Journal of Visualized Experiments

Considerations for studying near-real-time authentic examination experiences: A cross-disciplinary and multi-modal experimental design

--Manuscript Draft--

Article Type:	Methods Article - JoVE Produced Video
Manuscript Number:	JoVE60037R1
Full Title:	Considerations for studying near-real-time authentic examination experiences: A cross-disciplinary and multi-modal experimental design
Keywords:	Real-Time, Examination, Performance, Cross-disciplinary, Multi-Modal, Experimental
Corresponding Author:	Idalis Villanueva, Ph.D. Utah State University Logan, UT UNITED STATES
Corresponding Author's Institution:	Utah State University
Corresponding Author E-Mail:	idalis.villanueva@usu.edu
Order of Authors:	Idalis Villanueva, Ph.D.
	Jenefer Husman
	Darcie Christensen
	Kate Youmans
	Md Tarique Khan
	Paul Vicioso
	Shawn Lampkins
	Matthew C Graham
Additional Information:	
Question	**Response**
Please indicate whether this article will be Standard Access or Open Access.	Standard Access (US$2,400)
Please indicate the **city, state/province, and country** where this article will be **filmed**. Please do not use abbreviations.	Logan, UT, USA

TITLE:

A Cross-Disciplinary and Multi-Modal Experimental Design for Studying Near-Real-Time Authentic Examination Experiences

AUTHORS AND AFFILIATIONS:

Idalis Villanueva[1], Jenefer Husman[2], Darcie Christensen[1], Kate Youmans[1], Md. Tarique Hasan Khan[1], Paul Vicioso[1], Shawn Lampkins[2], Matt Graham[2]

[1] Department of Engineering Education, College of Engineering, Utah State University

[2] Department of Educational Psychology, College of Education, University of Oregon

Corresponding Author:

Idalis Villanueva (idalis.villanueva@usu.edu)

Email Addresses of Co-Authors:

Jenefer Husman (jhusman@uoregon.edu)
Darcie Christensen (darc.c@aggiemail.usu.edu)
Kate Youmans (kate.youmans@aggiemail.usu.edu)
Md. Tarique Hasan Khan (tariquehasan@aggiemail.usu.edu)
Paul Vicioso (paulvicioso@gmail.com)
Shawn Lampkins (slampkin@uoregon.edu)
Matt Graham (mgraham@uoregon.edu)

KEYWORDS:

real-time, examination, performance, cross-disciplinary, multi-modal, experimental

SUMMARY:

An experimental design was developed to investigate the real-time influences of an examination experience to assess the emotional realities students experience in higher education settings and tasks. This design is the result of a cross-disciplinary (e.g., educational psychology, biology, physiology, engineering) and multi-modal (e.g., salivary markers, surveys, electrodermal sensor) approach.

ABSTRACT:

Over the past ten years, research into students' emotions in educational environments has increased. Although researchers have called for more studies that rely on objective measures of emotional experience, limitations on utilizing multi-modal data sources exist. Studies of emotion and emotional regulation in classrooms traditionally rely on survey instruments, experience-sampling, artifacts, interviews, or observational procedures. These methods, while valuable, are largely dependent on participant or observer subjectivity and is limited in its authentic measurement of students' real-time performance to a classroom activity or task. The latter, in particular, poses a stumbling block to many scholars seeking to objectively measure emotions and other related measures in the classroom, in real-time.

The purpose of this work is to present a protocol to experimentally study students' real-time

responses to exam experiences during an authentic assessment situation. For this, a team of educational psychologists, engineers, and engineering education researchers designed an experimental protocol that retained the limits required for accurate physiological sensor measurement, best-practices of salivary collection, and an authentic testing environment. In particular, existing studies that rely on physiological sensors are conducted in experimental environments that are disconnected from educational settings (e.g., Trier Stress Test), detached in time (e.g., before or after a task), or introduce analysis error (e.g., use of sensors in environments where students are likely to move). This limits our understanding of students' real-time responses to classroom activities and tasks. Furthermore, recent research has called for more considerations to be covered around issues of recruitment, replicability, validity, setups, data cleaning, preliminary analysis, and special circumstances (e.g., adding an additional variable in the experimental design) in academic emotions research that rely on multi-modal approaches.

INTRODUCTION:

Psychologists have long understood the importance of humans' emotions in elucidating their behaviors[1]. Within the study of education, Academic Achievement Emotions (AEE) has become the focus of emotion research[2]. Researchers that use AAE argue that the situational contexts students find themselves in are important to consider when examining students' emotions. Students may experience test-related, class-related, or learning-related emotions that involve multi-component processes, including affective, physiological, motivational, and cognitive components. AEE is expressed in two forms: valence (positive/negative) and activation (focused/unfocused energy). Positive activating emotions, such as enjoyment, may increase reflective processes like metacognition, whereas positive deactivating emotions such as pride may result in low levels of cognitive processing. Negative activating emotions such as anger and anxiety may spark engagement, whereas negative deactivating emotions such as hopelessness may dampen motivation[3-5]. Academic emotions contribute to how we learn, perceive, decide, respond, and problem-solve[2]. To regulate academic emotions, an individual must possess self-efficacy (SE) [6-8], which is their confidence in their ability to employ control over their motivation, behavior, and social environment[6]. Self-efficacy and academic emotions are interrelated, where lower self-efficacy is tied to negative deactivating emotions (e.g., anxiety, anger, boredom) and higher self-efficacy is tied to positive activating emotions (e.g., happiness, hope, excitement)[6-8]. SE is also believed to be strongly tied to performance[6-8].

