Exploring Knowledge and Skill-Based Performance of STEM Students to Digital, Written and Video-Based Tutorials for Cell Culture Techniques

Naqash Gerard
Department of Engineering Education, Herbert Wertheim College of Engineering
University of Florida
Gainesville, FL, USA
naqash.gerard@ufl.edu
ORCID: 0000-0003-3118-0692

Idalis Villanueva Alarcón
Department of Engineering Education, Herbert Wertheim College of Engineering
University of Florida
Gainesville, FL, USA
i.villanueva@ufl.edu
ORCID: 0000-0002-8767-2576

Jing Pan
Department of Mechanical and Aerospace Engineering
Herbert Wertheim College of Engineering
University of Florida
Gainesville, FL, USA
jingpan@ufl.edu
ORCID: 0000-0001-6081-3204

***Abstract*—This research full paper describes a measuring performance using assessments allows instructors to evaluate if students have developed the requisite skills and knowledge needed to advance in their field. Cognitive factors alone do not contribute to performance. Researchers are increasingly utilizing physiological sensing techniques and biological markers to gain a deeper sense of the science of learning and performance. This exploratory study seeks to understand how cognitive and training performance looks like for a group of seven undergraduate biomedical and chemical engineering students participating in a summer research experience. Students were tasked to learn about cell culture lab techniques using video-based tutorials (experimental condition) or digitally written tutorials (control condition). A quiz was designed, face and content validated by the authors and provided to the students before and after the cell culture training. Preliminary observations suggest improved laboratory student performance, using video-based instruction. Future work will seek to better understand how lab training interventions based on video and other forms of instructional aids can improve laboratory student performance.**

Keywords: Engineering education, Physiological assessment, Academic emotional arousal, Video-based tutorials, Cell culture training.

I. Introduction

In STEM education, both knowledge and skill-based learning are essential for students to not just put together important concepts but to apply them in a real-world context. Combining knowledge and skills into learning provides a leading educational paradigm for innovation within learning environments [1]. However, in the context of laboratory training settings, little is known about how knowledge and skill-based learning occur and what instructional tools best support lab-based student learning and performance. Furthermore, measuring performance in a laboratory setting is not straightforward as typical assessment measures such as exams or surveys are seldom used in these settings.

For undergraduate students learning about lab skills in a course, Amida and colleagues [2] found that a multifaceted combination of instructional aids (e.g., manuals, videos, guides) and reflective self-reports yielded the most effective learning outcomes for students. On the other hand, other scholars argue that instructional aids are not enough to assess learning and performance in labs and that the actual immersive lab experiences (e.g., undergraduate research experience (URE)) is what determines performance [3].

Expanding upon UREs, a recent systematic literature review showed a gap in research for effective implementation of URE in other STEM fields outside of biology-related fields [4]. Generally organized over a summer, UREs allow students to spend the bulk of their time in a research lab to help devise, refine, and conduct a research-driven lab project. These types of experiences are believed to enhance their problem-solving and critical-thinking skills in laboratory settings [4]. However, the combination of instruction aids with laboratory experiences to assess learning and performance (hybrid models of UREs) are less understood.

Scholars are increasingly calling for a more thorough investigation of the impacts that planning and implementing UREs, including trainings using pre/post-survey designs can have in analyzing " real (instead of perceived) student achievement results..." [4, p. 14]. Thus, there is a dire need to explore the impacts of instructional aids within laboratory training to understand how undergraduates improve their research knowledge and skill-based performance. This exploratory study serves as a first attempt to understand what happens to undergraduate STEM undergraduate students, some in engineering and others in non-engineering disciplines, when exposed to a summer research experience involving biological laboratory techniques, in near-real-time.

II. Background

Traditional measures such as self-reports (e.g., surveys, interviews) are frequently utilized by educational researchers to observe changes in knowledge and skill-based performance in different training settings due to an intervention; however, these instruments are not free from flaws [5]. Pre- and post-test self-reports for lab-based experiences provide valuable insights of a student's exposure, but they do not help identify the real impacts learning of a URE. Additional limitations to self-reports include its overreliance on a person's retrospective or subjective perspectives and memories, higher susceptibility to bias, and limited consideration of authentic contexts, which limits researchers' understanding of nuanced barriers students may face as they are gaining lab skills [6, 7, 8].

In recent years, educational psychologists and STEM education researchers have begun to combine self-reports with physiological sensor technologies to gain both the perceptual and real experiences of participants as they engage on a learning and training task [8, 9]. These physiologically-embedded computing systems have been used in classrooms to monitor students' engagement levels [8, 9]. By tracking physiological signals like heart rate, skin conductance, and brain activity, these systems can infer student's cognitive and emotional state in near-real-time, offering an objective measure of performance. When combined with subjective measures offer,physiologically-embedded computing systems offers a holistic approach to study performance [8, 9]. For example, if a student appears disengaged or stressed, a physiologically-embedded system could adapt the learning content or provide targeted support to help students re-engage with experiential learning [8]. In one study, researchers used electrodermal activity (EDA), a measure of physiological stress while undergraduate computer science and engineering students were programming and found that students who experienced higher physiological stress levels tended to make more errors in their code [9].

