An Insight Into the Relationship Between Confidence, Self-efficacy, Anxiety and Physiological Responses in a CS1 Exam-like Scenario

Keith Nolan
Maynooth University
Kildare, Ireland
keith.nolan@mu.ie

Susan Bergin
Maynooth University
Kildare, Ireland
susan.bergin@mu.ie

Aidan Mooney
Maynooth University
Kildare, Ireland
aidan.mooney@mu.ie

ABSTRACT

Computer Science typically has one of the highest attrition rates in tertiary level education. Many reasons have been put forward as to the cause, including, for example, no prior formal experience of programming, high workloads and poor mental health of students. Recent advancements in wearable technology have made it possible to accurately and easily measure aspects of physiological response associated with emotional arousal which can be indicative of stress, such as heart rate variability and skin conductance. A novel opportunity now exists to monitor learners in real-time and gain an insight on their physiological responses during a learning task. Such information, perhaps coupled with known factors that influence learning success, could provide new insight to allow educators to better tailor module design, delivery, and assessment.

This paper builds on a study which concluded that there was no correlation between self-reported anxiety and Heart Rate and Electrodermal Activity. The goal of this paper is to investigate the relationship between measures such as self-reported anxiety, programming self-efficacy, confidence in responses and physiological responses during a controlled exam-like setting. An out-of-the-box psychological test was used to measure self-reported anxiety, a well-established questionnaire was used to measure self-efficacy and wearable sensors were used to measure physiological arousal, before and during the exam. Study design and methodology are described in detail in this paper. While no significant results were found, perhaps the most interesting finding is that students confidence in their answers weakly correlates with their physiological response when completing multiple choice programming questions. While the findings presented may not be major, they are novel and will provide direction for future research in the area.

CCS CONCEPTS

• **Social and professional topics → CS1**;

KEYWORDS

Confidence, self-efficacy, physiological signals, CS1, anxiety

1 SETTING THE SCENE

CS1 is the first CS module that students are exposed to in most Computer Science tertiary level courses and forms the foundations for many subsequent modules. CS1 typically focuses on the fundamental concepts of Computer Science programming. These basics typically include the syntax of a language, different decision expressions, different method applications and constructs associated with the language. In our university CS1, taught through Java, includes concepts ranging from creating a first program to the development of static methods within a class, in line with ACM curricula standards [9]. Within this range of concepts there are certain concepts that appear to be easy for students such as system output and variable assignment. In contrast there are concepts such as nested loops and arrays which year on year cause problems [22].

In Ireland, the non-progression rate for first year Computer Science students is currently at 25%, the highest among all disciplines in higher education [16]. This non-progression rate typically has multiple cited causes. For example, students typically have no formal exposure to CS prior to CS1 and this leads to concepts being difficult and hard to grasp [7]. These high attrition rates are mimicked in the US. In 2005, Beaubouef et al. investigated why there was such a significant attrition rate in their university citing reasons such as poor problem solving skills and poorly designed CS1 courses [2]. They outline in the first year of their CS course that there are approximately 400 registered students, however, by their final year, roughly only 20–30 students graduate with a CS degree.

Petersen et al. investigated why students drop out and suggested that reasons for dropping out are usually a set of interrelated factors. In the study, people who dropped out cited reasons such as changing priorities, falling behind in class and the intricate level of detail that is needed amongst others [24].

Concern for the mental health of students is growing internationally with many interventions being developed to aid students [15]. In 2012 Headstrong conducted the *My World Survey* which outlined that in any given 100 students, irrespective of subject, 40 suffer from depression and 38 from anxiety with the three main stressors identified as, college, money and work [12]. Nolan et al. conducted a comprehensive literature review of the role of anxiety

when learning to program [21] which drew on the literature, and identified a multitude of factors related to student stress including, task complexity, the modality of programming assessments and general anxiety of using a computer. In addition to this, Nolan et al. conducted a survey with 182 first year Computer Science students showing that Computer Science students in general are more anxious that the average university student [23].

Understanding when a student is stressed or anxious could allow educators to put interventions in place to aid these students in a meaningful way. With physiological sensors becoming more accessible to the general public through innovations such as smart watches and wearable technology, the use of sensors in the classroom can allow educators key insights into how a student is engaging with a class and with the course material. It is well known that when a student becomes stressed they begin to disengage with the material being presented in the class and so, as educators, knowing when this happens for students is critical so that appropriate interventions can be put in place [28, 32].