Research that has examined classroom emotions have relied on self-reports, observations, interviews, and artifacts (e.g., exams, projects)[9,10]. Although these methods provide rich contextual information about students' classroom experiences, they have significant limitations. For example, interviews, observations, and self-reports rely on individuals' introspections[10]. Other methods have sought to examine academic emotions more proximally than prior researchers, such as those based on experience-sampling approaches where researchers ask students to report on their emotions during the school day[11]. Although this research allows us to more accurately report students' emotions, this work relies on self-report methods and does not allow for real-time reporting as students have to pause their work on the exam to address the experience survey.

Recently, researchers have begun to address concerns about self-report measures through the use of biological or physiological measures of emotion[9], that combined with other instruments or techniques such as surveys, observations, or interviews consists of a multi-modal form of data collection for educational and psychological research[12]. For example, biological techniques, including salivary biomarkers, are being used to understand the role biological processes have on cognition, emotion, learning, and performance[13-15]. For cognitive processes, androgens (e.g., testosterone) have been linked to different spatial recognition patterns in adults and children[16,17] whereas hypothalamic-pituitary-adrenocortical hormones (e.g., cortisol) and adrenergic hormones (e.g., salivary α-amylase or sAA) are linked to stress responsiveness amongst individuals[18-20].

Electrodermal activity (EDA) represents a physiological measure of the activation of the autonomic nervous system (ANS) and is linked to an increased activation of the system, cognitive load, or strong emotional responses[21-23]. In examination activities, EDA is affected by physical mobility[21,22], bodily and ambient temperatures[24-27], and verbalization of thoughts[28], as well as sensitivity and degree of connectivity of the analog-digital electrodes to the skin[29].

Although these can be limitations to using EDA, this technique can still provide valuable insight into what happens during near-real-time examinations and can serve as a promising tool to explore AEE and by extent, self-efficacy. As a result, an accurate picture of students' AEE can be obtained through a combination of survey methods, to determine the valence of an emotion, and physiological and biological data, to measure the activation of that emotion. This paper builds upon a prior publication on examination activities[30] and expands the scope of that work to include multi-modal approaches (using experience-sampling surveys, EDA sensors, and salivary biomarkers) in an examination scenario. It is important to mention that the protocol described below allows for multiple participant data to be collected at the same time within a single experimental setting.

As researchers consider experimental designs of this nature, disciplinary knowledge and approaches must be integrated in a way that compliments and sustains the main research goal. As new instruments and methods are added, additional validation considerations are needed. In this work, we will explore an experimental study where surveys and electrodermal sensors were used for one of the semesters (experimental design A) and salivary biomarker collection (i.e., cortisol and sAA) was added to the subsequent semester (experimental design B).

PROTOCOL:

Procedures were approved by the Institutional Review Board (IRB) under a general review at Utah State University for studies on human subjects and use of these constructs. The representative results include two semesters of an engineering statics course, each with a slightly different experimental setup, at a western institution of higher education in the United States. Practice exams, whose content paralleled the actual exams, were developed by the course instructor and were used for our study. Please note that the protocol outlined below describes concurrent steps, and some steps may overlap.

1. Considerations for Experimental Designs and Integration of Disciplinary Practices

1.1. Experimental Design with Surveys and Electrodermal Sensors

1.1.1. Electrodermal sensors are sensitive. Participants' startle responses, if unintentionally activated, can create a significant spike in EDA response. This is particularly important when considering multiple participants for data collection, whose actions may enhance these startle responses. As such, be sure to set up the workspace carefully to minimize as many distractions as possible. As shown in **Figure 1**, include a testing shield if exploring examination experiences for an individual or group of individuals.

NOTE: To increase the ecological validity of the testing environment, plan to provide any material that a student would be using on their actual exam (e.g., workbooks, equation sheets) to allow participants to reflect upon and work out any needed exam problems

1.1.2. Electrodermal sensors provide signal every 1/4th of a second. To allow an event to be defined and studied, implement a plan to collect a precise measure of the onset of a task. When time synching electrodermal sensors with surveys, make sure that the presentation of the survey question is synchronized to the electrodermal sensor by using the internal clock of the computer to establish a data collection timeframe (see **Figure 1**). If using any Bluetooth-enabled electrodermal sensors (e.g., see **Table of Materials**), synch times in Greenwich Meridian Time (GMT) to account for time zone changes and daylight savings time differences during data collection procedures[30].