This research employed a multimodal method [10, 11] combining eye-tracking, electrodermal activity, and salivary biomarker analysis before and after undergraduate lab skills training session. Using a pre/post quiz, to understand knowledge and skills, seven STEM undergraduate students in their junior year were asked to study provided material in two forms: a digital written lab book (control) and video-based tutorials embedded into the digital written lab book (experimental condition). These two instructional aids were studied to thoroughly assess the factors that impacted lab skill acquisition as they prepared to put into practice cell culture cryopreservation techniques. This exploratory but full research study took place during the Summer of 2023.

Additional goals of this exploratory study served: to explore the utility of eye tracking to assess student attention [12] and the use of physiological electrodermal activity (EDA) to assess cognitive load and emotional arousal [13] knowledge and skill-based performance that happens when undergraduate STEM students participate in a summer URE.

A. Use of Multimodal Physiological Tools (EDA and Eye Tracker) in Task Performance Studies

Eye tracking (ET) and electrodermal activity (EDA) are frequently used physiological methods to assess cognitive processes and emotional arousal levels during learning tasks [14]. Eye tracking uses cameras or sensors to monitor and record eye movements and fixations to a region of interest, offering valuable insights into how visual attention is allocated [15]. Research studies have utilized eye-tracking technology to evaluate the focus and concentration of engineering students during exams [17, 18]. Researchers discovered that individuals who achieved high-performance levels dedicated a greater amount of time to concentrate on specified rote-recall information, such as formulas, while allocating less attention to peripheral content, such as headers. Suggesting selective attention to assigned tasks [19].

EDA quantifies fluctuations in sweat gland function using electrodermal activity sensors, which indicate the stimulation of the sympathetic nervous system and is an indicator of physiological emotional arousals [16]. Researchers have recently integrated eye tracking with EDA to assess both cognitive load and emotional arousal levels. A study monitored the eye movements and EDA of 105 undergraduate computer science students while they were taking coding exams at private university in the Midwest [20]. The study observed that the students experienced an elevated blinking rate, pupil dilation, and increased skin conductance when they encountered complex syntax errors suggesting excessive cognitive load and high levels of mental effort when taking the coding exam [19, 20]. Taken together, eye tracking enables the analysis of visual attention and scanning patterns while EDA measures physiological arousal due to an emotive response [28].

For this exploratory study, eye tracking and EDA were used in a multi-modal approach through measures of emotional arousal and cognitive load using an online quiz assessment before (pre) and after (post) a lab training experience for a URE summer program focused on cell culture [10, 27, 32]. The use of knowledge assessment (e.g., quiz) and physiological tools in STEM education or UREs is relatively nascent. To our knowledge, no other study has combined these tools to understand laboratory-specific skill acquisition in STEM education.

III. Research Design and Methods

This full research study was conducted as an exploratory quasi-experimental study [21] where a group of seven STEM education undergraduate students were placed in either a control or experimental condition to infer the causality between the independent variable (type of instructional aid) and dependent variable (lab-skill performance). The control condition consisted of an instructor-designed digital, written tutorial on laboratory cell culture techniques while the experimental condition consisted of the same tutorial with video-based contents and instructions; the latter group was also provided the digital, written tutorial for additional support. The research hypothesis was:

A. Research Hypothesis and Rationale

Hypothesis 1. *There will be an increase in pre- and post lab-based performance for URE students who utilized video-based learning materials compared to those who utilized digital, written learning materials only.*

We hypothesized that video-based supplementary material will help students to cognitively establish connections to the lab training that will lead to an improved quiz performance compared to text-only materials. According to the literature, the use of educational videos in the context of teaching and trainings can enhance engagement, support mental model development, strengthen mastery, recall and increase of complex technical concepts essential in engineering education [22, 33]. This prediction was also found in prior research in which video-based learning improved skill acquisition over text-based learning alone [23] and provided additional sensory modalities that support students' constructions of mental schemas via dynamic visuals and demonstrations [24].

B. Data Analysis and Interpretation

This study had 2 types of sample sizes. The first is participant sample size which included control (n=3) and experimental (n=4) groups. The second sample uses ET and EDA physiological data of participants specifically. Eye tracking data was collected using a GazePoint GP3 eye tracker at 60 Hz, resulting in n (avg) = 144,000 data points for the control group and n (avg) = 90,000 data points for the experimental group. This high-resolution data allowed for detailed analysis of pupil dilation and blink rate to assess visual attention patterns and cognitive load. Concurrently, EDA was recorded continuously using an Empatica E4 wristband. EDA data was preprocessed to extract baseline tonic (skin conductance level, SCL) and phasic (skin conductance responses, SCRs) components, providing insights into overall arousal and event-related physiological responses. To integrate these diverse data streams, Pearson correlation analyses were conducted between EDA metrics and eye-tracking parameters for the control and experimental groups. The participant sample size served to describe average performance scores in the pre- and post-quiz and time of quiz taking.