2 BACKGROUND AND RELATED WORK

2.1 Anxiety in Computer Science

Recent findings indicate that 38 out of 100 college students have experienced some sort of anxiety during their college degree. In a systematic literature review conducted by Nolan et al., CS students were found to be notably anxious,and this is caused by a multitude of factors [12, 21].

A longitudinal study by Connolly et al., conducted with undergraduate students learning to program during their first year of study showed that novice programmers experience increased anxiety when they receive errors in their code [8]. The purpose of their study was to investigate the variance of anxiety within the students throughout the year. This was done by conducting a survey which measured anxiety levels both before and after the students first year of CS.

Scott et al. conducted a study with two hundred and thirty-nine students [29]. Of note, they found that students often worry when completing debugging tasks and that they would start to feel anxious when they try to find and fix programming bugs.

A relationship between programming-task complexity and anxiety was found as a result of a study conducted by Chang [6]. The study was survey based and measured the perceived task complexity and self-reported anxiety levels. The perceived task complexity levels were measured through an in-house designed questionnaire and self-reported anxiety levels were measured using the Computer Attitude Scale [20]. The study consisted of 307 participants and results were based on three different levels of programming task complexity from easy to hard. Results showed that there was a significant relationship between perceived programming task complexity and anxiety levels, that is, as the perceived programming task complexity increased, so too did the perceived anxiety levels.

2.2 Physiological Sensor Technology

Given the advancements recently in sensor technology, it is now easy for people to access and monitor some of their physiological signals, for example their heart rate [25, 27].

A Photoplethysmogram (PPG) is a sensor which detects changes in the volume of blood flow by measuring the difference in the light reflected into the sensor. Heart rate-based algorithms can used to determine measures such as Heart Rate Variability (HRV), which is a known indicator of the interplay between the sympathetic and parasympathetic nervous system [13]. This interplay can provide an insight into the fight-fright response in the body. If there is more variability in the beat-to-beat data, the person could have perceived a threat to an external stimulus, and so the systems that controls the beating of the heart begin to fight for control. Using this knowledge, HRV measures can be used as an indicator of emotional arousal [18].

Electrodermal Activity (EDA) is one of the most commonly used measures for physiological response, with studies focusing on a variety of tasks from measuring attention to predicting abnormal behaviours [5]. EDA sensors work by measuring the resistance between two points on the skin. The Skin Conductance Response (SCR), a method of quantifying EDA, also known as the Electrodermal Response is the phenomenon that the skin momentarily becomes a better conductor of electricity when either external or internal stimuli occur that are physiologically arousing. Measures such as EDA and HRV have been used in the classroom to detect when students begin to lose engagement and how certain activities in the classroom may re-engage students.

McNeal et al. investigated the use of EDA over the course of a lecture to measure the engagement of a class during three key parts: the lecture, the movie and the dialogue [19]. Seventeen students participated in the study. These students wore an EDA sensor for the duration of the class and also filled out pre-class and post-class questionnaires on the teaching concept. From the study, they found that, a) students were more engaged during the movie than any other part of the class, and, b) the higher the emotional arousal of the participants, the more engaged they were with the material. In support of the use of EDA sensor technology, a study by Pho et al. found that the use of EDA sensor technology can show when a participant begins a stressful task as well as when the participant finishes the stressful task [25].

As well as the use of EDA, PPG has been used to detect emotional arousal while learning. A recent study [11] examined student's heart rate during a 50 minute lecture. Findings from the data suggest that student's heart rates dropped during the lecture indicating students were becoming disengaged. Following a question from a student in the class, heart rates increased likely induced by the alternate stimulus.