NOTE: If using a web server for the presentation of stimuli (e.g., test question, survey item, etc.), be sure to align the times between the server and the computer internal clock as these are not typically synched. Note that it may be necessary to pre-install a cross-platform web server (e.g., XAMPP or other Apache servers) to each computer used for the study. If intending to sync a web camera for video recording purposes, consider using security software that allows recording of the date, time, hour, minute, second and millisecond (e.g., 01/01/2000 04:01:02:05) of the video. Note that this video must also be synched with the computer's internal clock and the other devices (e.g., EDA sensor). Set the web cameras to measure the participant's face at different angles, if needed. We recommend that for a frontal faced web camera, the video is positioned parallel to the workstation surface and for downward facing web cameras to position the video at 30° to 45° from the workstation surface to the participant's face.

1.1.3. Place the electrodermal sensor on the non-dominant hand of the participant to minimize any noise in the signal due to movement or electrode contact error during data collection, as suggested in a prior protocol[30]. If researchers would like to minimize artifacts in the EDA due to movement, one alternative is to include a wrist gel pad in a location that is comfortable to the participant and that simultaneously allows them to rest their non-dominant hand on.

NOTE: The placement of the laptop computer, gel pad, sensor, exam sheets, and other elements in the study must be standardized to ensure repeatability across examination conditions and

semesters. As shown in **Figure 1**, painter's tape was used to center each item (e.g., laptops, exam sheets, cameras) of the experimental setup consistently across participants and semesters of data collection.

1.1.4. For electrodermal sensor readings, establish a period of time during which participants have achieved a relaxed state to establish baseline EDA data[31]. For this, either indicate a period of time at the beginning of the exam for participants to stare at the testing shield (~5–15 minutes) or program this cue into the laptop computer as part of the timestamping program. Upon completing this time period, participants can commence with any pertinent surveys and exam questions. In the same vein, assign a relaxation period at the end of the exam experience.

1.2. Experimental Design with Surveys, Electrodermal Sensors, and Salivary Biomarkers

1.2.1. When integrating electrodermal sensors with surveys and salivary biomarkers, ensure that disruptions are minimized to the best extent possible. As one strategy, create a training video to help participants understand how to provide their own salivary samples at set time periods of the exam according to manufacturing specifications (see **Table of Materials**) to minimize interruptions from the researchers.

NOTE: In this study, the researchers were interested in collecting saliva during 4 time-points: beginning, middle, end, and post-exam. However, researchers can choose other times they deem appropriate for their study. Also, we used the swab collection method[32] instead of its passive drool method[33] for ease of use and faster sample collection times. Also, we selected cortisol[34] and sAA[35] kits (see **Table of Materials**) and followed manufacturer specifications in its processing. However, if your group does not have a biological lab to conduct these forms of testing, there are other providers that may be able to analyze the samples[32,36].

1.2.2. When collecting saliva samples, have a cooler with dry ice with an internal temperature of -20 °C; this will prevent room-temperature degradation of enzymes for the cortisol samples[34]. If collecting salivary alpha amylase, its stability is much longer (~5 days at room temperature and allowing for 5 freeze-thaw cycles[35]). If collecting both, as was the case in this study, follow the guidelines needed to store cortisol salivary samples according to manufacturer recommendations[34,35].

1.2.3. If using the swab collection method[25], have the swab remain either in the inner cheek or under the tongue of the participant for 60 s. When handling the vials and sample collection caps, follow manufacturer protocols[34,35] and convey the information to participants prior to the commencement of the study.

NOTE: If the experiment is more granular (e.g., question by question data collection), make sure to record the onset and offset times of each salivary sample collection, as these may need to be accounted for in the EDA analysis. The same applies for the onset and offset of survey data collection times. For salivary data collection, our group developed a flagging system to allow participants to notify the researcher/proctor, when a salivary sample was ready to be collected.

Consider designating multiple proctors to assist with during an experimental session in case multiple salivary samples are ready to be collected and stored.

2. Setup and cleaning pre- and post-experiment

2.1. Surveys

2.1.1. In survey form, organize a scheduling process, designate participant IDs, and collect any demographic information, as needed. Also, establish or pre-label any pertinent survey questions in preparation for data export. This will enable faster and more efficient data cleaning, management, and statistical analyses.

2.1.2. Sync the survey presentation and exit times throughout the exam protocol. If integrating sensors or video, sync these technologies with the survey software as well.

2.1.3. As a matter of courtesy and in the interest of contributing to a cordial and welcoming research environment, and if instructors agree, set up an automated follow-up email containing responses to the exam questions to be sent to participants immediately or soon after their participation in the session.

2.2. Electrodermal Sensors

2.2.1. Plan to pre-schedule participants to an examination session/time, assess any medical information and dietary habits for EDA and saliva collection[30] and hand dominance for EDA collection[30], and remind participants to avoid consumption of sugary or caffeinated products the day of the experiment. This is important as certain medical conditions (e.g., metabolic disorders) and dietary habits (e.g., caffeine consumption) can influence EDA (and salivary values), as suggested in a prior protocol[30].