IV. PARTICIPANTS

For the participants, a group of seven undergraduate STEM education students, comprising three male-identifying and four female-identifying participants in a summer iGEM synthetic biology undergraduate summer research experience led by the third author. The junior and senior students were chosen through convenience sampling [25]. The third author pre-assessed the previous knowledge about laboratory cell culture techniques. All student participants reported to have basic knowledge on biology/microbiology but had not been trained in mammalian cell culture. The third author developed a NIH/3T3 Fibroblast Cell Culture Basics training for the students and in consultation with the second author assigned students to one of two groups - an experimental group (n=4) receiving video-based learning materials and a control group (n=3) receiving equivalent text-based materials.

Under Institutional Review Board approvals (IRB 202200802), the students were provided with a demographic questionnaire asking about their cell culture and bench laboratory experience in addition to their self-identified gender, race, ethnicity, age range, year of undergraduate study, age range, and GPA. This survey was provided one week before the cell culture training. The research team quickly assessed the level of experience self-reported by the participants to pre-assign them to a control group or experimental group. A control group was assigned to the students who had a lot of cell culture/bench lab experience, and the experimental group was assigned to students who had little-to-no experience. The rationale was that by allowing the inexperienced group to have the video-based tutorials, their skill acquisition may be like the control group after the training.

The third author had custom-created two versions of the NIH/3T3 Fibroblast Cell Culture tutorial a few weeks prior to the study. One version of the tutorial, aimed for the control group, contained digitally written laboratory protocols and PDF-generated resources. The second version of the tutorial, aimed for the experimental group, contained the same resources as the control group with the added feature of video-based content that included a suite of cell culture video demonstrations from the Journal of Visualized Experiments and YouTube. The students were required to take all necessary Environmental Health & Safety training for handling biological reagents, waste disposal and autoclaving procedures before the study. Upon identifying the control and experimental group, the third author allowed access to one of the two tutorials corresponding to their group assignment. The sharing of the content happened simultaneously and occurred two days before the study to ensure all participants received the same time frame to prepare for the NIH/3T3 Fibroblast Cell Culture Training.

Quiz materials were created by the second author using the cell culture tutorial as a reference and the 12 multiple-choice question quiz was shared with the third author for verification and accuracy of the content and to ensure that the content selected was the same across the two experimental conditions. The quiz consisted of 12 questions that covered the cell culture tutorial content. The first and second author custom-created a Canvas learning management system module to house the quiz as well as the protocol needed to collect the eye tracking and EDA data. This quiz included step-by-step instructions on how to calibrate the physiological tools to establish a baseline signal before the participants took the quiz.

On the day of the experiment, participants first completed a pre-quiz in the testing station found in the lab of the second author. The average duration of the pre-quiz was 40 minutes for the control group and 25 minutes for the experimental group. The student groups were then escorted to the third author's laboratory to put into practice what they had learned in the tutorial and to participate on a hands-on training session with a graduate student of the third author's lab. The cell culture technique training sessions averaged 2 hours for the control group and 1 hour 30 minutes for the experimental group. During this timeframe, the graduate student trainer annotated in a lab skill worksheet created by the third author, the number of cell culture lab skills that were demonstrably accurate by the participants. After completing the hands-on training, the groups were escorted back to the testing session of the second author's lab so that they can complete a post-quiz that had the same content as the pre-quiz. The average completion time for the post quiz was 24 minutes for the control group and 27 minutes for the experimental group.

V. Testing Station Setup and Apparatus Details

A. Testing Station Setup

During the pre- and post-quiz testing sessions, participants were asked to sit down in a desk station containing a laptop and a separately connected monitor, keyboard, mouse, and a custom-made Canvas quiz page. Underneath the monitor, a Gazepoint GP3 eye tracker was utilized to capture participants' gaze data throughout the quiz period. The GP3 angle was pre-set at around 45-degrees from the bottom of the monitor and pointing to the participants' eyes, although the degree was adjusted according to participant height by a member of the lab team of the second author. Proper angling of the eye tracker was verified by running the calibration protocol embedded in the eye-tracker software. The first author and a graduate student lab assistant helped to calibrate the eye tracker and EDA sensors as the participants sat down. The second author assisted with participant setup and salivary sampling. The third author and a lab assistant supporting laboratory training of participant pre- and post-quiz.

B. Apparatus Details

The gazepoint GP3 is a non-invasive, lightweight eye-tracker. GP3 is designed to capture eye gaze data in a way that enables unrestricted movement of the head during tracking. The GP3 system operates by projecting near-infrared light on the eyes at a wavelength of 850nm and recording corneal reflections with a sampling rate of 60Hz using sensors integrated into the glasses frame. Gaze position accuracy is 0.5-1 degrees over an 80-degree horizontal and 40-degree vertical range. The eye tracker is calibrated for each participant prior to use to map their individual gaze coordinates and was time-synced to the laptop internal clock. The GP3 records eye-tracking data with high temporal resolution, typically at 60Hz or higher, and each data point is timestamped, enabling detailed analysis of eye movements and gaze positions over time.

