instructions to the robot to perform a variety of tasks at home. The results of this eMuu showed that the speech recognition in the robotic conditions was decreased and also there was a considerable difference in recognition accuracy between the participants. Due to the noise of the robot's motors, there was a decline in speech recognition accuracy [18]. Feelix is a robot built by Canamero which is designed from LEGO Mindstorms. This robot has been used as a research platform for human-robotic interactions study. It uses a combination of two distinct emotion representations, discrete basic emotions, and continuous dimensions. Feelix perceives tactile stimuli which provide some information about the environment. Damasio proposes the somatic marker mechanism for modeling emotions. TABASCO model is based on the emotion appraisal theory. OCC theory of emotions is also one of the popular models. Most of these models focused on modeling emotional system as it seems to be in human beings and animals. Using these systems is helpful when improving our knowledge about the nature of emotion [17].

2.3 Stress Monitoring Systems

Affect recognition, and physiological monitoring has shown important results on human stress. For example, several studies have pointed to the presence of a contagion effect where stress can be transmitted from one person to another [13]. A 30-minute session was conducted in a laboratory in which 24 participants performed a variety of computerized tasks consisting of expressive writing, text transcription, and mouse clicking. During the stressful conditions; >79% of the participants showed that there was an increased typing pressure and 75% showed mouse contact [8].

A math problem-solving experiment was carried out in the lab with 75 participants to research how to detect affective states with data mining. To collect emotional data researcher used different data sources such as keyboard interactions, mouse interactions, webcam recording, computer screen recording, Kinect recording and physiological recording. The experiment collected participants' emotional data. The mathematical tasks consisted of 3 series which consists of 6 problems. For each problem, participants had to select one answer from a set of 4 options and fill in the 9-point Self-Assessment Manikin scale to report their valence (i.e., pleasure) and arousal (i.e., activation) state. After each group of tasks, participants had to type their feelings about it. Emotions were collected by giving less time than required to do some tasks or changing their difficulty level [9].

Another study was carried out to show a new way of measuring the emotion via the user's mouse motions combining the advantages of the other methods by being directly connected but not intrusive. A set of software tools for measurement was written, and an experiment for validation was executed. The results of this experiment showed that it is possible to use the computer mouse as a sensor for emotional states [10].

Another stress monitoring system was developed with an aim to design a system for the stress detection and classification related to the student in examination duration. In this, major focus on the input type of voice signal from the speech signal. Also, datasets were collected for the analysis of emotions in real time speech [11].

In today's world, Wearable sensors provide an estimation of the number of steps taken, physical activity levels, activity patterns and it will also show us the results of heart rate and skin conductance. A poor sleep quality group showed longer 3- 6 am screen on duration and shorter 9 am-12 pm duration. The high-stress group and the mentally unhealthy group used SMS less frequently and showed later mean timestamp [16].

Stress recognition using non-invasive technology is another system. An integrated hardware-software setup has been developed to achieve automatic assessment of the effective status of a computer user.

A computer-based "Paced Stroop Test" is designed to act as a stimulus to collect emotional stress. Many signal processing techniques are applied to the signals received to extract the most relevant features in the physiological responses and feed them into learning systems [12].

The Conductor's Jacket is a wearable physiological monitoring system that has been built into the clothing of an orchestral conductor. It aims to provide a testbed for the study of emotional expression as it relates to musical performance [13].

Stress and affective computing:

Stress is one of the most significant affective states that is a measurably important health factor. Research projects in this area are slowly but steadily emerging, such as the Affective Health System which aims to empower mobile users to find patterns of their stress levels throughout their daily lives. The term “Affective Computing” was first introduced by Rosalind Picard in 1995 where she defined it as “the one that relates to, arises from or deliberately influences emotion or other affective phenomena” (Picard R., 1997). Later, Picard introduced the term "affective medicine," highlighting research at the MIT Media lab towards emotionally aware and emotionally responsive computers, specifically targeted at certain categories of emotions closely related to health. The principal aim is to enable computers to recognize some of the most common emotions that people experience while interacting with today's technology, such as frustration, irritation or stress. [7] The computers can afterward facilitate the improvement of human-computer interaction and decrease the level of stress or related negative emotions.

Chapter 3: System and Experiment Description

This section provides descriptions of the device we have used to measure stress levels and the mobile application used in the project, the EEG and brain waves, and finally the experimental design.

3.1Neurosky:

Neurosky is a device which allows us to perform a low-cost EEG (electroencephalogram) by using inexpensive dry sensors. The basic outputs are,

- Brainwave bands
- Raw output

The device consists of eight main parts: ear clip, flexible ear arm, battery area, power switch, adjustable headband, sensor tip, sensor arm and inside think gear chipset.

Two dry sensors are used to detect and filter the EEG signals. The sensor tip detects electrical signals from the forehead of the brain. The headset's reference and ground electrodes are located on the ear clip, and the EEG electrode is on the sensor arm where it rests on the forehead just above the eye. Thinkgear is the technology inside every NeuroSky product that enables a device to interface with the wearers' brainwaves. It includes the sensor that touches the forehead, the contact and reference points located on the ear pad, and the onboard chip that processes all the data and provides this data to software and applications in digital form.

This chip processes all the data and provides this data to software and applications in digital form.

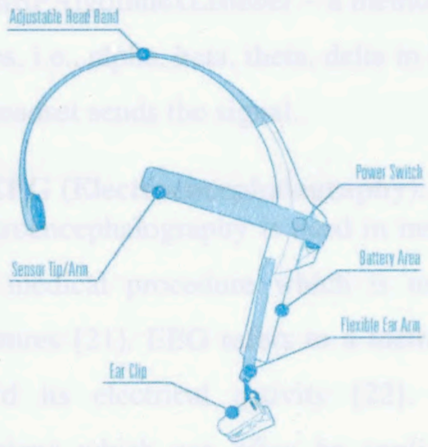

Figure 2 Components of Neurosky

By using neurosky, we apply a constant low voltage to our head which does not affect the brain. When the person feels pressure, the interaction between neurons creates electrical discharges. These electrical discharges can be measured using this neurosky, because of the sensors present in it. These patterns are characterized as waves with amplitudes and frequencies.

Each wave pattern is associated with some characteristics, defining our state of mind. The behavior of these wave patterns and their power levels at that point in time can be used to measure the mental stress.

3.2Mobile application:

The mobile app used in this project is an open source android application from Neurosky Inc. called as Algo_SDK_Sample. The application consists of four types of algorithms. They are attention, meditation, blink, and band power. The attention algorithm gives out the attention values of the participants. The band wave power algorithm is used to record the raw data of brain waves, and it also displays the waveforms. The app was modified to give out the numerical values of those brain waves into a .csv file.

The app is build using neuro sky library,

TgStreamReader – is a class from the neurosky library, used to connect our device with the neuro sky MindWave mobile through Bluetooth

setOnSignalQualityListener – a method in class NskAlgoSdk, which gives the quality of the signal of the MindWave mobile.

setOnBPAlgoIndexListener – a method in class NskAlgoSdk, which provides the values of brain waves, i.e., alpha, beta, theta, delta in floating point decimals which is automatically called when the headset sends the signal.

3.3 EEG (Electroencephalography):

Electroencephalography is used in medical and research domain areas. Electroencephalography is a medical procedure which is used to read scalp electrical activity generated by brain structures [21]. EEG refers to a method for calculating electro-physical monitoring of brain to record its electrical activity [22]. Electroencephalography is a completely non-incursive technique which can often be applied to adults, children, and patients without risk or any limitation [21].

Nondestructive, less side effect and less pain are characteristics of EEG. Also, this is a proper consideration for some brain diseases like memory loss and autism. EEG signals are categorized based on different states frequencies like eye open, eye close, eyeball movement, finger clenching. These signals have different frequency ranges from 0 Hz to 100 Hz. These signals help researchers to understand the characteristics of complex brains [22]. Current flows are created when brain cells are activated. Mostly K+, Na+, Ca++ and Cl- ions are present in brain electrical current which pumped in neuron membranes through channels [22]. On the head surface, large active neurons can generate recordable electrical activity. Scalp electrode's weak electrical signals are amplified and then stored in computer memory or displayed on paper. EEG found a powerful tool in the domain of neurology and clinical neurophysiology due to the capability to reflect the abnormal and normal electrical activity of the brain. Brain electrical activity starts from around 17-23 week of prenatal development. Nearly full neural cell, i.e., 1011 neural cells is already developed at the time of birth. Neural nets though synapses are connected to neurons. The number of neurons and synapses decreases with age. Brain can be divided into three sections: cerebrum, cerebellum, and brain stem. The cerebrum contains the cerebral cortex which is left and right hemisphere with a highly complex surface layer. This cortex is a part of the central nervous system. The cerebrum acquires initiation movement, complex analysis, and consciousness of sensation. EEG is highly effective as it comes from cerebral cortex electric activity [22].

3.3.1 History of EEG

Encephalography has experienced massive progress from more than 100 years of its history. Richard Caton discovered brain electrical currents in 1875. Caton observed the EEG with rabbits and monkey's brains. A German neurologist, Hans Berger used radio equipment in 1924 to amplify the brain's electrical activity measured. He discovered that without opening the skull, weak electric currents could be recorded. He observed that activity changed as per brain functional status like in sleep, anesthesia, and lack of oxygen. Berger uses electroencephalogram word as the description of human brain electric potentials. His suggestion was right, that brain activity changes from relaxation to alertness. Later in 1934, "human brain waves" paper was published by Adrian and Matthews explains the concept of regular oscillations around 10 to 12 Hz which named as "alpha rhythm" [21].