2.3 Self-efficacy in Computer Science

In recent years, self-efficacy has been a topic of increased research in the area of Computer science and in particular, first year Computer Science. Bandura defined perceived self-efficacy as "people's judgments of their capabilities to organize and execute courses of action required to attain designated types of performances" [1]. Self-efficacy mediates between an individual's knowledge and their actions. Somebody may possess the knowledge or skills which are necessary to perform a particular task, however, they may not succeed due to their own self doubt or self belief in their ability. Self-efficacy theory has emerged as an important means of understanding and predicting a person's performance. Bergin showed repeatedly that self-efficacy is the most significant predictor of

programming success [3, 4]. Quille et al., in several re-validation studies of Bergin's work, showed that self-efficacy was again a main predictor of success in first year computer science students [26]. Research has shown that those with high self-efficacy are more likely to undertake tasks that are more challenging and demonstrate a desire to learn and engage with material. In addition to this, the succession of a task will most likely increase the self-efficacy of the person involved. In contrast, those with low self-efficacy tend to experience tasks that are easy or at least easier as much more difficult that they actually are. This leads to the person experiencing forms of stress or possibly leading to anxiety. This would lead to the student not succeeding in solving the task and effectively reduce their self-efficacy.

2.4 Confidence and Perception

The use of multiple-choice questions (MCQs) is perhaps the quickest and easiest way to examine students knowledge within CS1 classes due to the generally large number of students within the classes. This form of assessment however typically allows the student to guess an answer and possibly guess a correct response without knowing the correct answers. As educators we have no way to gauge the confidence of the students answer and therefore we don't know if the student is guessing, is misinformed or is uninformed. By considering the students confidence in their answer if incorrect, we as educators might be able to determine if a student is uninformed (incorrect and not confident) or misinformed (incorrect but confident) [10]. Knowing when a student is either misinformed of uninformed can provide educators with information on how best to support individual students.

3 RESEARCH QUESTIONS

This paper attempts to answer the following research questions:

RQ1 : Does a students self-efficacy score before an exam align with their confidence in their answers throughout a MCQ?

RQ2 : Is there a relationship between a students physiological responses and their confidence when answering MCQ questions?

RQ3 : Do self-reported anxiety values align with participants confidence throughout an MCQ?

4 INSTRUMENTS

4.1 Background Survey

As part of this experiment, age, gender and other attributes pertaining to the setup of the experiment such as a students dominant hand were collected. The State Trait Anxiety Inventory (STAI), arguably the most commonly used tool in the evaluation of anxiety, was used to gather the students self-assessed anxiety level prior to the experiment [30]. State anxiety is defined as an unpleasant emotional arousal in face of threatening demands or dangers. A cognitive appraisal of threat is a prerequisite for the experience of this emotion [17]. Trait anxiety refers to the tendency to attend to, experience, and report negative emotions such as fears, worries, and anxiety across many situations [14]. The higher the Trait anxiety measure, the more susceptible somebody is to experience anxiety, i.e. somebody with high trait anxiety might respond negatively

to a stimulus whereas somebody with low Trait anxiety may not respond negatively to a similar stimulus.

As part part of this experiment, we included a programming self-efficacy questionnaire developed by Bergin et al. and re-validated by Quille et al. [3, 26].

4.2 Programming Comprehension Questions

As part of this experiment, a multiple choice programming quiz was designed. The questions in the quiz were designed around the course content of the CS1 module taught in the University. For a more in-depth review of the development and validation of the programming questions, see our prior work [22].

5 DESIGN

5.1 Participation

Participants were gathered on a voluntary basis from the CS1 student population and ethical approval was sought and granted to carry out this research.

5.2 Experimental Protocol

Participants were instructed to read an information sheet describing the experiment prior to commencement. If they had any issues or questions they were encouraged to ask for clarification. Once they completed reading the information sheet, they were asked to sign a consent form. The background survey, STAI and self-efficacy survey were then given to the participant for completion.

Following the completion of the questionnaires, the protocol remained the same as described in [22]. The one alteration to the protocol described before, was that two crucial questions were asked after each programming question. Participants were asked to first rate their confidence in their answer on a scale of "Not Confident" , "Slightly Confident", "Somewhat Confident", "Confident" and "Very Confident". Secondly , participants were asked to rate the difficulty of the programming question. The possible options were "Easy", "Medium" or "Hard".

5.3 Physiological Sensors

Participants wore two physiological sensors for the duration of the experiment: an EDA sensor and a PPG sensor. Both of these sensors were part of the Shimmer 3 GSR+, which is a wireless device used to record EDA and PPG signals.

Both the EDA and PPG were sampled at 51.2Hz to ensure subtle changes in the signals could be captured. The Shimmer was placed on the wrist of the non-dominant hand to ensure the participants could use the mouse and doodle with a pen and paper. The PPG sensor was placed on the tip of the middle finger with the EDA sensors placed on the tips of the index and ring fingers of the non-dominant hand.