EDA was assessed using a wireless, Bluetooth-enabled Empatica E4 non-invasive, wristbands equipped with both EDA and PPG sensors to enable the measurement of the sympathetic nervous system activity as well as heart rate. In addition, E4 houses additional sensors that capture accelerometry data for movement, internal thermopiles (for skin temperature detection), and has event-marking capabilities with a 5ppm time referencing accuracy. E4 has an internal memory capable of storing up to 60 hours of data with a 5s synchronization resolution. For this study, the research team solely used the EDA data and the event-marking capabilities. For the EDA, variations in the electrodermal conductivity of the skin, measured by electrical conductance changes in microSiemens was collected at a sampling rate of 4Hz to measure tonic (baseline) and phasic (reactive/responsive) levels. The EDA signal provides an indication of sympathetic nervous system arousal and physiologically induced emotional arousal levels.

Both EDA and ET were time synchronized to the quiz questions using the analytics from the Canvas quiz and the laptop internal clock. This allowed for a more nuanced data capturing that will be described in a future paper. Note that because EDA and ET data streams range between 4 Hz and 60 Hz, respectively, the physiological sample size for the control group was almost 200,000 and for the experimental group was 430,000. As such, while the participant numbers may be small, the real-time physiological data is substantiative to provide a robust understanding of the experiences that these students had during the study.

VI. Procedure

The experiment was conducted following Institutional Review Board (IRB)-approved procedures (IRB 202200802) and spanned two days. On the first day, all participants in the control group were placed in a lab station fitted with a computer monitor, keyboard, mouse, one ET positioned at the bottom of the monitor at a 45-degree angle pointing towards the participant's eyes, and one EDA wristband that was adjusted snugly to participants' non-dominant hand by a member of the research team. Before the start of the quiz, the eye tracker was calibrated using Gazepoint GP3 software for each participant. Eye movement data and EDA baselines were captured according to manufacturer recommendations as well as previous lab methods from the research team [26]. Both GP3 and E4 were time synched and consistently recorded throughout the entire duration of the quiz.

The raw gaze data was analyzed utilizing Gazepoint software [29]. The Gazepoint standard filter was employed to detect and analyze blinks and fixations. The eye fixation locations were determined by identifying anatomical regions of interest that were relevant to performance on the quiz [30]. Analyzed skin conductance data was used to extract phasic (reactive) EDA components from its tonic components (baseline), which served as indicators of sympathetic arousal in response to specific events [31]. Saliva samples were collected at five distinct time intervals: 1) 10 minutes before the quiz as a reference point, 2) after viewing the introductory video, 3) during the midpoint of the quiz, 4) immediately after the quiz, and 5) 10 minutes after the quiz. Note that for this study, salivary data or laboratory skill tracking will not be presented as data analysis is ongoing.

VII. Descriptive and Statistical Analysis

Descriptive summaries for the participant performance data, means and standard deviations, were calculated to describe differences in quiz completion time between the control and experimental groups. For the physiological electrodermal and eye-tracker data, in alignment with the assigned control and experimental group designations was used where $p < 0.05$

were statistically significant and ($0.05 < p \leq 0.10$) are marginally significant. To explore potential physiological correlates of performance, a Pearson correlation analysis was performed on the eye-tracking per group to identify any statistically significant relationships between EDA measures and ET metrics.

VIII. Results

A. Observation of Group Performance Results

The average time (minutes) and scores (%) that the participants in both groups for the pre-quiz and post-quiz are summarized in Table 1. As seen in Table 1, the control group participants exhibited a decrease in the time for the post-quiz compared to the pre-quiz. In the control group, the time to complete the post-quiz decreased by 10-20 minutes compared to the pre-quiz for all participants. The average time to complete the pre- and post-quiz decreased from 38.33 minutes (SD = 5.77) in the pre-quiz to 24.33 minutes (SD = 1.15) in the post-quiz. Similar reductions were observed in the experimental group, with all participants completing the post-quiz 8-14 minutes faster than the pre-quiz. The average time for the experimental group decreased from 41.00 minutes (SD = 2.94) in the pre-quiz to 27.00 minutes (SD = 3.65) in the post-quiz.

Regarding quiz scores, the control group showed an improvement from an average of 52.67% (SD = 17.62) in the pre-quiz to 66.67% (SD = 8.50) in the post-quiz. The experimental group, however, showed a slight decrease in average score from 77% (SD = 7.66) in the pre-quiz to 73% (SD = 17.30) in the post-quiz.

It's important to note that the experimental group started with a higher average pre-quiz score (77%) compared to the control group (52.67%). Despite the slight decrease in the experimental group's post-quiz score, they still maintained a higher average (73%) compared to the control group's post-quiz average (66.67%). The data suggests the control group showed improvement in both time and score, while the experimental group improved in time but showed a slight decrease in average score. However, the experimental group maintained higher scores overall.

