# Exploring Potential Relationships Between SE, Performance, and Electrodermal Arousal in Engineering Exams

**Proposal Submitted to:**
**The American Educational Research Association**
**July 2018**

**Abstract**

This exploratory, empirical study examined the potential relationships between self-efficacy and electrodermal activity (a measure of physiological arousal) within the context of problem performance in an engineering exam. Physiological arousal is commonly discussed as a source of SE (Bandura & Adams, 1977), which is known to be important for managing emotions (Bandura, 1994) and predicting exam scores (Gaylon, Blondin, Yaw, Nalls, & Williams, 2011). Yet, the dynamic relationship between SE and EDA in the context of authentic tasks (i.e., exam problems) is seldom explored. Preliminary findings suggests that for erroneously answered exam problems, SE and EDA were negatively correlated; no correlation was found for the opposite scenario. These findings point to the role actual exam problem performance may have at influencing SE and EDA relationships.

## Theoretical Framework

### *Self-efficacy*

Self-efficacy (SE) is a perceived set of beliefs held by a person about their ability to produce a given level of performance in a specific task (Bandura, 1994). When students have high SE, they are able to approach difficult tasks and problems by maintaining a strong commitment, withstand setbacks, and overcome failures (Bandura, 1994). Those with high SE have high expectations for themselves and set high goals, visualizing success to fight off doubt (Bandura, 1994). They are better able to manage emotions and believe they have a high level of control over potential threatening situations (Bandura, 1994). SE also relates to the anxiety that could come with tests, but does not necessarily moderate it (Barrows, Dunn, & Lloyd, 2013).

Gaylon et al. (2011) conducted a study with 165 students in an undergraduate educational psychology course and showed that SE is more related to exam scores than participation in class, especially at high levels of SE. Barrows et al. (2013) performed a study on 110 students across multiple departments (psychology, sociology, communication, business, english, music, and history) and found that exam grades and SE are strongly related.

SE comes from several sources including the vicarious experiences of observing others succeed in their efforts, verbal persuasion to cope, and states of physiological arousal by which people just their anxiety and vulnerability to stress (Bandura & Adams, 1977). Regarding the latter (physiological arousal), limited studies have been conducted to identify its correlations to SE in the context of engineering exams (Villanueva, Goodridge, & Call, 2018; Villanueva, Campbell, Raikes, Jones, & Putney, 2018).

### *Electrodermal Activity*

Electrodermal activity (EDA) is a commonly used biosignal to measure physiological arousal in the form of skin electrical conductivity (Boucsein, 2012). EDA is a measure of the sympathetic (fight/flight system) autonomic nervous system activity (Bouscein, 2012; Boucsein & Backs, 2009; Villanueva, Valladares, & Goodridge, 2016). EDA data can be divided into phasic (stimulus-specific and immediate) and tonic (baseline) forms (Boucsein, 2012). Phasic EDA evidence affective- or emotion-based physiological arousal, which can indicate mental activity or emotional strain (Boucsein & Backs, 2009). These changes can give insight into understanding student performance, emotional responses, and decisions made (Badami, VaezMousavi, Wulf, & Namazizadeh, 2012; Villanueva et al., 2016; Villanueva, Raikes, Ruben, Schaefer, & Gunther, 2014).

Regarding performance, EDA is a function of task difficulty, with lower levels of arousal measured when the problem is more difficult (Hammond & Jordan, 1984). It is assumed that this limited rise in arousal may have to do with inhibitory control (Hammond & Jordan, 1984), which is the ability one has to resist distractions and give full attention to relevant stimuli (Administration for Children and Families, n.d.). This study will begin to explore potential correlations between EDA arousal and SE in the context of exam problem performance.