6 FINDINGS

6.1 Participant Profile

Forty participants (28 male, 12 female) participated in this study. Table 1 presents the age and gender breakdown of the participants.

6.1.1 State Trait Anxiety Inventory. Self-reported State and Trait anxiety factors were collected before the experiment and Table 2

Table 1: Age and gender profile of participants

Age	Male (N=28)	Female (N=12)
17 - 19	24 (86%)	8 (67%)
20 - 22	0 (0%)	3 (25%)
23+	4 (14%)	1 (8%)

shows the State and Trait results of the participants broken down by gender.

Table 2: Average State, Trait and programming self-efficacy values for male and female participants

	Male	Female	p-value
State	49.82	52.83	0.155
Trait	54.32	55.5	0.38
Self-efficacy	29.74	28	0.257

Results were calculated using the scoring key provided with the STAI. Normalised results were obtained from the STAI manual [31]. The STAI Manual outlines normal State and Trait values for a normal college population broken down by gender. The manual reports that average State values for male and female students is 36.47 and 38.76 out of 100 respectively. The averages found in this study are considerably higher than the normative averages for College Students. The Trait averages for male and female students in a normal population are 38.30 and 40.40 respectively. Interestingly, both the State and Trait levels observed in this study are considerably higher than the normal values as outlined in the manual and supported by Nolan et al.'s survey of First Year Computer Science students [23]. Having a higher self-reported Trait anxiety score may make the student more likely to experience anxiety in situations.

6.1.2 Programming Self-efficacy. The programming self-efficacy scale was marked in the range of $10 - 40$. Higher values indicate higher levels of self-efficacy in programming. Table 2 shows self-efficacy broken down by gender and it can be observed that male students have an average score of 29.74 compared to an average score of 28 for females. One male failed to complete the survey correctly and so was excluded from the analysis. This result is in line with previous research in the area of self-efficacy with male students reporting higher self-efficacy for programming.

7 RESEARCH QUESTIONS ANSWERED

7.1 Does a students self-efficacy score before an exam align with their confidence in their answers throughout a Multiple Choice Questionnaire?

To approach this question, the self reported self-efficacy was compared against the participants confidence responses to the questions in the experiment. In order to gain an initial understanding of the data, a Pearson's correlation test was run over all variables. A strong correlation would return a value of $r >= 0.7$ or $r <= -0.7$. When the correlation test was run over the participants self-efficacy and all 9 confidence values, no strong correlations were found with values ranging from $0.13 <= r <= 0.30$ suggesting that there were little to no correlation relationships.

Given that there was no clear relationship between confidence and self-efficacy, the question was asked, how could confidence in the participants own ability before the experiment not align with their confidence during the experiment? A deeper relationship had to be explored between the variables. To begin this process, the nine confidence values were condensed into one value using Principal Component Analysis (PCA). This was done to allow for comparison across all 9 confidence values. An important point here is that the combined confidence values are in the range of $-1.617 - 5.434$. While these values seem random, the PCA algorithm normalises the results. While this is a large spread, the more negative a value is, the more confident someone was in their responses overall. It is important to note that if the participant is extremely confident (answered "Very Confident" on every question) this does not translate into a PCA value of -9. Figure 1 shows a scatter plot of the self-efficacy and combined confidence values. The trend line shows a downward trend however this is deceptive given that the more negative the confidence value the more confident the participant was. Therefore the trend line that is shown would suggest that the higher ones self-efficacy is the more confident one is in their responses. A Pearson's correlation was examined between the participants self-efficacy and the combined value for confidence. This returned a value of $r = -0.285$ suggesting that there is a weak negative correlation. While this correlation was weak it was significant at an $\alpha = 0.05$ level ($T = 1.83$).

Figure 1: Scatter plot of participants self-efficacy and their combined confidence values. In addition, the trend line within the graph would suggest a negative relationship.
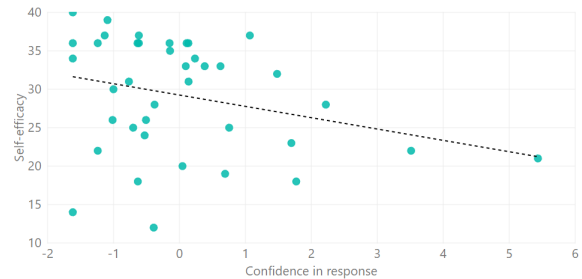

While the results are weak and would require further experimentation and investigation, there is a weak trend. This trend seems to suggest that the more confident a student is in their programming ability, the more likely they are to be confident in their responses to programming questions. While this may seem intuitive and obvious, the fact that the correlation is not strong would suggest that there may be some other factor/s influencing the relationship between self-efficacy and confidence in response.