TABLE I
Pre-Quiz and Post-Quiz Performance Comparison Between Control and Experimental Groups

		Pre-Quiz		Post-Quiz	
	Participant	Score (%)	Time (min)	Score (%)	Time (min)
Control	C-001	58	35	58	25
	C-002	67	35	75	23
	C-003	33	45	67	25
Average	-	52.66	38.33	66.66	24.33
SD	-	17.61	5.77	8.50	1.15
Experimental	E-001	67	37	75	23
	E-002	83	42	92	31
	E-003	75	41	50	25
	E-004	83	44	75	29
Average	-	77	41	73	27
SD	-	7.65	2.94	17.30	3.65

B. Physiological Data Results

As shown in Table 2, the eye-tracker blink rate and pupil dilation data for both the control and experimental groups are summarized during the pre-quiz and post-quiz phases. Blink rate, an indicator of cognitive fatigue, showed an increase in the control group from the pre-quiz (M=16.1, SD=7.12) to the post-quiz (M=20.47, SD=9.03). In contrast, the experimental group exhibited a decrease in blink rate from the pre-quiz (M=23.14, SD=7.56) to the post-quiz (M=11.91, SD=9.03). However, the differences between the control group (t = -3.18, p = 0.086) is marginally significant and experimental group (t = 2.18, p = 0.117) is not statistically significant for either the pre-quiz (t = -1.248, p = 0.134) or the post-quiz (t = 1.301, p = 0.125). Pupil dilation, an indicator of cognitive load, showed a slight decrease in the control group from the pre-quiz (M=16.13, SD=1.50) to the post-quiz (M=15.79, SD=2.55). The experimental group, however, demonstrated a small increase in pupil dilation from the pre-quiz (M=17.74, SD=2.77) to the post-quiz (M=18.39, SD=1.20). The differences between the control group (t = 0.53, p = 0.64) and experimental group (t = -0.41, p = 0.70) were not statistically significant for the pre-quiz (t = -0.897, p = 0.205) and marginally significant for the post-quiz (t = -1.822, p = 0.064). The eye-tracker data suggests that the experimental group experienced a decrease in cognitive fatigue (as indicated by the reduced blink rate) and a slight increase in cognitive load (as indicated by the increased pupil dilation) from the pre-quiz to the post-quiz phase. In contrast, the control group showed an increase in cognitive fatigue and a slight decrease in cognitive load. However, the lack of statistical significance in the differences between the groups suggests that these changes could be due to chance.

TABLE II
Eye-tracker data summary for control and experimental groups

	Blink Rate, blinks/minute						Pupil Dilation, mm/second					
	Pre-Quiz		Post-Quiz		t-test	p-value	Pre-Quiz		Post-Quiz		t-test	p-value
	Mean	SD	Mean	SD			Mean	SD	Mean	SD		
Control	16.10	7.12	20.47	9.03	-3.18	**0.086**	16.13	1.50	15.79	2.55	0.53	0.64
Experimental	23.14	7.56	11.91	9.03	2.18	0.117	17.74	2.77	18.39	1.20	-0.41	0.70
t-test	-1.248		1.301		-	-	-0.897		-1.822		-	-
p-value	0.134		0.125		-	-	0.205		**0.064**		-	-

The EDA analysis, as shown in Table 3, provides a detailed insight into the physiological emotional arousal of participants during the pre-quiz and post-quiz for control and experimental groups. In the pre-quiz, the experimental group exhibited higher mean phasic (M=1.17, SD=2.29) and tonic (M=1.36, SD=2.21) compared to the control group (phasic: M=0.358, SD=0.61; tonic: M=0.39, SD=0.36). However, the differences between control group (t = -0.063, p = 0.995) and experimental group (t = -0.102, p = 0.926) were not statistically significant for either phasic (t = -0.58, p = 0.29) or tonic EDA (t = -0.73, p = 0.24). During the post-quiz, the experimental group continued to show higher mean phasic EDA (M=1.12, SD=1.29) compared to tonic EDA (M=0.42, SD=0.36) compared to the control group (phasic

EDA: M=0.68, SD=1.17; tonic EDA: M=0.09, SD=0.08). The differences between the control group (t = 0.812, p = 0.502) and experimental group (t = 0.881, p = 0.443) were not statistically significant for phasic EDA (t = -0.46, p = 0.33) and marginally significant for tonic EDA (t = -1.47, p = 0.10). The higher tonic EDA values in the experimental group suggest that participants in this group experienced greater baseline physiological emotional arousal across the experiment compared to the control group suggesting potential autonomic changes in arousal, that may have resulted from changes from spontaneous fluctuations in EDA to a sustained stimulus (e.g., quiz questions) over a period. However, the lack of statistical significance indicates that the differences in EDA between the groups could be due to chance.