*SE and EDA Correlations in the Context of Exam Problem Performance*

Although physiological arousal is commonly discussed as a source and outcome of SE (Bandura, 1994), very little research has explored the dynamic relationship between the two. Physiological states are related to self-efficacy beliefs and the beliefs individuals have about their own abilities can effect physiological states (Bandura, 1994). Although there is a significant body of work linking both exam arousal and SE relation to test performance, these studies typically rely on prospective or retrospective self-reports, which limits its closer-to-real-time explorations (Pekrun & B hner, 2014; Frenzel, Pekrun, & Goetz, 2007) of what happens during an exam. Studies are just begin to explore these nearer-to-real time correlations on engineering exams, (Jones, Campbell, & Villanueva, 2016; Villanueva, Campbell, Raikes, Jones, & Putney, 2018) but has yet to provide more granular insights on exam problems. The research hypotheses for this study are: (1) SE and EDA will be positively correlated and (2) the relationship between SE and EDA are influenced by the students' performance on a specific problem.

**Methods**

All procedures were approved by the Utah State University Institutional Review Board (IRB) for studies on human subjects.

*Research Design*

This study is an exploratory empirical study involving customized experimental setup that includes a laptop computer that is interfaced with two web-cameras, a timestamping program created in MATLAB, and a customized exam protocol that recorded student responses as well as the time it took them to respond to the questions. A representative experimental layout can be found in Figure 1.

*Experimental Setup, EDA, and SE*

For the experiment, students were asked to complete the second practice exam in its entirety. Within each exam problem, was a question about SE asking students about their overall ability to perform on the particular exam problem presented. Soon after their response to the SE question, potential multiple-choice solutions to the particular exam problem were presented and students were prompted to work and provide a solution to the problem. During the entire exam period, EDA arousal signals were collected per participant.

Collection of EDA happened at 4 Hz (data collected every of a second) for a ~3-hour exam via an Empatica E4 wrist sensor (Empatica, Boston, MA). The EDA signal was separated into phasic and tonic levels using Ledalab software (Benedek & Kaernback, 2010) that applies continuous decomposition analysis to the raw EDA data. Outliers in the phasic EDA data were determined by extracting 5% of the maximal EDA response as reported previously (Bouscein, 2012; Villanueva, Campbell, Raikes, Jones, & Putney, 2018). SE response were collected at the beginning of each test question. Correct responses to each exam question was collected from the course instructor (who designed the practice and actual exams in the course) and performance data was collected from each student via our custom-created program.

*Exam Context and Participants*

An engineering statics course was chosen for this study because it is typically the first required engineering course students take (Villanueva, Goodridge, & Call, 2018). At the U.S. western university where this study was conducted, there is a higher-than-average non-traditional, first-generation, and rural student enrollment (Office of Analysis, Assessment, & Accreditation, 2018). The statics course has three midterm exams and one final exam. This analysis focused on the second practice exam in Spring 2018, which was around week 8 of a 16 week semester and whose length and difficulty paralleled the actual exam. The practice exam contained 20 multiple choice questions.

Students were recruited to take a practice exam for this research study from an engineering statics course in preparation for their actual exam. As an incentive, students were offered extra credit and a $5 gift card. Attending this study also fulfilled the class assignment requirement of completing a practice examination. The total students for this two-semester study were 90 students across three exams. This study will present the findings for the second exam, representing 27 students.

**Results**

An initial assessment of the time spent per exam question was conducted across the participants to test the effort on each question. The logarithm of the number of minutes spent on each exam was tested against the total number of correct and incorrect exam problem responses. A pooled t-test testing for equal variances was conducted and a Satterthwaite t-test testing for unequal variances was conducted. Findings indicated that there were no statistically different responses between time and number of correct and incorrect responses ($p>0.05$) suggesting that students expended the same amount of time per problem (Table 1).

Student's self-reported SE levels were compared against the number of correct and incorrect exam problem responses (Figure 2). A Satterthwaite t-test to test for unequal variances were conducted. Findings suggest that there is a significant difference between the correct and incorrect response groups ($t = -4.97$; $p < .001$) where SE levels were reported higher for the correct responses. Student responses to the problems were plotted for each exam question to identify the frequencies of correct and incorrect responses by the student group. Approximately 55% of the exam responses were correct. The data set was then split into two parts: (a) problems in where the majority of the students answered them correctly and (b) problems where the majority of the students answered them incorrectly.