3.3.2 Brainwaves classification

For acquiring basic brain patterns, subjects are requested to close their eyes and relax. Usually, pattern range is from 0.5 to 100 µV in amplitude and 100 times lower than ECG signals. Various frequencies are visible in power spectrum contribution of sine waves. The individual brain can make dominant frequencies even though the spectrum is ranging from 0 Hz up to one half of sampling frequency. There are four groups of brain waves (Figure 1) [21]:

beta (>12 Hz), alpha (8-12 Hz), theta (4-7 Hz), delta (0.5-4 Hz).

Brainwave Type	Frequency range	Mental states and conditions
Delta	0.1Hz to 3Hz	Deep, dreamless sleep, non-REM sleep, unconscious
Theta	4Hz to 7Hz	Intuitive, creative, recall, fantasy, imaginary, dream
Alpha	8Hz to 12Hz	Relaxed, but not drowsy, tranquil, conscious
Low Beta	12Hz to 15Hz	Formerly SMR, relaxed yet focused, integrated
Midrange Beta	16Hz to 20Hz	Thinking, aware of self & surroundings
High Beta	21Hz to 30Hz	Alertness, agitation
Gamma	30Hz to 100Hz	Motor Functions, higher mental activity

Table 1: Different types of brain waves and their frequencies

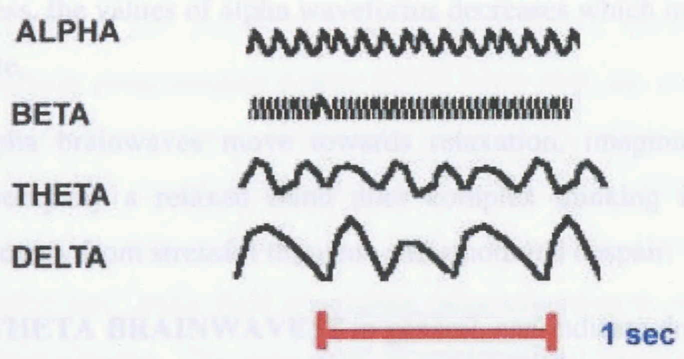

Figure 3: Brainwave samples with dominant frequencies belonging to beta, alpha, theta, and delta band

"BETA BRAINWAVES" are in general, the characteristics of an engaged mind which as we know is highly alert and well-focused. Beta activity is quick-connecting, fast action. When attention is directed towards the outside world, this Beta activity tends to dominate the normal waking state of consciousness. This activity is detected typically in the frontal lobes (where decisions are managed). Beta may be absent or reduced in areas of brain damage, and it is seen on both sides of the brain in a geometric distribution. In general, it was regarded as a normal rhythm but in those who are alert, anxious or have their eyes open, and it tends to be the dominant rhythm.

When the person experience stress, the values of beta waveforms increases to indicate the change in behavior.

Beta brainwaves are engaged when the brain is processing activities such as

- public speaking, lectures or teaching information.
- Conversations with others that command your full attention.
- problem-solving ability and assessment of situations.

"ALPHA BRAINWAVES" in general represent a relaxed awareness in mind mainly as they are very slow compared to Beta. When the brain sets itself to rest, this rhythm is seen. Alpha waves values are majorly increased by closing the eyes and relaxing yet are offset by opening one's eyes or any effort. Alpha is the rhythm seen in normal adults who are very relaxed and is typically regarded as the common mode of relaxation beyond the age of 13. These are usually best detected in the front regions of the head, on each side of the brain. When the person experience stress, the values of alpha waveforms decreases which means he/she is readily losing the relaxing state.

Alpha brainwaves move towards relaxation, imagination and good thinking. These waves accompany a relaxed mind after complex thinking into a mode of relaxation, meditation, recovery from stressful thoughts and emotional despair.

"THETA BRAINWAVES" in general, can indicate drowsiness, daydreaming, the first stage of sleep. Theta activity is not seen in awake adults (unless engaged in a meditative practice) but is perfectly normal in children up to 13 years.

The Theta state can be termed as a gateway to hypnagogic states that lay between being awake and falling asleep. It can promote flashes of mental imagery as one becomes receptive to brain/mind information beyond one's awareness. Theta has been defined as a part of learning, memory, and reductions in stress. When the person experience stress, the values of theta waveforms decreases.

Theta brainwaves are related to accessing subconscious information that eludes the conscious mind, hypnagogic states like 'daydreaming' and reductions such as heart rate and breathing.

3.3.3 Applications

According to R. Bickford research, EEG technology applications are used to detect the state of mind (the activity of the brain) humans and animals such as below [21]:

- Coma, brain death, and monitor alertness.
- Test pathways. Test epilepsy drug effects. Test drugs for convulsive effects.
- Monitor cognitive engagement. Monitor human and animal brain development.
- Locate damage areas of head injury, stroke, tumor, etc.
- Produce biofeedback situations.
- Investigate epilepsy and locate seizure origin.
- Control anesthesia depth
- Assist in experimental cortical excision of epileptic focus
- Investigate sleep disorder and physiology.

3.4 Experiment Design

The goal of this project is to collect data on the stress level of 20 students from introductory computer programming course (CS1) while they are coding. The data collected will be used to evaluate whether stress level varies significantly when a CS1 student writes a computer program. If we find a significant rise in the stress level of a CS1 student during coding, we will further analyze the data to see if there is any correlation between stress level during coding and the gender, age, major field of student, prior exposure to programming and math background of an average student. The stress level of a student will be measured by using a device called Neurosky which measures brain waves and calculates the stress level from that data. An email has been sent to all the students who belong to the CPSC 1301 course at Columbus State University requesting them to participate in the experiment in a lab. Those students who were interested in responded to that email. The interested students were thus requested to the lab. Once the participant attended the lab, he/she was asked to complete a perceived stress survey. Then the participant was given with a headset(neurosky) to wear to collect the brain wave data from participants who are studying the CS1 course. The entire procedure is divided into two steps. In the first step, simple additions and subtractions were given to participants to make them focused. In the second step, a simple program was given to the participants after the completion of their first task. An interval of 1minute was given to participants between these two steps. The brain waves of the participants were recorded from the 6 participants during these two tasks. Each step was performed for 6-7 minutes depending upon the convenience of the participants. A 2-3 minutes interval was given to the participant between the two tasks. Initially, the students were asked to wear this device to their head. The device was connected using Bluetooth to an Android device which has the application called algo_sdk_sample.

In step 1, simple math was given to participants. While students were doing that task, the app recorded their brain waves and displayed them in the form of graphs. The app also saved the data in an excel sheet.

1) 29 + 73

2) 323 + 126

3) 32 + 39

4) (17 + 21) - 36

5) 234 + 729

6) 729 + 81

7) 802 - 246

8) 777 + 329

9) 79 + 128

10) (69 + 37) - 46

11) (777 - 129) + 36

12) 421 - 239

13) 678 + 321

Figure 4: Basic math test for task 1

The app recorded the alpha, beta and theta values for 6-7 minutes during task 1. Once the task was completed, the app was closed, and the device was disconnected. Before starting task 2, the device was again connected to the android device.

In step 2, a simple program was given to the participants with a time limit of 6-7 minutes. While students were doing this task, their brain values were recorded using the app.

> Write a program that prompts the user for two integers and then prints out all prime numbers within that range. For example, if the user enters 10 and 20, the program prints all primes between 10 and 20:
>
> 11, 13, 17, 19
>
> Note: A prime number is prime if it is not divisible by any number except 1 and itself.

Figure 5: Sample program for task 2

After successful completion of both the tasks, the numerical values of alpha, beta and theta values were compared to check whether the students had felt stress during programming.

3.5 Statistical hypothesis testing:

A statistical hypothesis is a hypothesis that is testable based on observing a process that is modeled via a set of random variables. In simpler terms, it is a method of statistical inference. It's a procedure that involves formulating a statistical hypothesis and using a sample data to decide on the validity of the formulated statistical hypothesis. We can use this four-step procedure to do any hypothesis testing:

1. Set up the null and alternative hypotheses.

2. Define the test procedure, including the selection of significance level and power.

3. Calculate test statistics and Z value.

4. Conclude that data are consistent or inconsistent with the null hypothesis, i.e., decide about the null hypothesis.

Explained:

Commonly, two statistical data sets are compared, or a data set obtained by sampling is compared against a synthetic data set from an idealized model. A hypothesis is proposed for the statistical relationship between the two data sets, and this is compared as an alternative to an idealized null hypothesis that proposes no relationship between two data sets.

Hypothesis tests are used in determining what outcomes of a study would lead to a rejection of the null hypothesis for a pre-specified level of significance. The process of distinguishing between the null hypothesis and the alternative hypothesis is aided by identifying two conceptual types of errors (type 1 & type 2), and by specifying parametric limits on, e.g., how much type 1 error will be permitted.

An alternative framework for statistical hypothesis testing is to specify a set of statistical models, one for each candidate hypothesis, and then use model selection techniques to choose the most appropriate model. Statistical hypothesis testing is sometimes called confirmatory data analysis.

Z test and its importance:

The Standard Normal model is used in hypothesis testing, including tests on proportions and the difference between two means. The area under the whole of a normal distribution curve is 1, or 100 percent. The z-table helps by telling us what percentage is under the curve at any point. When the curve is standardized, we can use a Z Table to find percentages under the curve.

The z-test is a hypothesis test in which the z-statistic follows a normal distribution. The z-test is best used under the central limit theorem, as the number of samples gets larger, the samples are considered to be approximately normally distributed. When conducting a z-test, the null and alternative hypotheses, alpha and z-score should be stated.

A z-test is a statistical test used to determine whether two population means are different when the variances are known.

Z -test for the single mean is used to test a hypothesis on a specific value of the population mean.

Statistically speaking, we test the null hypothesis H0: $\mu = \mu 0$ against the alternative hypothesis H1: $\mu >< \mu 0$ where μ is the population means and $\mu 0$ is a specific value of the population that we would like to test for acceptance. Finally, the population standard deviation is known.