TABLE III
EDA DATA SUMMARY FOR CONTROL AND EXPERIMENTAL GROUPS

	EDA, microSiemens											
	Pre-Quiz						Post-Quiz					
	Phasic		Tonic		t-test	p-value	Phasic		Tonic		t-test	p-value
	Mean	SD	Mean	SD			Mean	SD	Mean	SD		
Control	0.36	0.61	0.39	0.36	-0.06	0.96	0.68	1.17	0.09	0.08	0.81	0.50
Experimental	1.17	2.29	1.36	2.21	-0.10	0.93	1.12	1.29	0.42	0.36	0.88	0.44
t-test	-0.58		-0.73		-	-	-0.46		-1.47		-	-
p-value	0.29		0.24		-	-	0.33		**0.10**		-	-

C. *Pearson Correlation Results*

A bivariate Pearson correlation analysis was performed to investigate the relationships between physiological measures, including pupil dilation, EDA tonic values, EDA phasic values, and blink rate as students were taking the pre- and post-quiz. The results revealed varying degrees of correlation between these variables, although none of the correlations reached statistical significance.

TABLE IV
PEARSON CORRELATION FOR THE CONTROL GROUP

	Pupil Dilation & Tonic EDA		Pupil Dilation & Phasic EDA		Bink Rate & Tonic EDA		Blink Rate & Phasic EDA	
	Pre-Quiz	Post-Quiz	Pre-Quiz	Post-Quiz	Pre-Quiz	Post-Quiz	Pre-Quiz	Post-Quiz
r	-0.85	0.47	0.38	0.22	-0.26	0.95	-0.36	-0.51
p-value	0.34	0.68	0.74	0.85	0.83	**0.19**	0.76	0.65
R^2	0.72	0.22	0.14	0.04	0.06	0.90	0.12	0.26
Variance	72%	22%	14%	4%	6%	90%	12%	26%

Upon analyzing the data in Table 4, we observed a marginal positive correlation between the post-quiz blink rate and post-quiz EDA tonic for the control group, with a Pearson correlation coefficient of r = 0.95, p = 0.19, although not significant but close to marginal significance for the control group.

TABLE V
PEARSON CORRELATION FOR THE EXPERIMENTAL GROUP

	Pupil Dilation & Tonic EDA		Pupil Dilation & Phasic EDA		Bink Rate & Tonic EDA		Blink Rate & Phasic EDA	
	Pre-Quiz	Post-Quiz	Pre-Quiz	Post-Quiz	Pre-Quiz	Post-Quiz	Pre-Quiz	Post-Quiz
r	0.71	-0.07	0.06	0.71	0.73	-0.26	-0.60	-0.25
p-value	0.28	0.92	0.93	0.28	0.26	0.73	0.31	0.74
R^2	0.50	0.005	0.003	0.51	0.54	0.06	0.36	0.06
Variance	50%	0%	0%	51%	54%	6%	36%	6%

After examining the data in Table 5, we found no statistically significant differences in pupil dilation, blink rate, EDA phasic, and EDA tonic for the experimental group.

IX. DISCUSSION

The study aimed to compare the effectiveness of video-based and text-based instructional materials in relation to students' learning outcomes and physiological responses. The results indicate that both groups experienced an observable though not statistically significant decrease in the time taken to complete the post-quiz compared to the pre-quiz. Nevertheless, the experimental group, which employed video-based learning material, appeared to have better results on the post-quiz despite a slight decrease in the mean score compared to the pre-quiz. The control group, which utilized text-based instructional material, appeared to have a greater improvement in average scores from the pre-quiz to the post-quiz.

After conducting an in-depth investigation, we discovered one noteworthy participant in the control group who achieved a very low score on the pre-quiz and took a longer time to complete it compared to the other participants. This participant acknowledged using internet videos (i.e., YouTube) to prepare for the quiz despite the researchers' instructions to refrain from using online tools; this participant also acknowledged not having prior cell culture technique but having general lab experience. Nevertheless, this same participant exhibited significantly better performance compared to the other two participants in the control group during the post-quiz. This participant achieved a 50% increase in post-quiz score compared to their pre-quiz score and demonstrated the greatest improvement in learning from the cell culture training among all the participants in the control group. This student also became one of the highest performers of the group towards the end of the URE, as reported by the third author (data not shown).

Based on the eye-tracker results, it can be inferred that the experimental group showed a decrease in cognitive fatigue (as indicated by a decrease in blink rate) and a slight increase in cognitive load (as indicated by an increase in pupil dilation) from the pre-quiz to the post-quiz phase. In contrast, the control group demonstrated an increase in cognitive fatigue and a slight decline in cognitive load. This may suggest that the experimental group, having more lab experience, afforded the space to focus more on the task and concentrate on the application of the learned content compared to the control group.

Interestingly, the experimental group showed higher emotional arousal across the pre- and post-quizzes for both phasic and tonic EDA although the post-quiz event appeared to have the highest tonic EDA levels with a marginal significance compared to the control group. This may suggest either a long-standing cognitive stress prior to the summer experience perhaps due to a desire to perform accurately to the experiment and training task or due to frustration about the experimental setup and time duration. Further experiments would need additional qualitative data (post-interview with participants) to nuance these findings further.