For each context (correct exam problem performance vs. incorrect exam problem performance), the logarithm of the median phasic EDA was plotted against the Likert-scale responses from the SE questions. Regression analysis was conducted for the incorrect (Figure 3) and correct responses (Figure 4). Results suggested that for the incorrect responses, EDA was negatively correlated to SE ($r = -.128$; $p < .0001$). For the correct responses, EDA showed a negative correlation to SE but it was not statistically significant ($r = -.071$; $p = 0.055$).

**Discussion**

Time spent on each question was not statistically different between the correct and incorrect exam problems ($p > 0.05$). While other studies have shown that time spent on questions may depend on the difficulty of the question and student response rate (Ma, Agnihotri, Baker, & Mojarad, 2016), it was hard to determine if there was an ideal time for each question in this exam. Future work will explore the difficulty indices of each exam problem as they relate to the effort expended will be conducted. Next, SE levels were significantly higher on the correct exam problems. This is consistent with findings that SE is significantly related to exam scores (Gaylon et al., 2012; Barrows et al., 2013).

In terms of potential correlations between EDA and SE for each exam problem performance context, there was a significant negative correlation for the incorrect responses but not for the correct responses. What is interesting here is that this relationship is only evident when students are performing poorly. This indicates that actual performance may influence the relationship between SE and arousal, which is a unique finding. Future work will explore this phenomenon further with a higher number of participants.

**Conclusions**

The performance on a question may influence what the relationship between EDA and SE since the relationship was only significant when the question was answered incorrectly. This suggests that the relationship between correct answers is more complicated than simply an increase or decrease in EDA. These relations call for the exploration of these relationships on a larger number of participants as well as differing testing scenarios.

**Limitations**

This data analysis was only performed on a small subset of the data collected from this study. Future analysis will include a larger number of participants. The exam is a practice exam, which could change the approach of the students to the exam. The situation was set up to be as authentic as possible but may still not be fully representative of an actual exam on the course.

**Acknowledgements**

This material is based upon work supported by the National Science Foundation Grant No. 1661100 and the National Science Foundation Graduate Research Fellowship Program under Grant No. 120214. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

**References**

Administration for Children and Families (n.d.). *National Survey of Child and Adolescent Well-Being* (Methods Brief No. 1). Retrieved from https://www.acf.hhs.gov/sites/default/files/opre/inhibitory_control.pdf

Alsharif, N. Z. & Qui, Y. (2014). A Three-Year Study of the Impact of Instructor Attitude, Enthusiasm, and Teaching Style on Student Learning in a Medicinal Chemistry Course. *American Journal of Pharmaceutical Education, 78*(7), Article 132, 1-13. doi: 10.5688/ajpe787132

Bandura, A. (1994). SE. In V. S. Ramachaudran (Ed.), *Encyclopedia of Human Behavior* (Vol. 4, pp. 71-81). New York: Academic Press. (Reprinted in H. Friedman [Ed.], *Encyclopedia of Mental Health*. San Diego: Academic Press, 1998).

Bandura, A. (1997). *SE: The Exercise of Control.* New York: W. H. Freeman.

Bandura, A., & Adams, N.E. (1977). Analysis of SE theory of behavioral change. *Cognitive therapy and research*, *1*(4), 287-310. doi: 10.1007/BF01663995

Badami, R., VaezMousavi, M., Wulf, G., & Namazizadeh, M. (2012). Feedback about more accurate versus less accurate trials. *Research Quarterly for Exercise and Sport, 83*(2), 196-203. doi: 10.1080/02701367.2012.10599850

Barrows, J., Dunn, S., & Lloyd, C. A. (2013). Anxiety, SE, and college exam grades. *Universal Journal of Educational Research, 1*(3), 204-208. doi: 10.13189/ujer.2013.010310

Benedek, M. & Kaernbach, C. (2010). A Continuous Measure of Phasic Electrodermal Activity. *Journal of Neuroscience Methods, 190,* 80-91. doi: 10.1016/j.jneumeth.2010.04.028

Boucsein, W. & Backs, R. W. (2009). The Psychophysiology of Emotion, Arousal, and Personality: Methods and Models. In Duffy, V. G. (Ed.), *Handbook of Digital Human Modeling: Research for Applied Ergonomics and Human Factors Engineering* (pp. 35-1-35-18). Boca Raton, FL: CRC Press.