Steps:

Step 1: State the Null hypothesis, H_0.

Step 2: State the Alternate Hypothesis, H_1. The fact that we are looking for scores "greater than" a certain point means that this is a one-tailed test.

Step 3: Draw a picture to help you visualize the problem.

Step 4: State the alpha level. If you aren't given an alpha level, use 5% (0.05).

Step 5: Find the rejection region area (given by your alpha level above) from the z-table. An area of 0.05 is equal to a Z-score of 1.645.

Step 6: Find the test statistic using this formula:

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

Step 6: If Step 6 is greater than Step 5, reject the null hypothesis. If it's less than Step 5, you cannot reject the null hypothesis.

When Z test proves the Hypothesis testing as wrong?

Considering the right Z table, when the Z value calculated from the above formula is greater than the Z value at α (Alpha), then our Null Hypothesis falls under the "Rejection region." Therefore, we reject the Null Hypothesis. Its vice versa for the left Z table.

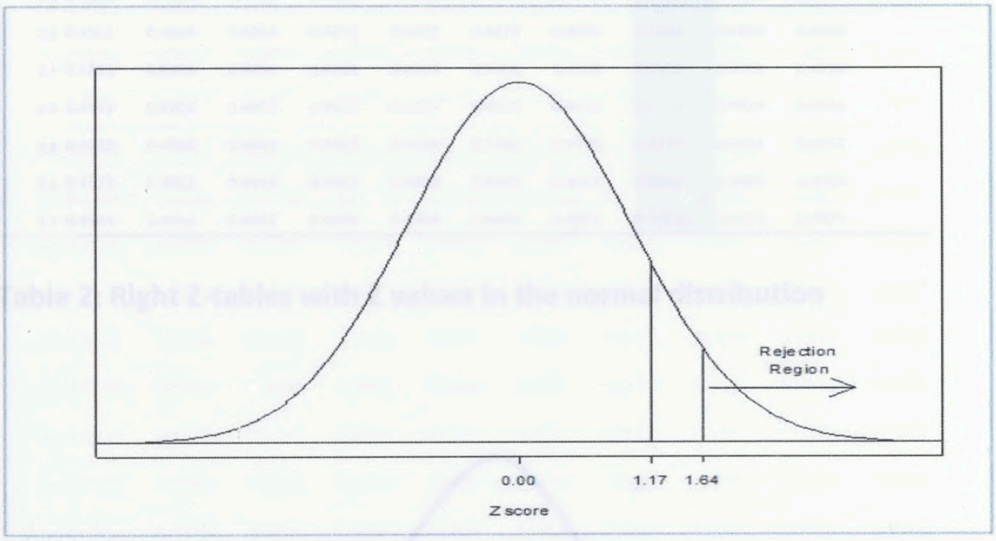

Figure 6: Probability distribution curve with the rejection region after z =1.645 at alpha 0.05

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974

Table 2: Right Z-tables with Z values in the normal distribution

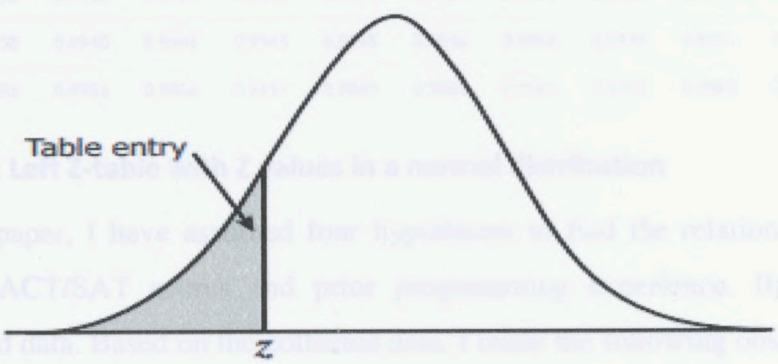

Figure 7:Probability distribution curve with rejection region before z = -1.805 at alpha 0.05

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.0120	0.0160	0.0199	0.5239	0.0279	0.0319	0.0359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6064	0.1064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964

Table 3: Left Z-table with Z values in a normal distribution

In this paper, I have assumed four hypotheses to find the relation between stress and gender, major, ACT/SAT scores and prior programming experience. Based on the pattern of the collected data. Based on the collected data, I made the following observations:

1. The average stress value of students who don't have prior programming experience is greater than the average stress value of students who have prior programming experience.

2. Average stress value of female students is greater than male students.

3. If the total SAT score of a student is less than the average SAT score (1060), then the average stress value of that student is high.

4. The average stress values of non-CS students are greater than average stress values of CS students.

Chapter 4: Discussion of Results and Conclusion

Analysis:

When a person feels pressure, the interaction between neurons creates electrical discharges. The sesnors in the neurosky can measure these electrical discharges as wave patterns. These wave patterns are characterized as brain waves. Each wave pattern is associated with some characteristics, defining our state of mind. The behavior of these wave patterns and their values at that point in time can be used to measure the mental stress. The electrical activity in a brain fluctuates with the level of stress. When the brain is not very active, i.e., the stress level is very low, alpha waves are more dominant. In stressful situations, the power of alpha and theta waves falls showing the change in response under stress. Beta waves rise under stress or mental tasks [26].

Here I performed three types of analyses to detect and measure stress.

1. Asymmetric alpha wave forms. (Stress gives too little of alpha waves). Asymmetric beta and theta wave froms (high beta and low theta values for stress).
2. Measuring stress by assuming basic metric beta/(alpha * theta) for every second.
3. Calculating averages of the metric and correlate them with age, gender, ACT/SAT scores and with prior programming experience.

Contribution of the metric: It is known that the behavior of alpha, beta and theta waveforms changes during stress. The beta waveforms tend to increase during stress whereas the alpha and theta waveforms decrease during stress [26]. Keeping this fact in mind, we thought of assuming a metric compromising these three brainwave form values. The metric value must increase/decrease with the stress, and the metric values must differ similarly with time. The metric beta/ (alpha*theta) gives similar values ranging between 0-1 and increases with stress, because as mentioned in the above paragraph, beta values (numerator) increases during stress and alpha, theta values (denominators) decreases making the whole metric value big and vice versa. For example, assume beta, alpha, theta values as 12, 8, 4 giving the metric value as 0.375. Similarly, assume beta, alpha, and theta values as 10, 9, 6 then the metric value is 0.185. Here in the first case, the beta value is high compared to second case and alpha, theta values are less, that means the participant has high stress in the first case, so as the metric value (0.375).

Participant1:

1) Changes in alpha, beta and theta values:

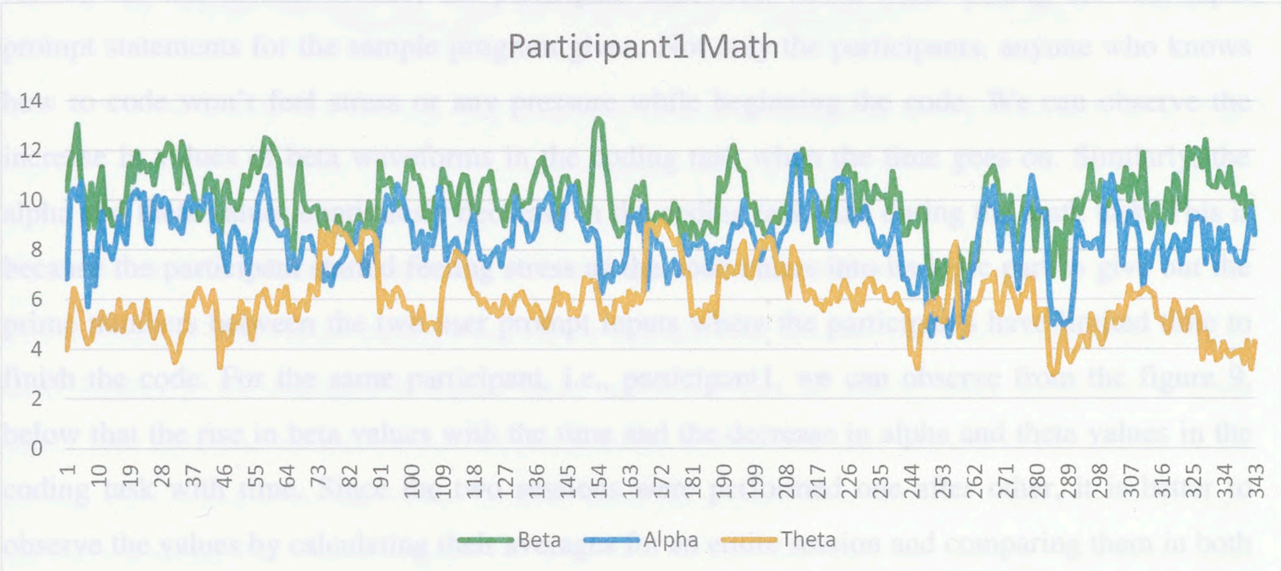

Figure 8: beta, alpha, theta values of participant1 during math task.