2.2.2. Before participants arrive, make sure sensors are properly calibrated, software updates have been taken care of, and sensors have been cleaned with 70% alcohol wipes[30].

2.2.3. When fitting the EDA sensor on participant's wrists, make sure to place it on the participant's non-dominant hand. To fit the EDA sensor:

2.2.3.1. Place the sensor with the button facing down towards the thumb.

2.2.3.2. With their palms facing up towards their face, have participants draw an imaginary line from the space between the second and third finger of their non-dominant hand to their mid-wrist area and place the sensor electrodes there.

2.2.3.3. Ask participants to fit the sensor straps in a way that is not too tight or too loose.

NOTE: A representative image of this fitting can be found in **Figure 2**.

2.2.4. When starting the sensor, be sure to follow manufacturer protocols[31] to ensure the sensors are set up to collect data. In this experiment, the protocol is tailored to use with a particular brand of sensors (see **Table of Materials**) although researchers are welcome to use any physiological sensor of their choosing.

2.2.4.1. For devices used here, depress the sensor button for three seconds. A green light will blink intermittently, followed by a red blinking light, and then a fade out occurs.

2.2.4.2. During fade-out, to make sure the sensor is ON, press the button once for less than 1 s. If it blinks red, it is indicating that it is recording data.

2.2.5. When turning the sensor OFF, press the button for 3 s. The sensor will turn off if the lights on the bottom of the wristband go from green to fade.

2.2.6. To retrieve the data from the sensor, connect it to the computer and upload the data in the managing software system according to manufacturer recommendations[31].

2.3. Salivary Biomarkers

2.3.1. As stated before, pre-assess any medical conditions or dietary habits that may influence salivary values during analysis. Also, remind participants not to wear any lip balm, make-up, or products near the lips when they arrive to the session, as this could introduce contaminants that may influence cortisol and salivary alpha amylase samples. If participants arrive wearing these products, gently guide them to a restroom or provide appropriate wipes that would remove these products without introducing other chemicals (e.g., water on a napkin versus make-up remover towels). Finally, clear experiment rooms of food or drinks that have a strong smell (e.g., pizza, oranges) that may enhance salivary production among participants.

2.3.2. Upon participants' arrival to the experimental room, hand participants 1 ounce of water poured into a cup in their presence. Ask them to swish and swallow the water. This is done to clear the mouth of any food residues that may influence the cortisol and salivary alpha amylase data.

2.3.3. If collecting EDA data in conjunction to saliva, gently remind participants to minimize hand movement in the hand that has the EDA sensor. As such, participants will need to be informed that any saliva sample collection provided has to be done in their dominant hand. To facilitate this process, it is recommended that the experimental setup includes pre-labeled vials and a stand to minimize any loss of samples (refer to **Figure 1**).

2.3.4. When collecting salivary samples, wear fresh nitrile gloves to minimize any dust particulate or any other contaminant from hand oils to be transferred to the salivary sample vial.

2.3.5. As indicated previously, immediately transfer the samples to a cooler that has an internal

temperature of -20 °C.

3. Increasing ecological validity in light of surveys, electrodermal sensors, and salivary biomarkers

3.1. Concerning exam authenticity

3.1.1. To provide an authentic testing experience, align the exam content with course content. For this, review the course content in conjunction with a group of content experts including the course instructor.

3.1.2. Select an evaluation (test or assessment) of the course content that can be replicated in an experimental setting or that can complement existing course content (e.g., practice exam).

NOTE: Depending on the Institutional Review Board policies of your institution, using real exams may not be allowed due to its potential harm to students' grades in the course. As such, an equivalent experience (e.g., practice exam) may be considered instead.

3.1.3. Alongside the instructor, develop an answer key and exam problems and its solutions to be used to collect performance data at a granular level (i.e., question by question) and/or macro-level (i.e., entire exam) depending on the goals of the research

3.1.4. Ask the instructor to also provide any additional materials that are typically used in their exams (e.g., cheat sheets) or any allowable materials (e.g., textbooks, list of references) typically used in their courses. Experimenters should be prepared to provide these tools to the participants.

3.1.5. Make sure that the testing environment parallels the experimental setup (e.g., exam times, offering of exam—testing center or classroom, etc.) and its features such as desk space, lighting, temperature of the room, among others.

3.2. Concerning survey inclusion

3.2.1. Depending on the amount of survey questions, it will be important to account for the approximate times it might take participants to complete the survey questions while they are taking their exam.

3.2.2. Allot additional test taking time to account for interruptions and design the examination program to return students to a particular exam problem if a survey prompt interrupted them. Also, be sure this interruption time is consistent across participants (e.g., beginning, middle, and end of exam).

3.2.3. Depending of the type of experimental design, if granular type of responses are needed (e.g., question by question), plan to present the exam problem first, then prompt participants to

respond to the survey question, and then allow participants to enter their response (e.g., open-text, multiple choice, etc.). This will allow participants to first view the problem and respond to the survey question according to the presented problem. If the experimental design is on a macro-level, make sure that participants are allowed to reflect on the exam experience up to that point before providing a response.