Boucsein, W. (2012). *Electrodermal Activity* (2nd ed.). New York: Springer Science+Business Media, LLC.

Frenzel, A. C., Pekrun, R., Dicke, A. L., & Goetz, T. (2012). Beyond quantitative decline: Conceptual shifts in adolescents' development of interest in mathematics. *Developmental Psychology, 48*(4), 1069-1082. doi: 10.1037/a0026895.

Gaylon, C. E., Blondin, C. A., Yaw, J. S., Nalls, M. L., & Williams, R. L. (2011). The Relationship of Academic SE to Class Participation and Exam Performance. *Social Psychology of Education, 15*(2), 233-249. doi: 10.1007/s11218-011-9175-x

Hammond, G. R. & Jordan, P. M. (1984). Bilateral Electrodermal Activity During Performance of Cognitive Tasks of Varying Difficulty Levels. *Physiological Psychology, 11*(4), 256-260. doi: 10.3758/BF03326804

Office of Analysis, Assessment & Accreditation. (2018). Utah State University. Accessed 23 July 2018.

Ma, Y., Agnihotri, L., Baker, R., & Mojarad, S. (2016). Effects of Student Ability and Question Difficulty on Duration. In T. Barnes, M. Chi, & M. Feng (Eds.). *Proceedings of the 9th International Conference on Educational Data Mining (EDM) 2016* (pp. 143-151). Raleigh, NC: International Educational Data Mining Society.

Pecchinenda, A. & Smith, C. A. (1996). The Affective Significance of Skin Conductance Activity During a Difficult Problem-solving Task. *Cognition and Emotion, 10*(5), 481-503. doi: 10.1080/026999396380123

Pekrun, R., & B hner, M. (2014). Self-report measures of academic emotions. In R. Pekrun & L. Linnenbrink-Garcia, (Eds.), *International handbook of emotions in education* (pp. 561-566). London: Routledge Press.

Villanueva, I., Goodridge, W., & Call, B. (2018). An Initial Exploration of Engineering Students' Emotive Responses to Mechanics and Statics Problems, *Proceedings of the American Society of Engineering Education Annual Conference and Exposition, Mechanical Engineering Division*, Salt Lake City, UT.

Villanueva, I., Raikes, A., Ruben, N., Schaefer, S., & Gunther, J. (2014). The Use of Physiological Tools to Identify Changes in Affective Responses for Graduate Students Recently Admitted into a Scientific Discipline. In *2014 IEEE Frontiers in Education Conference Proceedings* (pp. 1-5). Madrid, Spain. doi: 10.1109/FIE.2014.7044316

Villanueva, I., Valladares, M., & Goodridge, W. (2016). Use of Galvanic Skin Responses, Salivary Biomarkers, and Self-reports to Assess Undergraduate Student Performance

During a Laboratory Exam Activity. *Journal of Visualized Experiments,* 108. doi:10.379153255

Villanueva, I., Campbell, B., Raikes, A., Jones, S., & Putney, L. A. (2018). Multi-Modal Exploration of Engineering Students' Emotions and Electrodermal Activity in Design Activities. *Journal of Engineering Education,* In Press.

Villanueva, I., Christensen, D., Youmans, K., Vicioso, P., Khan, T., Husman, J., & Lampkins, S. (2018). An Experimental Method to Study the Biological, Physiological, and Motivational Elements Experienced by Undergraduate Students in Authentic Examination Activities. Manuscript Submitted for Publication in Journal of Visualized Experiments.