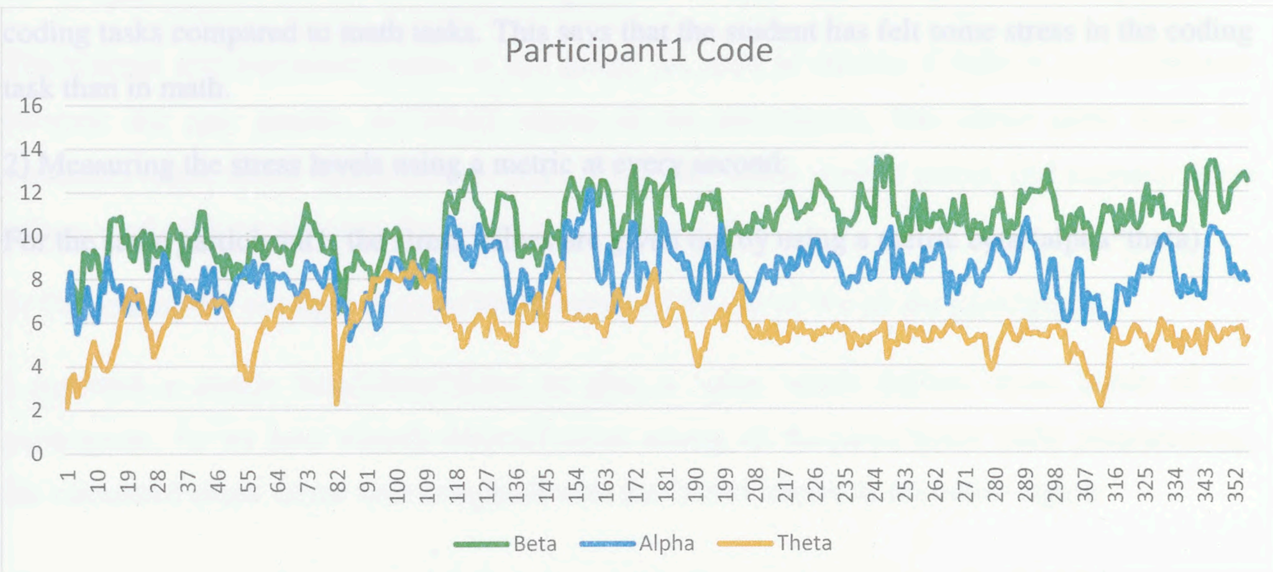

Figure 9: beta, alpha, theta values of participant1 during programming.

From the figure 8 and 9, there are changes in the values of alpha, beta and theta values during math and programming tasks. The alpha, beta and theta curves are asymmetric. Initially, the beta values in coding sessions are less than the beta values in math task as the participant has prior programming experience and good at programming, but when the time goes on, the beta values gradually increases exceeding the beta values during math task. Similarly, alpha, theta values decreases with increase in time when compared to the math task. The beta values in the code session are less at first because the participant didn't feel stress while writing the user input prompt statements for the sample program given. Not only the participants, anyone who knows how to code won't feel stress or any pressure while beginning the code. We can observe the increase in values of beta waveforms in the coding task when the time goes on. Similarly, the alpha and theta values continuosly decrease in the coding task than during the math task. This is because the participant started feeling stress as the code enters into its logic part to give out the prime numbers between the two user prompt inputs where the participants have limited time to finish the code. For the same participant, i.e., participant1, we can observe from the figure 9, below that the rise in beta values with the time and the decrease in alpha and theta values in the coding task with time. Since the two sessions were performed one after other, it is better to observe the values by calculating their averages for an entire session and comparing them in both tasks. When we calculate their averages, the beta values during math task is less than the average beta value during the coding task. Similarly, the average values of alpha, theta values are less in coding tasks compared to math tasks. This says that the student has felt some stress in the coding task than in math.

2) Measuring the stress levels using a metric at every second:

For the same participant1, the stress values are given out by using a metric beta/(alpha*theta).

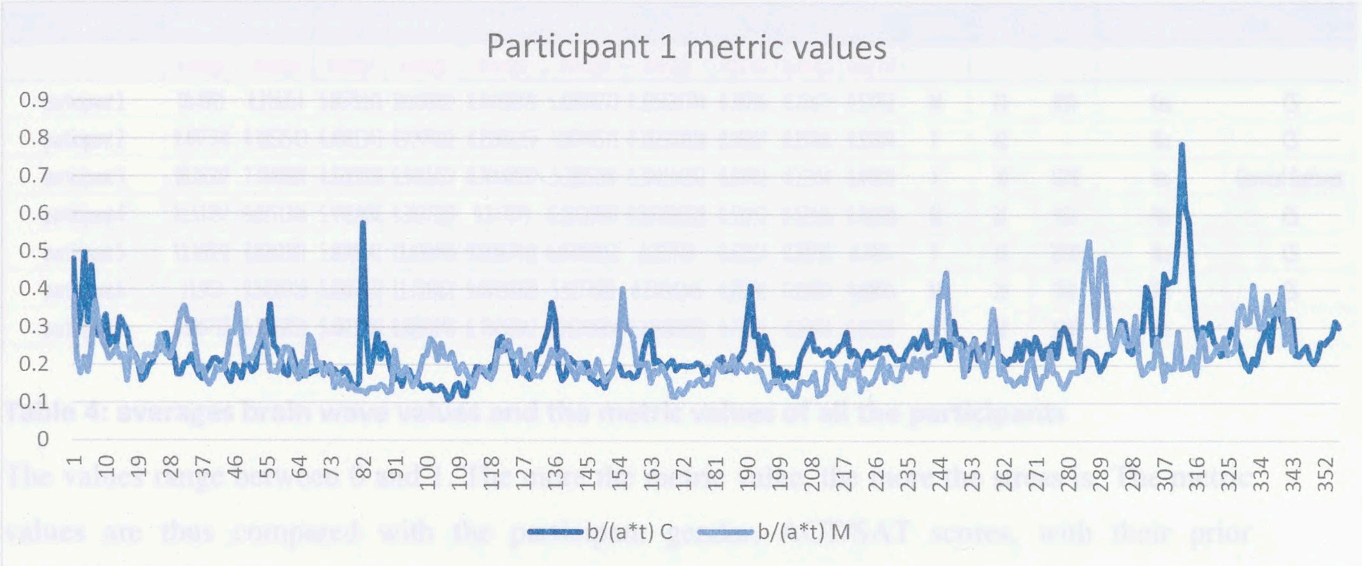

Figure 10: metric values during math and coding tasks of participant1

The average of beta values during math task is less than the average beta value during the coding task. Similarly, the averages of alpha, theta values are less in coding tasks compared to math tasks. This says that the student has felt stress in the coding task than in math in the middle of the coding session and at the end of the session. Each participant gave 300-400 values on an average for each session. The values depend on what part of the code they are writing. The participant values vary much when they are trying to code the logic part than writing user input prompt statements.

The average and maximum values of this metric are used to observe if there is any correlation between the age, gender, ACT/SAT scores of the participants. The above plots show the variations of the stress values in both tasks. Thus, the metric used to detect, and measure stress values worked good and gave the values range between 0-1.

3) Calculating the averages and maximum values of the metric for all the participants:

I assumed a metric beta/(alpha*theta) to give a value which defines stress levels of the participants. As we have already detected stress among all the participants while programming, the calculated stress values thus compared with the factors shown in the below figure.

Student	beta Code	alpha Code	theta Code	beta Math	alpha Math	theta Math	b/(a*t) Code		b/(a*t) Math		Gender	Age	ACT/SAT	Prior Programing	Major
	average	average	average	average	average	average	average	max val	average	max val					
participant 1	10.4991	8.216354	5.8075931	10.030935	8.45098979	5.82000722	0.235631744	0.78726	0.21523	0.52951	M	21	850	Yes	CS
participant 2	9.497594	8.1852543	6.4981343	8.9205612	8.25586257	7.08749571	0.182340918	0.34097	0.15746	0.55374	F	42	-	No	CS
participant 3	10.26507	7.7849094	4.6283068	9.9461427	8.74148537	5.3805199	0.324114162	0.80703	0.21914	0.49803	F	20	1090	No	General Business
participant 4	10.51297	9.0711746	5.9382856	9.3927189	9.177079	6.18426897	0.205016328	0.53152	0.17346	0.38198	M	21	950	No	CS
participant 5	11.10331	8.8241101	5.8049641	11.899496	9.81167441	6.07098161	0.237749	0.61514	0.20754	0.7495	F	18	1090	No	CS
participant 6	11.952	9.5436918	5.8587922	11.721451	9.64393929	5.9277603	0.22491346	0.76581	0.21083	0.40453	M	29	950	Yes	CS
participant 7	6.830737	6.1243859	5.4811703	6.6869446	6.78456942	6.16249604	0.228289902	0.77451	0.16844	0.80265	F	18	1080	No	CS

Table 4: averages brain wave values and the metric values of all the participants

The values range between 0 and 1. The more the metric value, the more the stress is. The metric values are thus compared with the participant gender, ACT/SAT scores, with their prior programming experience and the major they study.

Participant2:

1) Changes in alpha, beta and theta values:

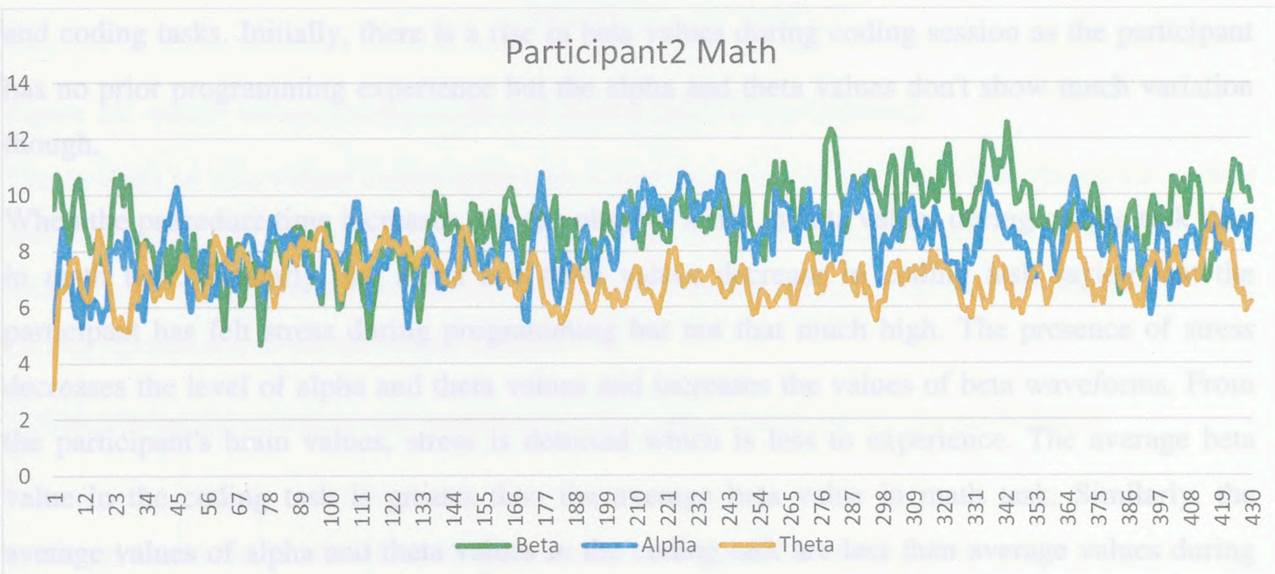

Figure 11: beta, alpha, theta values of participant 2 during math task.