The bivariate Pearson correlation analysis examined the relationships between students' physiological measures, specifically pupil dilation, tonic EDA, phasic EDA, and blink rate. Although none of the correlations achieved statistical significance two noteworthy findings showed a strong positive correlation between the post-quiz pupil dilation and tonic EDA for control and experimental groups and a strong positive correlation found between post-quiz blink rate and post-quiz EDA tonic (for control group). Both suggest a potential relationship between cognitive load and emotional reactive arousal responses to a task.

To summarize, the study demonstrated that the experimental group, who used the video-based instructions, achieved a slightly better score on the post-quiz. On the other hand, the control group, which used text-based instructions, showed a greater improvement from the pre-quiz to the post-quiz. However, these differences were not statistically significant. The control group performed better in the post-quiz as compared to the pre-quiz, although the difference was not statistically significant. Meanwhile, the experimental group performed almost the same in both the pre-quiz and post-quiz. The eye-tracker data showed a slightly significant rise in blink rate for the control group, suggesting that they experienced more cognitive fatigue compared to the experimental group. The reason for this could be that the control group studied only with text-based instruction, whereas the experimental group received video-based instruction. This difference in instructional methods may have resulted in reduced fatigue for the experimental group, as video-based instruction is believed to facilitate faster learning. The observed effect can be attributed to the utilization of video-based instructions, that help in reducing fatigue and facilitating faster learning within the experimental group.

X. Conclusion and Next Steps

The study suggests that video-based instructions might have advantages compared to text-based material, such as reducing learner fatigue and potentially improving the learning process. With that said, we cannot discard the influence that prior experience and perhaps disciplinary areas of study may have played.

Future studies will aim to increase sample sizes to identify the potential statistical significance of the findings. Also, the research team will begin to randomize the treatments as well as expand modalities of instruction as it was suggested that a combination of text and video (per the control participant exception observed) may be conducive to increase in performance. More work will continue to explore the influences that instructional aids can have in research lab training for undergraduate students as this is an underexplored area of study.

XI. Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant No. CBET # 2339756. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. We also are grateful to Dr. Jing Pan's graduate student Yongchen Tai for providing invaluable support with participant recruitment, data collection, and pre- and post-quiz and lab training for the participants. We also thank Suneet Jain for his vital contribution to this study and assistance with data analysis, without which this work would not have been completed.

XII. Author Contribution

Author 1: Naqash Gerard
- Designed data collection protocols and collected the data
- Contributed data or analysis
- Performed the analysis
- Write the paper

Author 2: Idalis Villanueva Alarcón
- Conceived and designed the study
- Collected the data
- Contributed data or analysis
- Write and review the paper

Author 3: Jing Pan
- Conceived and designed the study
- Collected the data
- Write and review the paper

References

[1] Letsatsi, M. T., Agarwal, A., & Seretse, O. M. (2019). The Battle towards Skill Based Competency Integration to Knowledge Based Competency in the Sustainable Development of Growing Economy. Int. J. Recent Technol. Eng, 8(1).

[2] Amida, A., Chang, I., & Yearwood, D. (2020). Designing a practical lab-based assessment: A case study. Journal of Engineering, Design and Technology, 18(3), 567-581.

[3] Beheshtian, C., Garcia, V. E., Zhu-Hui Ng, T., Alkhatib, S., Quang, E., Cho, K. J., ... & Kadandale, P. (2023). Does exposure to research experiences have different learning outcomes than prior exposure to lab techniques in non-research settings?. Biochemistry and Molecular Biology Education, 51(2), 180-188.

[4] Ahmad, Z., & Al-Thani, N. J. (2022). Undergraduate Research Experience Models: A systematic review of the literature from 2011 to 2021. International Journal of Educational Research, 114, 101996.

[5] Jordan, K. L., & Sorby, S. A. (2014, June). Intervention to improve self-efficacy and sense of belonging of first-year underrepresented engineering students. In 2014 ASEE Annual Conference & Exposition, 24-803.

[6] Ma, C. W., Cheng, P. S., Chan, Y. S., & Tipoe, G. L. (2023). Virtual reality: a technology to promote active learning of physiology for students across multiple disciplines. Advances in Physiology Education, 47(3), 594-603.

[7] Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. Proceedings of the national academy of sciences, 111(23), 8410-8415.

[8] Fairclough, S. H. (2009). Fundamentals of physiological computing. Interacting with computers, 21(1-2), 133-145.

[9] Hernandez-Cuevas, B. Y. (2022). Physiological Computing Education: Towards Teaching Using Physiological Signals (Doctoral dissertation, University of Alabama Libraries).

[10] Villanueva Alarcón, I., Anwar, S., & Atiq, Z. (2023). How multi-modal approaches support engineering and computing education research. Australasian Journal of Engineering Education, 1-16.

[11] Kammermeier, P., Buss, M., & Schmidt, G. (2001). A systems theoretical model for human perception in multimodal presence systems. IEEE/ASME Transactions on mechatronics, 6(3), 234-244.