NOTE: Theories and hypotheses are important to consider in this step as the choice of the particular kind of presentation of an item (e.g., survey, exam) will matter. For example, if studying self-efficacy, this is best assessed at the level of the test question, while academic achievement emotions are typically asked pre-, during-, and post-exam.

3.3. Concerning electrodermal activity sensors

3.3.1. To ensure participants are not being overly stressed due to the experimental protocol, include calibration and relaxation periods throughout the exam experience. One strategy could be to allow participants to refocus their attention between questions. Beginning with a simple-to-respond question (e.g., "what day of the week are we in?") and allow participants 30 s to rest in between each exam question.

NOTE: Keep in mind that understanding the design of the exam questions itself and predicting what students' reactions may be important (e.g., increased cognitive loads or neural efficiencies[37]) as they could influence the salivary marker and EDA data collection. For example, the exam questions should all be in the form of essay entry, which would require hand movement that can influence EDA data[24,25] or an exam may be designed by varying levels of difficulty, which could influence students' cognitive loads or neural efficiencies[37].

3.3.2. Ensure that the time stamping program will account for any changes in the examination experience (e.g., calibration periods, onset and offset of in-between calibration questions, survey questions onset and offset, start and finish of the exam). This is an important step as it will allow for data source matching, which will determine the intervals or events to be processed and analyzed.

3.4. Concerning salivary biomarker use

3.4.1. Be mindful of when to collect salivary biomarkers.

NOTE: Salivary bio-marker studies typically are explored through a pre-pre-mid-post-post-post design[32-36]. As cortisol takes 20 minutes to respond to stress[14], these time lags are needed to observe cortisol onset and recovery. In the case of students' preparing for an exam, participants may be worried about taking the exam, and, thus, a before-onset measure may not be possible. It is also important not to interrupt students frequently during the exam. In our study, we opted to collect saliva once prior to onset, once during, immediately after, and 20 minutes after the exam as quietly as possible to minimize disruptions. A sample testing timeline is provided in **Figure 3**.

3.4.2. In the examination program, include timed prompts to cue participants when it is time to collect saliva. Include a 60-s timer so participants are aware of the duration of the salivary collection. Return participants to the problem they were working on in the exam once the 60 s are complete.

4. Considerations for data processing and analysis

4.1. Survey

4.1.1. Be sure that data outputs are labelled and organized appropriately to allow for effective data management and ensure statistical programs (e.g., SPSS, SAS) are able to perform any needed analysis.

4.1.2. Identify any potential outlier data based upon standards for survey outlier detection[38] as well as any determined through the demographic data collected previously (e.g., medical conditions).

4.1.3. Determine the type of statistical analysis and/or modeling to conduct based upon the established research question(s) and/or hypotheses

4.2. Electrodermal Activity

4.2.1. Note that electrodermal data outputs may vary by company. For the device used in this study[31], data outputs are presented as a single column with a starting time measured in GMT followed by the frequency of data collection and the EDA measured in microSiemens. The EDA data then increments according to the frequency of data collection. Since the data is dependent on the time of onset, convert this time to UNIX time according to manufacturing protocols and previous protocols[30]. This will allow a more seamless synchronization of the EDA data changes throughout the experiment.

4.2.2. Identify and remove any potential manufacturer sources of outliers, such as sensor malfunction, incomplete data collection, or poor contact of the electrodes in the skin. These will be identified by negative values or constant near-zero continual data segments in the data output sheet.

4.2.3. Identify and remove any potential user-generated sources of outliers such as erratic movements (e.g., hand hitting desk or nervous tapping), survey or salivary biomarker collection periods, or vast changes in body temperatures or blood volume pressure readings.

4.2.4. To remove noise due to movement, do the following series of steps, outlined in **Figure 4**:

4.2.4.1. First, scan through the participants' accelerometer (ACC) profiles, also provided by the wrist sensor. Note that the data will have X, Y, and Z columns indicating three-dimensional

horizontal, vertical, and spatial hand movements respectively. Calculate the moving average of this accelerometer data according to the Euclidean Distance (L2-Norm)[39] equation to calculate the total movement:

$$\mathrm{L2-Norm} = \sqrt{\mathrm{X}^2 + \mathrm{Y}^2 + \mathrm{Z}^2}$$

4.2.4.2. Calculate the standard deviation of the Euclidean distance values for the entire participant set and rank-order them. Calculate the average values of the Euclidian distance values too.

4.2.4.3. Calculate the coefficient of variance of the Euclidean distance values to determine the signal-to-noise ratios[40] according to the following equation:

$$\text{Coefficient of Variance} = \frac{\text{average}}{\text{standard deviation}}$$

NOTE: Coefficient of variance values that exceed a score of 1 indicates an outlier and must be removed from analysis according to recommendations in handling signaling data[33].

4.2.4.4. Once the noise due to movement is removed, determine the needed threshold to filter the data. For this, calculate the upper and lower limits of the 95% of the standard deviation of the signals. Any data outside these ranges can be either removed from the dataset/analysis or imputed according to the researcher´s goals and objectives. For this study, we opted to average the outside ranges with the determined acceptable data.