7.2 Is there a relationship between a students physiological responses and their confidence when answering MCQ questions?

Given that the physiological responses that were collected are Heart Rates and Electrodermal Activity, both signals were analysed separately. Given the nature of collecting wireless sensor data, six sets of Electrodermal Activity and Heart Rate data were either incomplete or corrupt and so were disregarded for the analysis.

7.2.1 Confidence and Heart Rate. Similarly to Research Question 1, an approach to correlate the data was taken. This was done to inspect if there is a general (normal) relationship between the data. A Pearson's correlation test was conducted. As Figure 2 shows, there is no obvious relationship between Confidence and Heart Rate. The correlation test, there was a weak correlation with a value of $r = 0.177$. This correlation is shown in Figure 2 by the positive trend line showing the slight positive relationship between Confidence and Heart Rate. This correlation was not significant at the $\alpha = 0.05$ level. Again, given that there is a somewhat weak correlation, perhaps with further experimentation and investigation stronger results might be found. However there appears to be a slight relationship between confidence and heart rate where the more confident one is the lower their heart rate is.

Figure 2: Scatter plot of participants Heart Rate and their combined confidence values. In addition, the trend line within the graph would suggest a positive relationship between Heart Rate and Confidence.
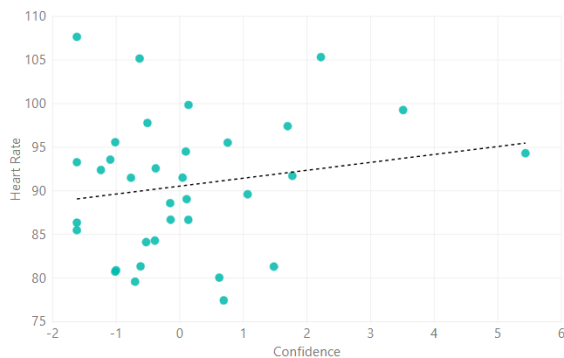
7.2.2 Electrodermal Activity. Similar to the case of confidence and Heart Rate, a Pearson's correlations between Confidence and Electrodermal Activity was calculated. Initially there was a correlation of $r = 0.1$ which indicates a weak correlation and no real relationship. This correlation was not significant. Figure 3 shows the scatter plot of the two variables along with a trend line depicting a slight positive relationship. In addition to this, it is hard to suggest that the lower the Electrodermal Activity the more confident you are given the influence of the outliers and weak correlation between Confidence and Electrodermal Activity. However with further investigation, perhaps a relationship can be found.

Figure 3: Scatter plot of participants Electrodermal Activity and their combined confidence values. The trend line overlaid on the graph shows a slightly positive relationship between the factors however this may be skewed due to the outliers.
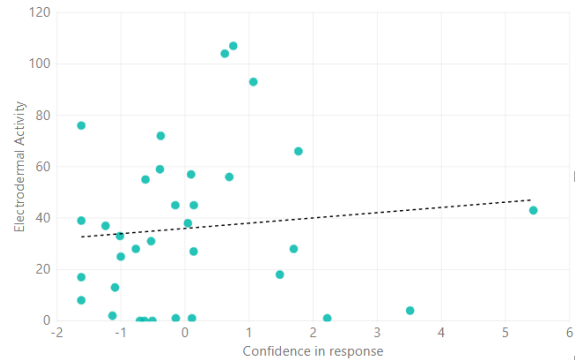
7.3 Does self-reported anxiety values align with participants confidence throughout an MCQ?

Following the trends from the previous research questions, State and Trait anxiety were correlated with the PCA confidence to examine if there existed a relationship. As there were two distinct variables, State anxiety and Trait anxiety, two correlation tests were run, namely Confidence values vs State and Confidence values vs Trait.