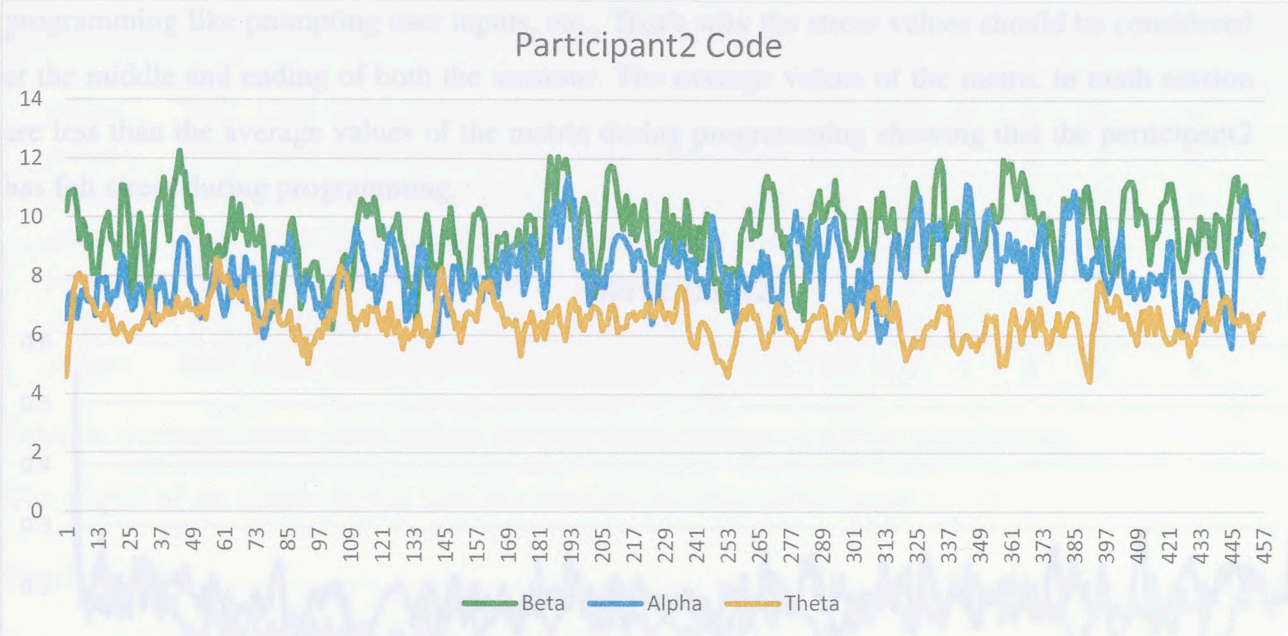

Figure 12: brain values of participant2 during coding task

From the figure 12, we can observe that the beta, alpha, and theta values changes during math and coding tasks. Initially, there is a rise in beta values during coding session as the participant has no prior programming experience but the alpha and theta values don't show much variation though.

When the procedure time increases, we can observe a rise in beta values during coding task than in math task. Similarly, the alpha and theta values decrease in coding task saying that the participant has felt stress during programming but not that much high. The presence of stress decreases the level of alpha and theta values and increases the values of beta waveforms. From the participant's brain values, stress is detected which is less to experience. The average beta value in the coding task is greater than the average beta value in math task. Similarly, the average values of alpha and theta values in the coding task are less than average values during math task. These values at the middle and end of both the sessions say that the participant2 felt stress while programming which is not much in quantity.

2) Measuring stress values using the metric:

The stress values of the participants are given by a metric beta/(alpha * theta). Initially, the brain waveforms don't show any stress because no one feels stress while writing fundamental steps in programming like prompting user inputs, etc., That's why the stress values should be considered at the middle and ending of both the sessions. The average values of the metric in math session are less than the average values of the metric during programming showing that the participant2 has felt stress during programming.

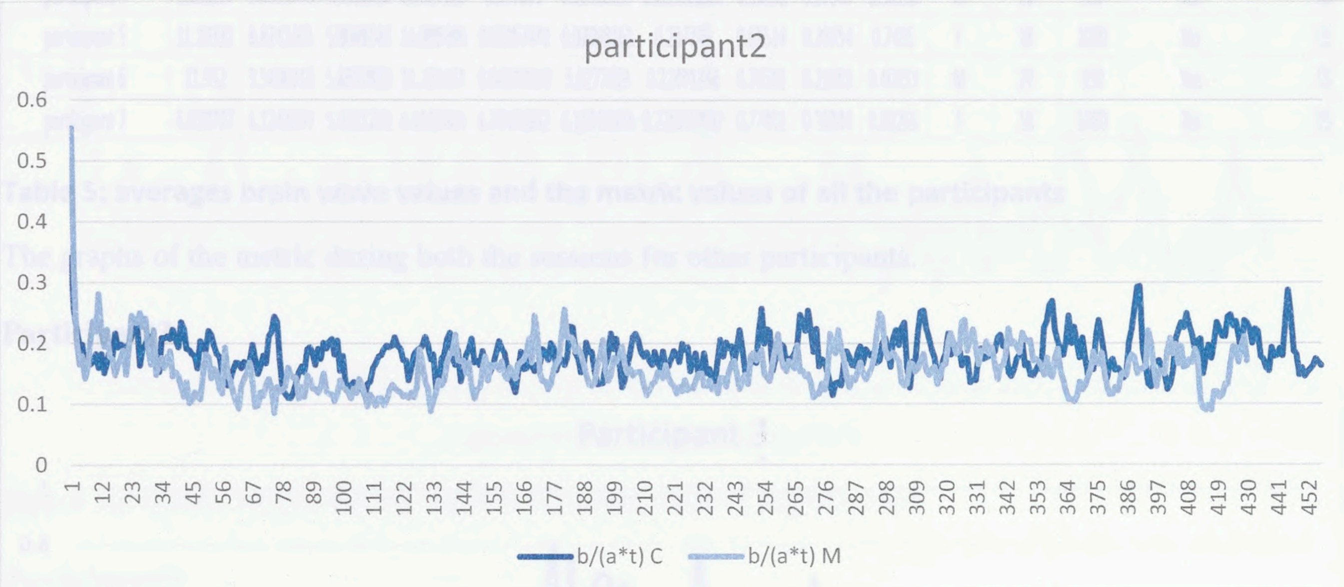

Figure 13: metric values during math and coding tasks of participant2.

The average of beta values during math task is less than the average beta value during the coding task. Similarly, the averages of alpha, theta values are less in coding tasks compared to math tasks. This says that the student has felt stress in the coding task than in math in the middle of the coding session and at the end of the session.

3) Calculating the averages and maximum values of the metric for all the participants:

I assumed a metric beta/(alpha*theta) to give a value which defines stress levels of the participants. As we have already detected stress among all the participants while programming, the calculated stress values thus compared with the factors shown in the below figure.

Student	beta Code	alpha Code	theta Code	beta Math	alpha Math	theta Math	b/(a*t) Code		b/(a*t) Math		Gender	Age	ACT/SAT	Prior Programing	Major
	average	average	average	average	average	average	average	max val	average	max val					
participant 1	10.4991	8.216354	5.8075931	10.030935	8.45098979	5.82000722	0.235631744	0.78726	0.21523	0.52951	M	21	850	Yes	CS
participant 2	9.497594	8.1852543	6.4981343	8.9205612	8.25586257	7.08749571	0.182340918	0.34097	0.15746	0.55374	F	42	-	No	CS
participant 3	10.26507	7.7849094	4.6283068	9.9461427	8.74148537	5.3805199	0.324114162	0.80703	0.21914	0.49803	F	20	1090	No	General Business
participant 4	10.51297	9.0711746	5.9382856	9.3927189	9.177079	6.18426897	0.205016328	0.53152	0.17346	0.38198	M	21	950	No	CS
participant 5	11.10331	8.8241101	5.8049641	11.899496	9.81167441	6.07098161	0.237749	0.61514	0.20754	0.7495	F	18	1090	No	CS
participant 6	11.952	9.5436918	5.8587922	11.721451	9.64393929	5.9277603	0.22491346	0.76581	0.21083	0.40453	M	29	950	Yes	CS
participant 7	6.830737	6.1243859	5.4811703	6.6869446	6.78456942	6.16249604	0.228289902	0.77451	0.16844	0.80265	F	18	1080	No	CS

Table 5: averages brain wave values and the metric values of all the participants

The graphs of the metric during both the sessions for other participants.