[12] Gerard, N., Yousuf, T., Johar, A. H., Asgher, U., Malik, I., Ul Hasan, A., & Shafait, F. (2021). Detection of Subject Attention in an Active Environment Through Facial Expressions Using Deep Learning Techniques and Computer Vision. In Advances in Neuroergonomics and Cognitive Engineering: Proceedings of the AHFE 2020 Virtual Conferences on Neuroergonomics and Cognitive Engineering, and Industrial Cognitive Ergonomics and Engineering Psychology, July 16-20, 2020, USA (pp. 326-332). Springer International Publishing.

[13] Conway, D., Dick, I., Li, Z., Wang, Y., & Chen, F. (2013). The effect of stress on cognitive load measurement. In Human-Computer Interaction–INTERACT 2013: 14th IFIP TC 13 International Conference, Cape Town, South Africa, September 2-6, 2013, Proceedings, Part IV 14 (pp. 659-666). Springer Berlin Heidelberg.

[14] Paul, E. S., Harding, E. J., & Mendl, M. (2005). Measuring emotional processes in animals: the utility of a cognitive approach. Neuroscience & Biobehavioral Reviews, 29(3), 469-491.

[15] Frutos-Pascual, M., & Garcia-Zapirain, B. (2015). Assessing visual attention using eye tracking sensors in intelligent cognitive therapies based on serious games. Sensors, 15(5), 11092-11117.

[16] Boucsein, W. (2012). Electrodermal activity. Springer Science & Business Media.

[17] Rosch, J. L., & Vogel-Walcutt, J. J. (2013). A review of eye-tracking applications as tools for training. Cognition, technology & work, 15, 313-327.

[18] Jamil, N., Belkacem, A. N., & Lakas, A. (2023). On enhancing students' cognitive abilities in online learning using brain activity and eye movements. Education and Information Technologies, 28(4), 4363-4397.

[19] Anderson, O. R., Love, B. C., & Tsai, M. J. (2014). Neuroscience perspectives for science and mathematics learning in technology-enhanced learning environments. International Journal of Science and Mathematics Education, 12, 467-474.

[20] Hutt, S., Mills, C., White, S., Donnelly, P. J., & D'Mello, S. K. (2016). The Eyes Have It: Gaze-Based Detection of Mind Wandering during Learning with an Intelligent Tutoring System. International Educational Data Mining Society.

[21] May, D. R., & Luth, M. T. (2013). The effectiveness of ethics education: A quasi-experimental field study. Science and engineering ethics, 19, 545-568.

[22] Mazza, J. T. (2015). The relationship between video-based instruction and student achievement in an online math course (Doctoral dissertation, Capella University).

[23] Mayer, R. E. (2002). Multimedia learning. In Psychology of learning and motivation (Vol. 41, pp. 85-139). Academic Press.

[24] Merkt, M., Weigand, S., Heier, A., & Schwan, S. (2011). Learning with videos vs. learning with print: The role of interactive features. Learning and Instruction, 21(6), 687-704.

[25] Stratton, S. J. (2021). Population research: convenience sampling strategies. Prehospital and disaster Medicine, 36(4), 373-374.

[26] Khan, T. H., Villanueva, I., Vicioso, P., & Husman, J. (2019, October). Exploring relationships between electrodermal activity, skin temperature, and performance during. In 2019 IEEE Frontiers in Education Conference (FIE) (pp. 1-5). IEEE.

[27] Hogervorst, M. A., Brouwer, A. M., & Van Erp, J. B. (2014). Combining and comparing EEG, peripheral physiology and eye-related measures for the assessment of mental workload. Frontiers in neuroscience, 8, 322.

[28] Chesnut, M., Harati, S., Paredes, P., Khan, Y., Foudeh, A., Kim, J., ... & Williams, L. M. (2021). Stress markers for mental states and biotypes of depression and anxiety: A scoping review and preliminary illustrative analysis. Chronic Stress, 5, 24705470211000338.

[29] Cuve, H. C., Stojanov, J., Roberts-Gaal, X., Catmur, C., & Bird, G. (2022). Validation of Gazepoint low-cost eye-tracking and psychophysiology bundle. Behavior Research Methods, 54(2), 1027-1049.

[30] Krupinski, E. A., Tillack, A. A., Richter, L., Henderson, J. T., Bhattacharyya, A. K., Scott, K. M., ... & Weinstein, R. S. (2006). Eye-movement study and human performance using telepathology virtual slides. Implications for medical education and differences with experience. Human pathology, 37(12), 1543-1556.

[31] Figner, B., & Murphy, R. O. (2011). Using skin conductance in judgment and decision making research. A handbook of process tracing methods for decision research, 163-184.

[32] Jayasekaran, S. R., Farooq, U., Anwar, S. (2023, June). Impact of extra credit for practice questions on programming students' participation and performance. In 2023 ASEE Annual Conference Exposition.

[33] G. Aslam and I. Villanueva Alarcón, "Exploring Contributions of U.S. Engineering Education Research Centers," 2023 IEEE Frontiers in Education Conference (FIE), College Station, TX, USA, 2023, pp. 1-6, doi: 10.1109/FIE58773.2023.10343418.