7.3.1 State Anxiety. A Pearson's correlation was run on Confidence values and State anxiety, which returned a very weak correlation of $r = 0.035$. This suggests that there was little to no correlation between Confidence and State anxiety. This correlation is seen in the trend line in Figure 4.

As can be seen, the trend line shows no real relationship between the variables. This correlation was not significant either. Given the fact that there was an extremely weak correlation coupled with a non-significant correlation, there is no way to hypothesize a type and direction of relationship between State anxiety and Confidence.

Figure 4: Scatter plot of participants State Anxiety and their combined confidence values. The trend line overlaid on the graph shows no significant relationship between the factors.

7.3.2 Trait Anxiety. Given that State and Trait anxiety are closely correlated ($r = 0.842$) it was expected that a similar result to the State anxiety vs Confidence would be found here. A Pearson's correlation was run on Confidence values and Trait anxiety. This returned a weak but better correlation of $r = 0.1435$ when compared to the correlation between Confidence and State anxiety ($r = 0.035$). Figure 5 shows the scatter plot of Trait anxiety and Confidence with the trend line showing a weak correlation. The trend line is positively sloped which shows the positive relationship between the variables. This correlation is not significant however. While there is a slightly stronger relationship between Trait anxiety and Confidence, there is still a need for further investigation into the area however it does appear that there may be a relationship between Trait anxiety and Confidence.

Figure 5: Participants Trait Anxiety and Confidence values displayed on a scatter plot. The trend line on the graph shows the positive relationship between the factors found.
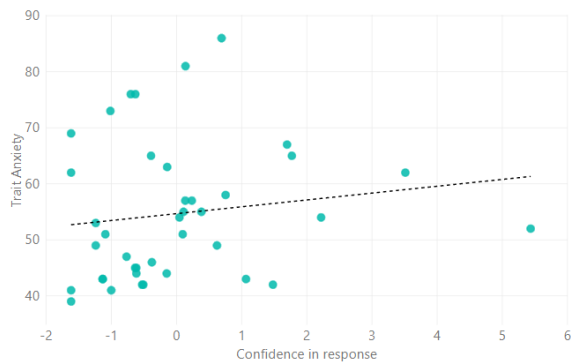
8 THREATS TO VALIDITY

While every precaution was taken when designing and setting up this study, the following are some potential limitations:

(1) All correlations presented here are weak and many are not significant. While this would suggest that there is no relationship between the variables, further study focused on different profiles of students would be of benefit.

(2) The students who participated in this study were self-selecting students. With this in mind, the students who are represented in the study may be more confident in their programming ability and on average are higher performing and therefore results should be interpreted with care.

(3) This experiment was carried out with just the participant and the researcher in the room, a further study would be valuable in a more authentic exam-like setting.

(4) This physiological signals that were collected were for a short period of time. Allowing the participant to wear the wireless technology before the experiment to obtain a longer and more accurate baseline might provide more significant findings.

9 CONCLUSION

This paper describes a study which is an adaption of previous work [22]. The study was designed to examine if a relationship can be found between student confidence in their responses to programming questions with Programming Self-efficacy, Electrodermal Activity, Heart Rate and State and Trait anxiety.

Based on the findings from Research Question 1, there were no significant correlations between Programming Self-Efficacy and confidence of response to each question, but a weak negative correlation between Self-efficacy and confidence in question responses combined exists. While this result might be somewhat intuitive and expected, it is important to report on the result. This finding should be the starting point of further research.

Research Question 2 examined the relationship between confidence in student's answers and their physiological responses. Similarly to Research Question 1, when the correlations were tested, there were extremely weak correlations observed. This is somewhat unsurprising. Intuition would lead us to believe that the more relaxed (low heart rate and low sweat rate) you are during an exam, the more confident you should be in your answers. The positive correlations observed here do support this intuition however, caution should be taken with this. Further investigation is required to strengthen this.

Research Question 3 explored the relationship between a student's confidence in their answers and their psychological responses. Given that there are two psychological variables, State and Trait anxiety, they were tested independently. The correlation tests that where conducted showed that there was a weak relationship between both State and Trait anxiety and confidence in responses. This is perhaps unsurprising as one could be confident in their responses to questions but still anxious due to the nature of the setting, for example a multiple choice questionnaire.