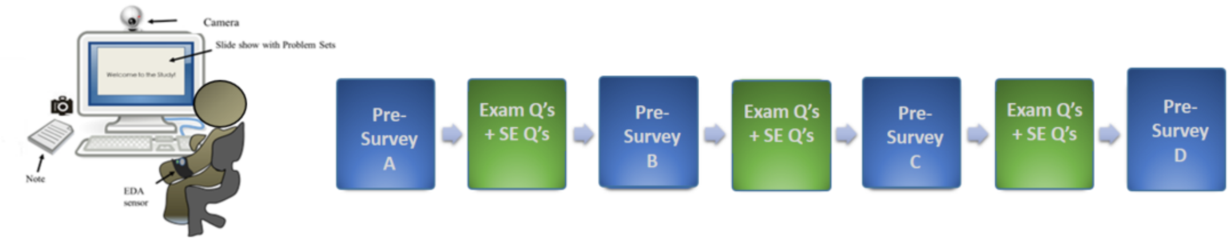
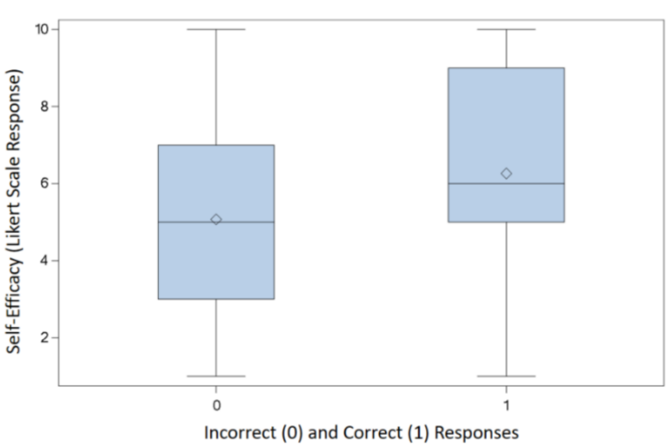



*Figure 1.* Representative diagram of experimental setup; adapted from Villanueva, Goodridge & Call, 2018.

*Table 1.*

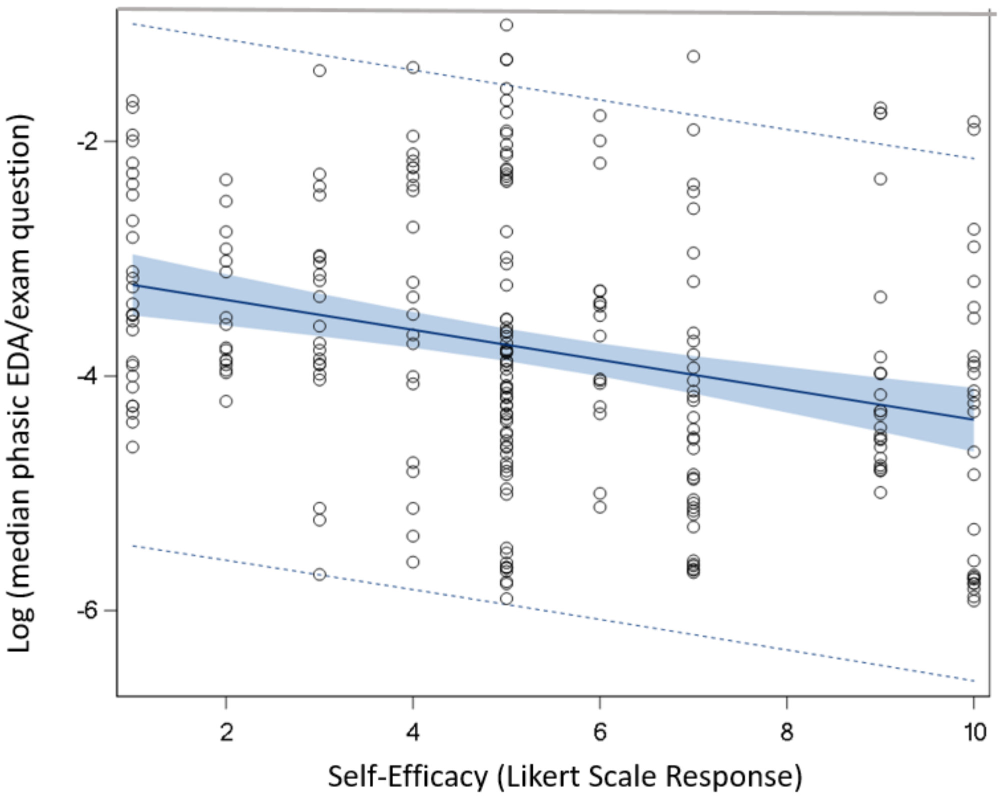
Analysis for the number of minutes spent on the exam problems compared to the number of correct and incorrect responses.