Participant3:

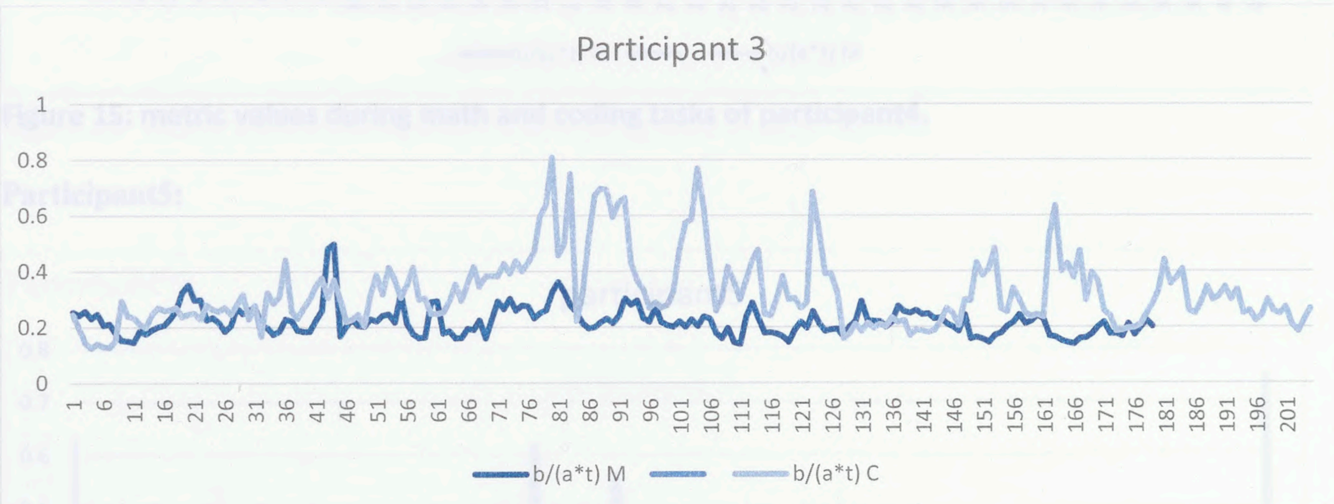

Figure 14: averages brain wave values and the metric values of all the participants

Participant4:

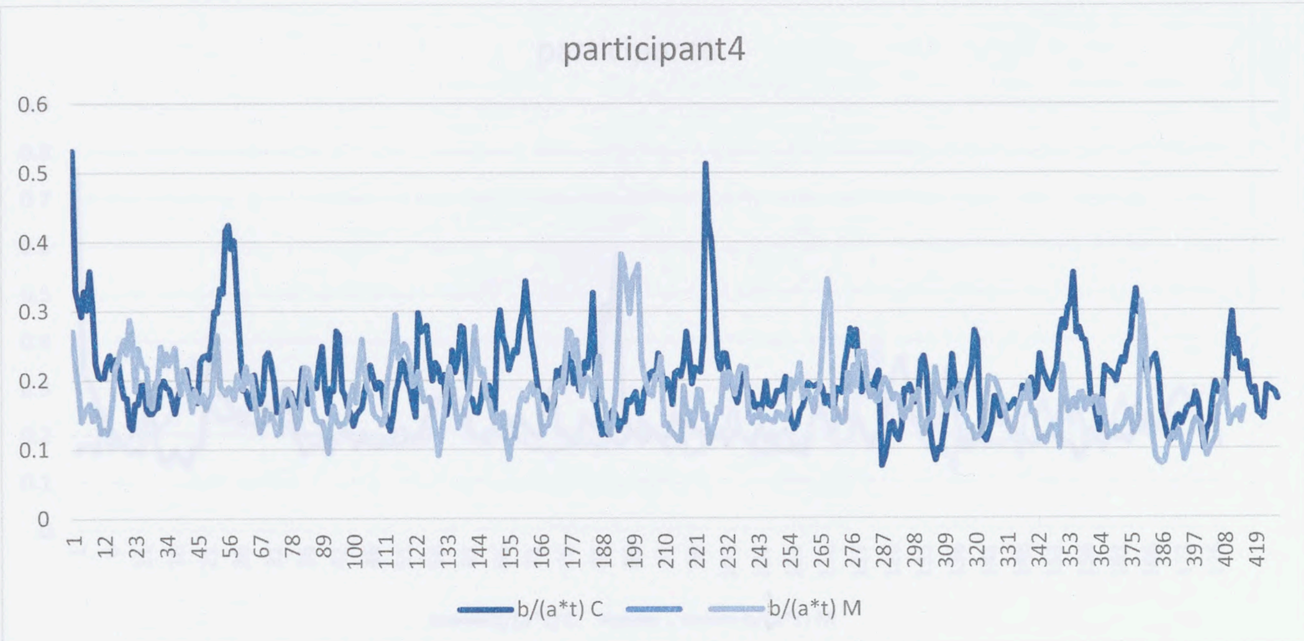

Figure 15: metric values during math and coding tasks of participant4.

Participant5:

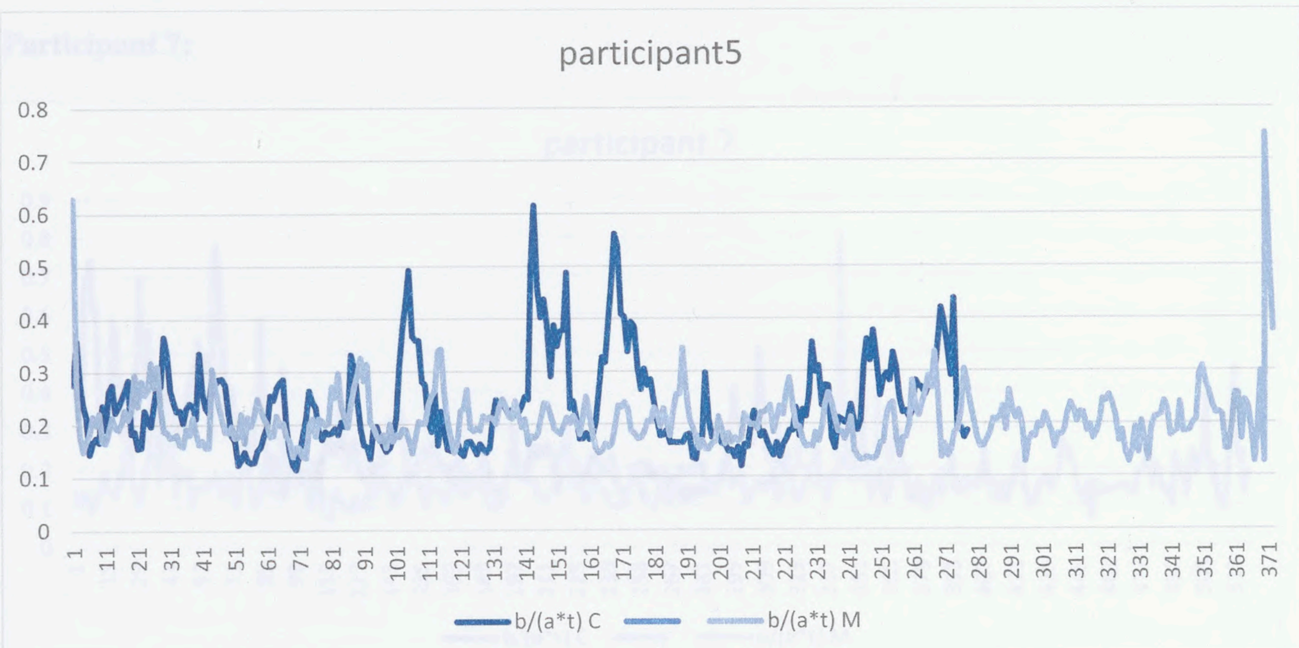

Figure 16: metric values during math and coding tasks of participant5.

Participant 6:

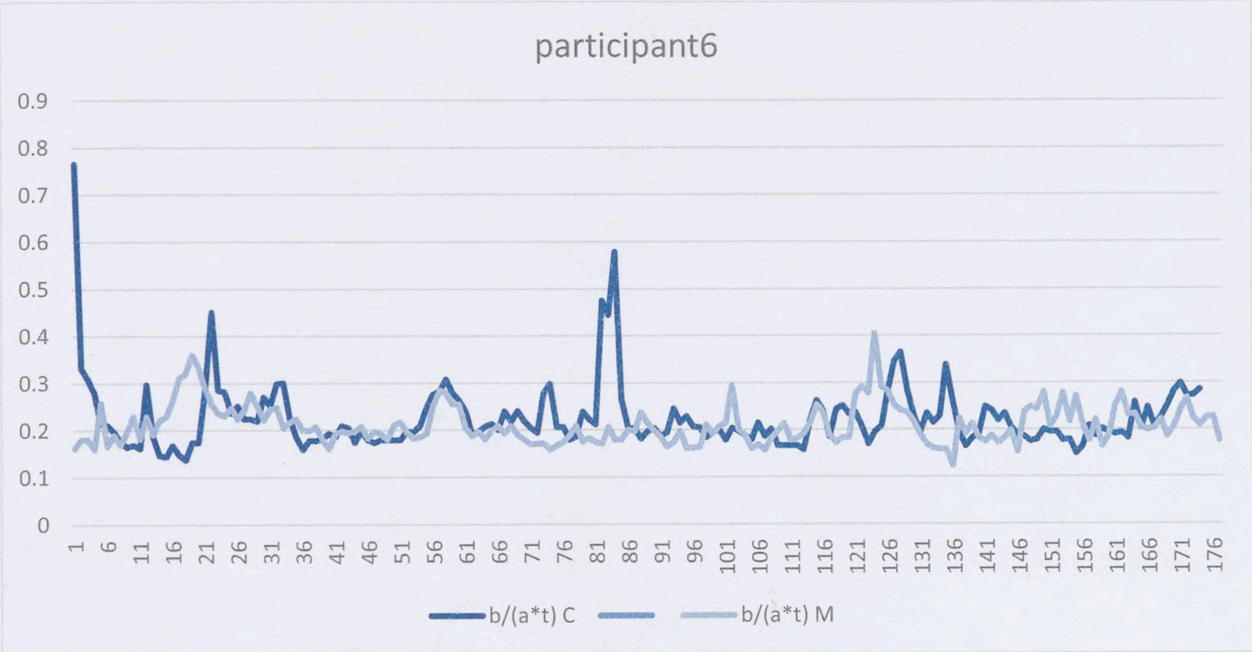

Figure 17: metric values during math and coding tasks of participant6.