This work, while no major findings are reported, is a novel contribution to the community as it can now be used to direct future research in the area. Aspects of this research are interesting and promising, however further research is required. A second larger study would prove valuable in this case. Given the low number of students participating in the study and the self selecting nature of the participants, a larger and broader subset of the First Year CS community might give more significant findings. With more significant findings, it would also be possible to suggest methods that might increase retention rates within the First Year CS community.

REFERENCES

[1] Albert Bandura. 1986. Social foundations of thought and action. *Englewood Cliffs, NJ* 1986 (1986).

[2] Theresa Beaubouef and John Mason. 2005. Why the High Attrition Rate for Computer Science Students: Some Thoughts and Observations. *SIGCSE Bull.* 37, 2 (June 2005), 103–106. https://doi.org/10.1145/1083431.1083474

[3] Susan Bergin and Ronan Reilly. 2005. Programming: Factors That Influence Success. In *Proceedings of the 36th SIGCSE Technical Symposium on Computer Science Education (SIGCSE '05)*. ACM, New York, NY, USA, 411–415. https://doi.org/10.1145/1047344.1047480

[4] Susan Bergin and Ronan Reilly. 2006. Predicting introductory programming performance: A multi-institutional multivariate study. *Computer Science Education* 16, 4 (2006), 303–323. https://doi.org/10.1080/08993400600997096 arXiv:https://doi.org/10.1080/08993400600997096

[5] John T Cacioppo, Louis G Tassinary, and Gary Berntson. 2007. *Handbook of psychophysiology*. Cambridge University Press.

[6] Sheeson E. Chang. 2005. Computer anxiety and perception of task complexity in learning programming-related skills. *Computers in Human Behavior* 21, 5 (September 2005), 713–728. https://doi.org/10.1016/j.chb.2004.02.021

[7] C. Connolly and E. Murphy. 2005. Retention initiatives for ICT based courses. In *Proceedings Frontiers in Education 35th Annual Conference*. S2C–10. https://doi.org/10.1109/FIE.2005.1612215

[8] Cornelia Connolly, Eamonn Murphy, and Sarah Moore. 2009. Programming Anxiety Amongst Computing Students - A Key in the Retention Debate? *IEEE Transactions on Education* 52, 1 (February 2009), 52–56. https://doi.org/10.1109/TE.2008.917193

[9] The Joint Task Force on Computing Curricula, Association for Computing Machinery, and IEEE-Computer Society. 2013. *Computer Science Curricula 2013: Curriculum Guidelines for Undergraduate Degree Programs in Computer Science*. ACM, Inc. https://doi.org/10.1145/2534860

[10] Donald A Curtis, Samuel L Lind, Christy K Boscardin, and Mark Dellinges. 2013. Does student confidence on multiple-choice question assessments provide useful information? *Medical Education* 47, 6 (2013), 578–584. https://doi.org/10.1111/medu.12147 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/medu.12147

[11] Diana Darnell and Paul Krieg. 2014. Use of heart rate monitors to assess student engagement in lecture (721.25). *The FASEB Journal* 28, 1_supplement (2014), 721–25.

[12] Barbara A Dooley and Amanda Fitzgerald. 2012. *My world survey: National study of youth mental health in Ireland*. Headstrong and UCD School of Psychology.

[13] Gernot Ernst. 2017. Heart-Rate Variability - More than Heart Beats? *Frontiers in Public Health* 5 (Sep 2017). https://doi.org/10.3389/fpubh.2017.00240

[14] Yori Gidron. 2013. *Trait Anxiety*. Springer New York, New York, NY, 1989–1989. https://doi.org/10.1007/978-1-4419-1005-9_1539

[15] Margaret Grigg and Shekhar Saxena. 2004. Promoting mental health nursing research in low and middle income countries. *International Nursing Review* 51, 4 (December 2004), 194–195. https://doi.org/10.1111/j.1466-7657.2004.00268.x

[16] David Harmon and Stephen Erskine. 2017. Eurostudent Survey VI. http://hea.ie/assets/uploads/2018/01/HEA-Eurostudent-Survey.pdf

[17] Theodore D. Kemper and Richard S. Lazarus. 1992. Emotion and Adaptation. *Contemporary Sociology* 21, 4 (1992), 522. https://doi.org/10.2307/2075902