4.2.4.5. Return to the EDA data and use the time stamped accelerometer data to identify the corresponding intervals of EDA (which have also been time-stamped).

NOTE: To sync accelerometer and electrodermal data, note that the recording frequencies are different (4 Hz for EDA and 32 Hz for ACC) so they must first be aligned. Since, inherently, there will be more ACC data than EDA data, use the average EDA values to account for this difference.

4.2.5. Once EDA data sets have been cleaned[41,42] though the filtered accelerometer data, proceed to separate out the tonic (baseline) and phasic (immediate, reactive) signals using prescribed tools (e.g.., Ledalab, EDA Explorer)[43,44]. For statistical analysis, primarily the phasic, filtered EDA data are used and values (e.g., magnitudes, number of peaks, latency times) are calculated based upon the research question/hypothesis and using methods described by Bouscien[22,23].

4.3. Salivary Biomarker

4.3.1. For both cortisol and salivary alpha amylase assays, follow manufacturer protocols[22-28] and technician recommendations about terms of use, storage, and handling samples.

4.3.2. Spin thawed samples at 1,500 x *g* at 4 °C. Be sure to carefully remove the swabs and that the vials have salivary supernatant at the bottom of the vial to ensure mucin separation.

4.3.3. As good practice, before following the assay protocols, do a buffer rinsing of the wells using a plate washer prior to processing. This is particularly important for cortisol.

4.3.4. Ensure that the optical density plate reader has been pre-programmed to the appropriate temperatures (e.g., sAA samples require incubation temperatures of 37 °C whereas cortisol samples requires room temperature readings) and wavelengths (i.e., sAA requires 405 nm and cortisol requires 450 nm and 490–492 nm reference filters). For sAA assays, it is recommended that the plate reader used has both a shaker and an incubator inside.

4.3.5. Follow manufacturer protocols[34,35] to calculate the concentration values of each sample and the corresponding intra- and inter-assay percent of coefficient of variation (%CV) equations to identify outliers from the data set (this is calculated differently compared to the equation provided previously). Please note that, for sAA, keep track of the lot numbers used in the controls as they are not standardized.

4.3.5.1. First, average the %CV of the controls by lot number and then average these values to get a grand average %CV score.

4.3.5.2. For samples, the manufacturer recommends that the intra-assay of samples should have a %CV under 10% while the controls should have an inter-assay %CV under 15%[34,35]. However, these %CV values will greatly depend on the laboratory conditions and equipment used to conduct the research. As such, consider alternate methods of immunoassay assay validation as needed[45].

4.3.6. Freeze saliva samples at -80 °C after the assay to allow verification of its validation. Do not freeze thaw more than once to prevent further enzymatic degradation of the samples or controls.

4.4. Data Triangulation

4.4.1. Depending on the research question or hypothesis, correlate relevant variables. Ensure that all outliers and data are properly pre-processed and filtered before use[46].

4.4.2. Determine if the sample size, data collection points, observed statistical power, and research questions or hypothesis necessitate amalgamating data[47], or utilizing repeated-measures analytic techniques[48-50].

4.4.3. Accounting for inter-individual differences in task time[51] and the delay in response of salivary biomarkers to stress[14], use timestamps or determine events to sync datasets together.

4.4.4. Using statistical models and software, analyze the data set and interpret findings.

REPRESENTATIVE RESULTS:

In this study, we were interested in studying the influences of self-efficacy, performance, and physiological (EDA sensors) and biological (sAA and cortisol) responses of undergraduate engineering students as they took a practice exam. The data shown is a representative sub-set of exams: (a) one that considered surveys and electrodermal sensors (experiment design A) and (b) one that included the same exam along with the salivary biomarker data (experiment design B). While we collected emotions data in this study, we will not present it, as our goal was to demonstrate granular data in real-time rather than at prescribed timepoints at the beginning, middle or end of the exam, which is where emotions data was collected.

As shown in **Figure 4**, the degree of difficulty of the exam according to the collective response of students was compared across the experimental designs. Also, the mean EDA as a function of students´ reported self-efficacy scores prior to completing the exam questions was plotted. Even though the degree of difficulty was the same for the two designs, opposing differences in the mean EDA values were found between the correct and incorrect responses across different self-efficacy scores. For experimental design A (EDA sensors and surveys), mean EDA increased for a mid-SE score for students who responded incorrectly to the exam questions compared to students who responded the questions correctly ($p < 0.001$). For experimental design B (EDA sensors, surveys, and salivary biomarkers), mean EDA values varied where an opposite effect was found for low SE scores ($p < 0.05$) and high SE scores ($p < 0.01$), respectively.