Participant 7:

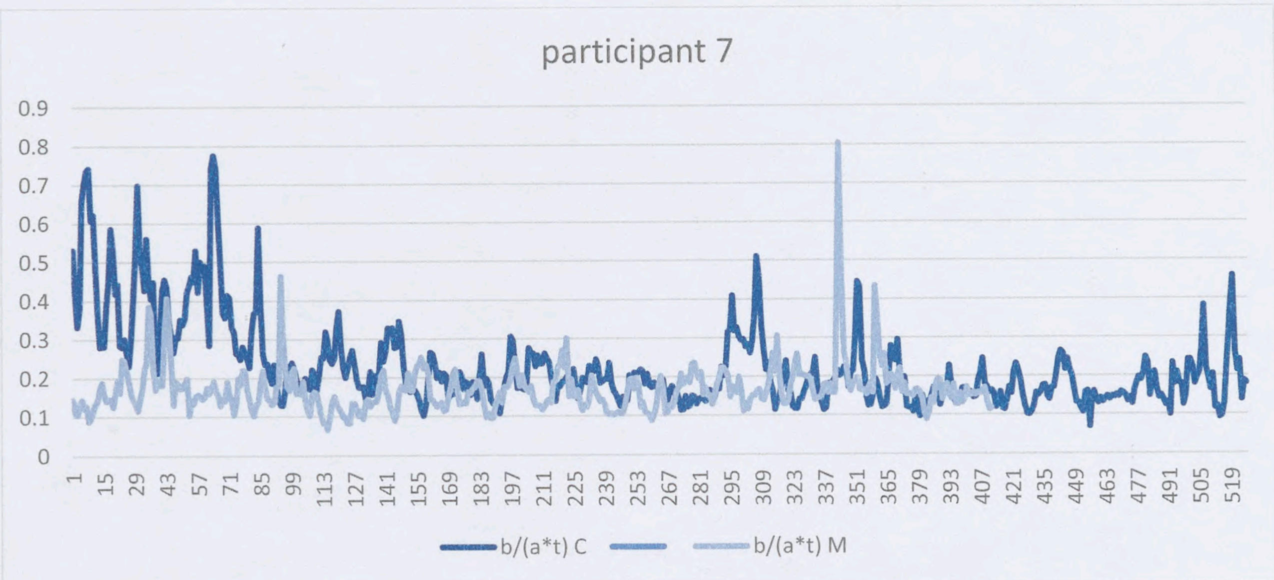

Figure 18: metric values during math and coding tasks of participant7.

The calculated stress values are thus used to find any correlation with gender, SAT scores, prior programming experience and major they study.

Student	beta Code	alpha Code	theta Code	beta Math	alpha Math	theta Math	b/(a*t) Code		b/(a*t) Math		Gender	Age	ACT/SAT	Prior Programing	Major
	average	average	average	average	average	average	average	max val	average	max val					
participant 1	10.4991	8.216354	5.8075931	10.030935	8.45098979	5.82000722	0.235631744	0.78726	0.21523	0.52951	M	21	850	Yes	CS
participant 2	9.497594	8.1852543	6.4981343	8.9205612	8.25586257	7.08749571	0.182340918	0.34097	0.15746	0.55374	F	42	-	No	CS
participant 3	10.26507	7.7849094	4.6283068	9.9461427	8.74148537	5.3805199	0.324114162	0.80703	0.21914	0.49803	F	20	1090	No	General Business
participant 4	10.51297	9.0711746	5.9382856	9.3927189	9.177079	6.18426897	0.205016328	0.53152	0.17346	0.38198	M	21	950	No	CS
participant 5	11.10331	8.8241101	5.8049641	11.899496	9.81167441	6.07098161	0.237749	0.61514	0.20754	0.7495	F	18	1090	No	CS
participant 6	11.952	9.5436918	5.8587922	11.721451	9.64393929	5.9277603	0.22491346	0.76581	0.21083	0.40453	M	29	950	Yes	CS
participant 7	6.830737	6.1243859	5.4811703	6.6869446	6.78456942	6.16249604	0.228289902	0.77451	0.16844	0.80265	F	18	1080	No	CS

Table 6: averages brain wave values and the metric values of all the participants

The ACT scores are converted into SAT scores by using an ACT-SAT calculator. For example, a student with ACT score ACTE 21, ACTM 24 gives an SAT score of 1090.

Testing the hypotheses:

Metric used to calculate stress level is beta/(alpha*theta).

1. Prior programming experience:

Null hypothesis: H_0: The average stress value of students who don't have prior programming experience is greater than the average stress value of students who have prior programming experience. Let's say, μ is the average stress value of students who do not have prior programming experience. Null hypothesis is: μ = population average (average calculated on all students)

Alternative hypothesis: H_1: The average stress value of students who don't have prior programming experience is less than the average stress value of students who have prior programming experience. H1: μ < population average

$p = 0.234007931$, $\sigma = 0.040989831$, $n = 7$, $\bar{x} = 0.235502062$,

$Z = \frac{\bar{x} - p}{\sigma/\sqrt{n}} = 0.096901696.$

Where, p = population mean,

σ = standard deviation

n = number of samples

$\bar{x}$ = sample mean

From the Right Z table, Z at alpha (5%) =1.645

Considering the right Z table, when the Z value calculated from the above formula is less than the Z value at α (Alpha) then our Null Hypothesis doesn't fall under the "Rejection region." Therefore, we accept the Null Hypothesis.

2. Gender:

Null hypothesis: H_0: Average Stress value of female students is greater than male students. Let's say, μ is the average stress value of female students. Null hypothesis is: μ = population average (average calculated on all female students)

Alternative hypothesis: H_1: Average Stress value of female students is less than male students.

H_1: μ < average stress value of male students

$p = 0.234007931$, $\sigma = 0.040989831$, $n = 7$, $\bar{x} = 0.248068027$,

$Z = \frac{\bar{x}\ p}{\sigma/\sqrt{n}} = 0.90799119$

Where, p = population mean,

σ = standard deviation

n = number of samples

$\bar{x}$ = sample mean

From the Right Z table, Z at alpha (5%) =1.645.

Considering the right Z table, when the Z value calculated from the above formula is less than the Z value at α (Alpha) then our Null Hypothesis doesn't fall under the "Rejection region." Therefore, we accept the Null Hypothesis.

3. Total SAT score:

Null hypothesis: H_0: If the total SAT score of a student is less than the average SAT score (1060), then the average stress value of that student is high. Let's say, μ is the total SAT score value of student. Null hypothesis is: μ = population average (average calculated on all students).

Alternative hypothesis: H_1: If the total SAT score of a student is greater than the average SAT score (1060), then the average stress value of that student is high.

$p = 0.234007931$, $\sigma = 0.040989831$, $n = 7$, $\bar{x} = 0.248068027$

$Z = \frac{\bar{x} - p}{\sigma/\sqrt{n}} = -1.421648924$

Where, p = population mean,

σ = standard deviation

n = number of samples

$\bar{x}$ = sample mean

From the Left Z table, Z at alpha (5%) = -1.805.

Considering the left Z table, when the Z value calculated from the above formula is greater than the Z value at α (Alpha), then our Null Hypothesis doesn't fall under the "Rejection region." Therefore, we accept the Null Hypothesis.

4. Major:

Null hypothesis: H_0: The average stress values of non-CS students are greater than average stress values of CS students. Let's say, μ is the average stress values of non-CS students. Null hypothesis is: μ = population average (average calculated on all students).

Alternative hypothesis: H_1: The average stress values of non-CS students are less than average stress values of CS students.

We do not have enough data to test this hypothesis.

Support to the hypotheses:

The findings of Alex Lishinki and Aman Yadav says [24], "female students in computer science have lower self-efficacy and comfort levels than their male peers. In every instance of the self-efficacy question, male students rated themselves on average more likely to do well than did female students," which alternatively support our hypothesis that the female students feel more stress than male students during programming.

The findings of Kinnuen and Simon says, "The "OK, what now?" experience relates to the moment right after reading the directions. It is characterized by students' perception of their ability or knowledge to do the assignment. This experience had two distinctive variations – not knowing how or where to start and not knowing how to do the assignment. The former experience was characterized by the feeling of really not knowing at all where to start and left students puzzled, confused and "froze up." The latter case refers to the situation where the student understands where to start but how, exactly, to do the rest of the assignment is still somewhat unclear. In this situation students used strategies, such as, looking for examples in the textbook, doing research online and started writing the code down even though one did not know exactly what to do. The two embodiments of the OK, what now experience differ in regard to how most students felt," supports our hypothesis that students who have prior programming experience feel less stress than the students who don't have any programming experience [23].

The opinion of students are collected after the experiment on how they feel about CS courses? The below table gives the opinion of the participants and their stress values.