| **Logarithm of number of minutes spent on each exam problem** | | | | | |
|---|---|---|---|---|---|
| **Responses** | **Mean (SD)** | **t-value** | | **p-value** | |
| | | Pooled | Satterthwaite | Pooled | Satterthwaite |
| Incorrect | 1.28 (2.55) | -.73 | -.73 | .46 | .46 |
| Correct | 1.34 (2.68) | | | | |



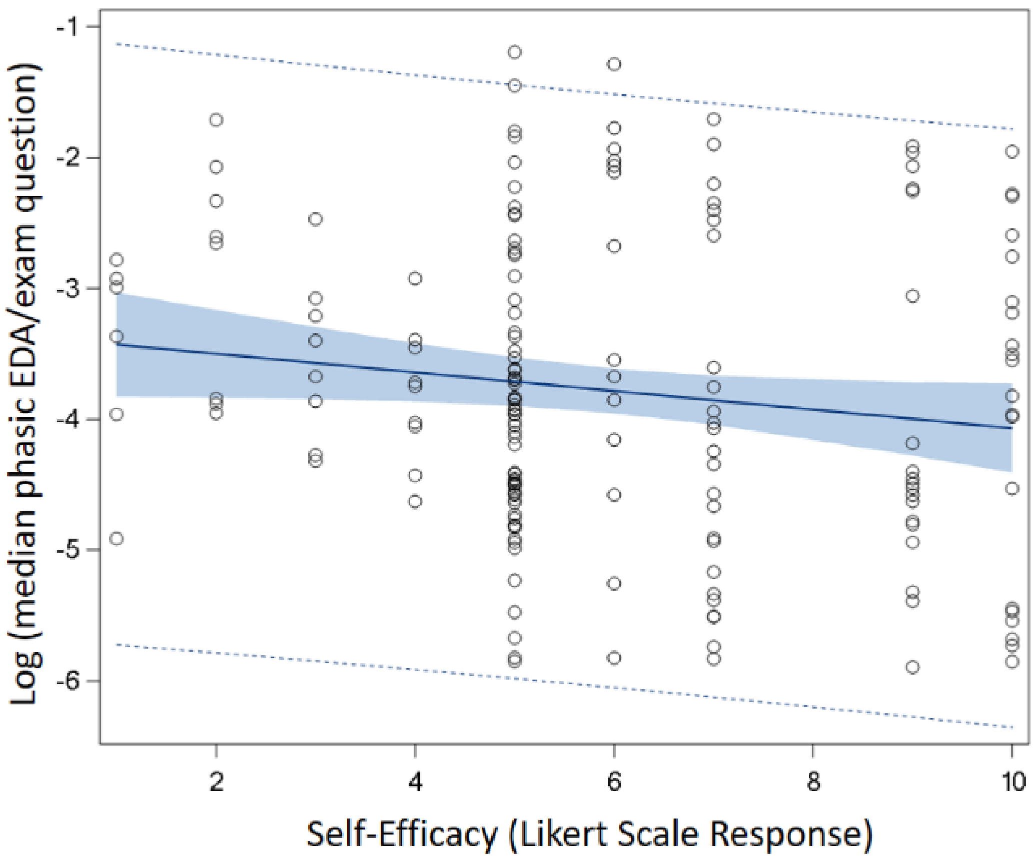

| t Value | Pr > \|t\| |
|---|---|
| -4.97 | <.0001 |

*Figure 2.* T-test analysis for the self-reported SE levels per question (incorrect – 0 and correct – 1) for the exam.




| r Value | Pr > \|t\| |
|---|---|
| -0.12769 | <.0001 |

*Figure 3.* Regression of the logarithm of the median phasic EDA (microSiemens) per incorrect exam questions to the SE Likert scale responses.




| r Value | Pr > \|t\| |
|---|---|
| -0.07109 | 0.0552 |

*Figure 4.* Regression of the logarithm of the median phasic EDA (microSiemens) per correct exam questions to the SE Likert Scale responses.