[18] Richard D. Lane, Kateri McRae, Eric M. Reiman, Kewei Chen, Geoffrey L. Ahern, and Julian F. Thayer. 2009. Neural correlates of heart rate variability during emotion. *NeuroImage* 44, 1 (2009), 213 – 222. https://doi.org/10.1016/j.neuroimage.2008.07.056

[19] Karen S. McNeal, Jacob M. Spry, Ritayan Mitra, and Jamie L. Tipton. 2014. Measuring Student Engagement, Knowledge, and Perceptions of Climate Change in an Introductory Environmental Geology Course. *Journal of Geoscience Education* 62, 4 (Nov 2014), 655–667. https://doi.org/10.5408/13-111.1

[20] Gary S. Nickell and John N. Pinto. 1986. The computer attitude scale. *Computers in Human Behavior* 2, 4 (1986), 301 – 306. https://doi.org/10.1016/0747-5632(86)90010-5

[21] Keith Nolan and Susan Bergin. 2016. The Role of Anxiety when Learning to Program: A Systematic Review of the Literature. In *Proceedings of the 16th Koli Calling International Conference on Computing Education Research (Koli Calling '16)*. ACM, New York, NY, USA, 61–70. https://doi.org/10.1145/2999541.2999557

[22] Keith Nolan, Aidan Mooney, and Susan Bergin. 2019. An Investigation of Gender Differences in Computer Science Using Physiological, Psychological and Behavioural Metrics. In *Proceedings of the Twenty-First Australasian Computing Education Conference (ACE '19)*. ACM, New York, NY, USA, 47–55. https://doi.org/10.1145/3286960.3286966

[23] Keith Nolan, Aidan Mooney, and Susan Bergin. 2019. A Picture of Mental Health in First Year Computer Science. In *Proceedings of the 10th International Conference on Computer Science Education: Innovation and Technology (CSEIT'19)*. Global Science and Technology Fourm.

[24] Andrew Petersen, Michelle Craig, Jennifer Campbell, and Anya Tafliovich. 2016. Revisiting Why Students Drop CS1. In *Proceedings of the 16th Koli Calling International Conference on Computing Education Research (Koli Calling '16)*. ACM, New York, NY, USA, 71–80. https://doi.org/10.1145/2999541.2999552

[25] M. Poh, N. C. Swenson, and R. W. Picard. 2010. A Wearable Sensor for Unobtrusive, Long-Term Assessment of Electrodermal Activity. *IEEE Transactions on Biomedical Engineering* 57, 5 (May 2010), 1243–1252. https://doi.org/10.1109/TBME.2009.2038487

[26] Keith Quille, Natalie Culligan, and Susan Bergin. 2017. Insights on Gender Differences in CS1: A Multi-institutional, Multi-variate Study.. In *Proceedings of the 2017 ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE '17)*. ACM, New York, NY, USA, 263–268. https://doi.org/10.1145/3059009.3059048

[27] Brian K Sandall. 2016. Wearable Technology and Schools: Where are We and Where Do We Go From Here? *Journal of Curriculum, Teaching, Learning and Leadership in Education* 1, 1 (2016), 9.

[28] Carmen Sandi. 2013. Stress and cognition. *Wiley Interdisciplinary Reviews: Cognitive Science* 4, 3 (Jan 2013), 245âĂŞ261. https://doi.org/10.1002/wcs.1222

[29] Michael James Scott and Gheorghita Ghinea. 2014. Measuring Enrichment: The Assembly and Validation of an Instrument to Assess Student Self-beliefs in CS1. In *Proceedings of the Tenth Annual Conference on International Computing Education Research (ICER '14)*. ACM, New York, NY, USA, 123–130. https://doi.org/10.1145/2632320.2632350

[30] Charles D Spielberger. 1989. *State-Trait Anxiety Inventory: Bibliography (2nd ed.)*. CA: Consulting Psychologists Press.

[31] Charles D Spielberger, R. Gorsuch, R. L.and Lushene, P. R. Vagg, and G. A Jacobs. 1983. *Manual for the State-Trait Anxiety Inventory*. CA: Consulting Psychologists Press.

[32] Habib Yaribeygi, Yunes Panahi, Hedayat Sahraei, Thomas P. Johnston, and Amirhossein Sahebkar. 2017. The impact of stress on body function: a review. *EXCLI Journal; 16:Doc1057; ISSN 1611-2156* (2017). https://doi.org/10.17179/excli2017-480