Opinions	b/(a*t) C	b/(a*t) M
Participant 1 I think that as a programmer I need to know how to figure out something that I do not know at all. This experiment made me realize that I was missing the skill to find a solution. It would be better if I am taught how to think instead of learning to program as a lecture.	0.235631744	0.215227611

Participant 2 I feel the CS course is very fast pace and show be broken down a little more.	0.182340918	0.157459973
Participant 3 Having more hands-on activities will make the course easier. Majority of the students today are visual and hands-on learners. Just lecturing makes it difficult for students to grasp the material.	0.324114162	0.21914036
Participant 4 I feel that it is better to learn the concepts while practicing more. Instead of giving very large projects which takes time, it is better to give small ones with more number of practice sessions.	0.205016328	0.173457497
Participant 5 We do have lots of labs that help, but during class, we will get a couple of projects that are long and cover stuff we have barely touched on. It stresses everyone out. Perhaps smaller projects that gradually increase in difficulty as the course progresses that are over what we were learning.	0.237749	0.207544627

Table 7: Opinion of students on how they feel about CS courses and their stress values.

Conclusion:

Eckerdal reports that the students' experience frustration and even depression during programming [25]. Students often felt frustrated, depressed, and humbled, but also eventually became confident, when they had succeeded in grasping the concept. Most of the students feel stress during programming whatever the quantity may be. The data collected using neurosky has clearly shown the difference in their brain waves during math and coding tasks. Although the data collected was small, they have supported all the assumed hypotheses. The hypotheses testing clearly showed that the assumed hypotheses are true with the data collected.

Future Work:

With a large data set, the fourth hypothesis, " The average stress values of non-CS students, are greater than average stress values of CS students." can be tested. Basing on these hypotheses and the opinions of the participants, an effective pedagogical strategy may be proposed which can help students improve their performance without experiencing any stress.

References:

[1] PICARD, R. W. *Affective computing.* MIT Press. Cambridge. 1997.

[2] Cenk Akbiyik, Faculty of Education. Computer Education and Instructional Technology Department.Kayseri, Turkey, Erciyes University, "Can Affective Computing Lead to More Effective Use of ICT in Education," (N.D), http://www.revistaeducacion.educacion.es/re352/re352_08ing.pdf, Last accessed on 04/01/2018.

[3] Hernani Costa and Luis Macedo, "Affective Computing," ATCM State of the Art, Theoretical Report CISUC, University of Coimbra, Portugal, (N.D), http://www.academia.edu/2890674/Affective_Computing, Last accessed on 04/01/2018.

[4] Jennifer S. Lerner, "Emotion and Decision Making," Riverside, California, 16 June 2014, https://scholar.harvard.edu/files/jenniferlerner/files/annual_review_manuscript_june_16_final.final_.pdf, Last accessed on 03/31/2018.

[5] Asha Kapur, Ajay Kapur, Naznin Virji-Babul, George Tzanetakis, and Peter F. Driessen, " Gesture-Based Affective Computing on Motion Capture Data," May 2005, http://www.karmetik.com/sites/default/files/publications/2005_acii_affect.pdf, Last accessed on 04/01/2018.

[6] Deepa Mathew, "Haptic Responses to Emotional Stimulation," University of Tampere School of Information Sciences Interactive Technology, M.Sc. thesis, 31 pages, April 2016, https://tampub.uta.fi/bitstream/handle/10024/99387/GRADU-1466515484.pdf?sequence=1, Last accessed on 03/31/2018.

[7] Grace Leslie, Rosalind Picard, and Simon Lui, "An EEG and Motion Capture Based Expressive Music Interface for Affective Neurofeedback," Media Lab, Massachusetts Institute of Technology, Singapore University of Technology and Design, (N.D), https://affect.media.mit.edu/pdfs/15.leslie-picard-lui.pdf, Last accessed on 04/01/2018.

[8] Javier Hernandez, Pablo Paredes, Asta Roseway, and Mary Czerwinski, “Under Pressure: Sensing Stress of Computer Users,” Toronto, ON, Canada, April 26 - May 01, 2014, https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/UnderPressure.pdf, Last accessed on 03/30/2018.

[9] Sergio Salmeron-Majadas, Olga C. Santos, Jesus G. Boticario," Exploring indicators from keyboard and mouse interactions to predict the user affective state," *Proceedings of the 7th International Conference on Educational Data Mining,* (N.D), https://pdfs.semanticscholar.org/c376/8a327382693c3cba7be7da37bc361bae29d0.pdf, Last accessed on 03/29/2018.

[10] Wolfgang Maehr, “eMotion Estimation of the User’s Emotional State by Mouse Motions,” Chalmers University of Technology, Gothenburg, (N.D), https://www.einarsen.no/can-you-detect-user-emotion-with-only-mouse-movements/, Last accessed on 03/29/2018.

[11] Y. H. Gulhane, S. A. Ladhake, *Proceedings of International Journal of Emerging Research in Management &Technology* ,"A Short Survey on Methodology for Stress Recognition," ISSN: 2278-9359 (Volume-5, Issue-12), 2004.

[12] Jing Zhai and Armando Barreto, "Stress Recognition Using Non-invasive Technology," NSF grants CNS-0520811,IIS-0308155, HRD-0317692 and CNS-0426125, Florida International University, Florida, https://www.aaai.org/Papers/FLAIRS/2006/Flairs06-077.pdf, Last accessed on 04/03/2018.

[13] Teresa Marrin and Rosalind Picard, ” Analysis of Affective Musical Expression with the Conductor.s Jacket”, M.I.T. Media Laboratory Perceptual Computing Section Technical Report No. 475 Appears in *Proceedings of the XII Colloquium on Musical Informatics*, Gorizia, Italy, September 1998.

[14] Eerik Vesterinen, "Affective Computing," Digital media research seminar, Co-learnit Ltd., Ionoff Ltd. Torkkelinkatu 7 a 25 FI-00500 Helsinki, Finland, spring 2001. https://pdfs.semanticscholar.org/9b85/a01a87e4e0830886c9725e1c1c387433cf04.pdf, Last accessed on 03/29/2018.

[15] Clifford Nass, Jonathan Steuer, and Ellen R. Tauber, "Computers are Social Actors," Human Factors in Computing Systems, CH194-4/94, Boston, Massachusetts USA April 24-28, 1994.

[16] Akane Sano, Andrew J. Phillips, Amy Z. Yu, Andrew W. McHill, Sara Taylor, Natasha Jaques, Charles A. Czeisler, Elizabeth B. Klerman, and Rosalind W. Picard, "Recognizing Academic Performance, Sleep Quality, Stress Level, and Mental Health using Personality Traits, Wearable Sensors, and Mobile Phones," *Proceedings of IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, Draft for Body Sensor Networks 2015.

[17] Hossein Mobahi, Shahin Ansari, "Fuzzy Perception, Emotion and Expression for Interactive Robots," 0-7803-7952-7/03, *IEEE,* 2003.

[18] Christoph Bartneck, "Interacting with an Embodied Emotional Character," *Proceedings of the Design for Pleasurable Products Conference (DPPI2004)*, Pittsburgh pp. 55-60. | DOI: 10.1145/782896.782911.

[19] Balaji Akula & James Cusick, "Impact of Overtime and Stress on Software Quality," *Proceedings of Conference on, The 4th International Symposium on Management, Engineering, and Informatics*, Orlando, FL, 10.13140/RG.2.2.12815.59041, January 2008. https://www.researchgate.net/publication/259781769_Impact_of_Overtime_and_Stress_on_Software_Quality, Last accessed on 04/16/2018.

[20] Daniel Graziotin , Fabian Fagerholm , Xiaofeng Wang and Pekka Abrahamsson, "Consequences of Unhappiness While Developing Software," 24 February 2017, https://www.arxiv.org/pdf/1701.05789.pdf, Last accessed on 04/21/2018.

[21] Measurement Science Review by M. Teplan, "*Fundamentals of EEG Measurement*," Volume 2, Section 2, September 2002.

[22] J Satheesh kumar and P Bhuvaneshwari, "*Analysis of Electroencephalography (EEG) Signals and Its Categorization*," Edited by R. Rajesh, K. Ganesh, Siau Ching Lenny Koh, Volume 38, Pages 1-4064 (2012), https://www.sciencedirect.com/science/article/pii/S1877705812022114, Last accessed on 04/22/2018

[23] Päivi Kinnunen, Beth Simon, "Experiencing Programming Assignments in CS1: The Emotional Toll," *Proceedings of the Sixth international workshop on Computing education research*, Aarhus, Denmark, August 9–10, 2010, http://ims.mii.lt/ims/konferenciju_medziaga/ICER'10/docs/p77.pdf, Last accessed on 05/01/2018.

[24] Alex Lishinki and Aman Yadav, "Students' Emotional Reactions to Programming Projects in Introduction to Programming: Measurement Approach and Influence on Learning Outcomes," *Proceedings of the 2017 ACM Conference on International Computing Education Research*, August 2017, http://www.alexlishinski.com/publication/icer-2017-emotions/, Last accessed on 05/01/2018.

[25] Eckerdal, A., McCartney, R, Mostrom, J. E., Sanders, K., Thompson, L. and Zander, C. *From Limen to Lumen:Computing students in luminal spaces.* ICER 2007.

[26] Ahmad Rauf Subhani, Likun Xia, Aamir Saeed Malik, "EEG Signals to Measure Mental Stress," http://www.ipedr.com/vol40/002-ICPSB2012-B00034.pdf, Last accessed on 05/01/2018.

I have submitted this thesis in partial fulfillment of the requirements for the degree of Master of Science.

1/4/19

Date

Gopi Satya Sainadh Raju Sarikonda

We approve of the thesis of Gopi Satya Sainadh Raju Sarikonda as presented here.

12-8-2018

Date

Lydia Ray

Lydia Ray, Ph.D.
Associate Professor, Thesis Advisor

12/12/2018

Date

Alfredo Perez, Ph. D.
Assistant Professor

12/11/18

Date

Shamim Khan, Ph.D.
Professor

12/13/18

Date

Yesem Peker, Ph.D.
Associate Professor

1/2/19

Date

Wayne Summers, Ph.D.
Distinguished Chairperson
Professor of Computer Science