To understand any potential salivary influences, the mean EDA as well as cortisol and sAA assay values for set data points in the exam (beginning, middle, end, and 20-minutes after the exam) were normalized (**Figure 5**) for experimental design B. It is important to note that the mean EDA values for this table were truncated at 60-second intervals during the pre-set timeframe to allow for comparisons between each salivary marker. The data suggests that EDA levels decreased from beginning to the end of the exam and these levels recovered by the 20-minute mark after the exam. These trends were paralleled in the cortisol and sAA data. Statistical significance, as determined through ANOVA, was found between EDA and sAA at the beginning and middle of the exam ($p < 0.05$ for both times) whereas EDA and cortisol showed significance between the middle and end of the exam ($p < 0.01$ and $p < 0.05$, respectively). By the 20-minute mark, EDA and sAA ($p < 0.01$) and cortisol and sAA ($p < 0.05$) began to show significance between each other.

FIGURE LEGENDS:

Figure 1. Experimental setup when using surveys and electrodermal sensors to study examination experiences. The image shows Experimental Design A (sensors and survey) and B (sensors, survey, and salivary biomarkers).

Figure 2. A schematic representation of how participants can fit and start the electrodermal sensor. Image A (in the left) shows the placement of the start button on the sensor while Image B (on the right) shows the placement of the EDA electrodes on the wrist of the participant.

Figure 3. Representation of an experimental timeline when surveys, salivary biomarkers, and electrodermal sensors are included.

Figure 4. Degree of difficulty. Degree of difficulty of the exam according to collective student performance and mean EDA as a function of self-efficacy scale ranking by participants for the correct and incorrect responses for experimental design A (**A** and **B**) and experimental design B (**C** and **D**). N = 15 participants per design; data is reported as mean ± standard error of the mean (represented in the error bars); dashed lines on panels A and C represent the limits for moderate ranges of difficulty (between 0.3 to 0.8)[52] ; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Figure 5. Normalized sAA, cortisol and mean EDA. Normalized sAA, cortisol and mean EDA for experimental design B compared at 60-s intervals at prescribed time periods during the exam (beginning, middle, end, 20 minutes after). N = 15; data is reported at mean ± standard error of the mean (represented in the error bars); *$p < 0.05$, **$p < 0.01$.

DISCUSSION:

Although physiological measures have been used in many authentic learning contexts, it is critical to design a study environment that is mindful of the limits of the current technology. Our design balances the need for an authentic testing environment and accommodates the technology. Comfortably limiting participant movement, reducing unintended interruptions, and timestamping participants' testing responses are all critical steps within the protocol.

The space and expense of the electrodermal sensor devices may make the study impractical for researchers with limited research funds. However, once purchased, these sensors have unlimited uses. Salivary biomarkers must be processed in a laboratory and have significant per-sample pre- and post-processing expenses. It is also important to consider the particular laboratory conditions and equipment used, as alternate salivary assay validation methods may be needed to identify inter- and intra-assay percentages of CV.

The protocol is a significant step forward in the application of multi-modal approaches in the study of academic emotions. The protocol maximizes the precision of EDA measurements by timestamping participant responses while replicating an authentic testing environment, which enables more objective real-time studies of student coursework and classroom studies, addressing a constraint that limited prior research studies focused on learning and performance. It is possible to modify the technique to include online learning activities that requires keystroke capture. It is also possible to use the protocol for deception studies in where the difficulty of the test or present text-based prompts are pre-designed to influence students' expectations for the test.

ACKNOWLEDGMENTS:

This material is based upon work supported in part by the National Science Foundation (NSF) No. EED-1661100 as well as a NSF GRFP grant given to Darcie Christensen (No. 120214). Any opinions, findings, and conclusions or recommendations expressed in this material do not necessarily reflect those of NSF or USU. We want to thank Sheree Benson for her kind discussions and recommendations for our statistical analysis.

Author contributions in this paper are as follows: Villanueva (research design, data collection and analysis, writing, editing); Husman (research design, data collection, writing, editing); Christensen (data collection and analysis, writing, editing); Youmans (data collection and analysis, writing, and editing); Khan (data collection and analysis, writing, editing); Vicioso (data collection and analysis, editing); Lampkins (data collection and editing); Graham (data collection and editing)

DISCLOSURES:

The authors have nothing to disclose.

REFERENCES:

1. William, J. What is an emotion? *Mind*. **9** (34), 188–205 (1884).
2. Pekrun, R., Linnenbrink-Garcia, L. Emotions in education: Conclusions and future directions. In Pekrun, R., Linnenbrink-Garcia, L. (Eds.), *International handbook of emotions in education.* 659–675. London: Routledge Press (2014).
3. Pekrun, R. The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*. **18** (4), 315–341 (2006).
4. Pekrun, R., Perry, R. P. Control-value theory of achievement emotions. *International Handbook of Emotions in Education*. 120–141 (2014).
5. Pekrun, R., Stephens, E. J. Academic emotions. In Harris, K. R. et al. (Eds.), *APA Educational Psychology Handbook.* 3–31. Washington, D.C.: American Psychological Association (2011).
6. Bandura, A. *Self-efficacy: The exercise of control*. New York, NY: W. H. Freeman & Co (1997).
7. Bandura, A. *Social foundations of thought and action: A social cognitive theory.* Upper Saddle River, New Jersey: Prentice Hall (1986).
8. Bandura, A. Guide for constructing self-efficacy scales. In Pajares, F., Urdan, T. (Eds.), *Self-efficacy beliefs of adolescents.* 307–337. Charlotte, NC: Information Age Publishing (2006).
9. Jarrell, A., Harley, J. M., Lajoie, S., Naismith, L. Success, failure and emotions: examining the relationship between performance feedback and emotions in diagnostic reasoning. *Educational Technology Research and Development.* **65** (5), 1263–1284 (2017).
10. Pekrun, R., Bühner, M. Self-report measures of academic emotions. In Pekrun, R., Linnenbrink-Garcia, L. (Eds.), *International Handbook of Emotions in Education*. 561–566. London: Routledge Press (2014).
11. Nett, U. E., Goetz, T., Hall, N. C. Coping with boredom in school: An experience sampling perspective. *Contemporary Educational Psychology*. **36** (1), 49–59 (2011).
12. Azevedo, R. Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues. *Educational Psychologist*. **50** (1), 84–94 (2015).
13. Spangler, G., Pekrun, R., Kramer, K., Hofman, H. Students' emotions, physiological reactions, and coping in academic exams. *Anxiety, Stress, & Coping*. **15** (4), 413–432 (2002).
14. Husman, J., Cheng, K. C., Puruhito, K., Fishman, E. J. Understanding engineering students stress and emotions during an introductory engineering course. *American Society of Engineering Education.* Paper ID #13148 (2015).
15. Vedhara, K., Hyde, J., Gilchrist, I., Tytherleigh, M., Plummer, S. Acute stress, memory, attention and cortisol. *Psychoneuroendocrinology*. **25** (6), 535–549 (2000).
16. Berenbaum, S. A., Moffat, S., Wisniewski, A., Resnick, S. Neuroendocrinology: Cognitive

effects of sex hormones. In de Haan, M., Johnson, M. H. (Eds.), *The Cognitive Neuroscience of Development: Studies in Developmental Psychology*. 207–210. Psychology Press (2003).
17. Lundberg, U., Frankenhaeuser, M. Pituitary-adrenal and sympathetic-adrenal correlates of distress and effort. *Journal of Psychosomatic Research*. **24** (3-4), 125–130 (1980).
18. Nater, U. M., Rohleder, N. Salivary alpha-amylase as a non-invasive biomarker for the sympathetic nervous system: Current state of research. *Psychoneuroendocrinology*. **34** (4), 486–496 (2009).
19. Denson, T., Spanovic, M., Miller, N., Cooper, H. Cognitive appraisals and emotions predict cortisol and immune responses: A meta-analysis of acute laboratory social stressors and emotion inductions. *Psychological Bulletin*. **135** (6), 823–853 (2009).
20. Van Stegeren, A. H., Wolf, O. T., Kindt, M. Salivary alpha amylase and cortisol responses to different stress tasks: Impact of sex. *International Journal of Psychophysiology*. **69** (1), 33–40 (2008).
21. Benedek, M., Kaernbach C. A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*. **190** (1), 80–91 (2010).
22. Boucsein, W., Backs, R. W. (2000). Engineering psychophysiology as a discipline: Historical and theoretical aspects. In Backs, R. W., Boucsein, W. (Eds.), *Engineering psychophysiology. Issues and applications*. 3–30. Mahwah, NJ: Lawrence Erlbaum (2000).
23. Boucsein, W., Backs, R. W. The psychophysiology of emotion, arousal, and personality: Methods and models. In V. G. Duffy (Ed.), *Handbook of digital human modeling.* 35–38. CRC: Boca Raton (2009).
24. Turpin, G., Shine, P., Lader, M. H. Ambulatory electrodermal monitoring: effects of ambient temperature, general activity, electrolyte media, and length of recording. *Psychophysiology*. **20**, 219–224 (1983).
25. Posada-Quintero, H. F. et al. Timevarying analysis of electrodermal activity during exercise. *PLoS ONE*. **13** (6), e0198328 (2018).
26. Lobstein, T., Cort, J. The relationship between skin temperature and skin conductance activity: Indications of genetic and fitness determinants. *Biological Psychology*. **7**, 139–143 (1978).
27. Scholander T. Some measures of electrodermal activity and their relationships as affected by varied temperatures. *Journal of Psychosomatic Research.* **7**, 151–158 (1963).
28. Schwerdtfeger, A. Predicting autonomic reactivity to public speaking: don't get fixed on self-report data! *International Journal of Psychophysiology*. **52** (3), 217–224 (2004).
29. Braithwaite, J. J., Watson, D. G., Jones, R., Rowe, M. A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments. *Psychophysiology*. **49** (1), 1017–1034 (2013).
30. Villanueva, I., Valladares, M., Goodridge, W. 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** (e53255), 1–8 (2016).
31. Empatica. *E4 wristband from Empatica: User's manual. Empatica*. 1–32 (2018).
32. Salimetrics. *Collection methods: SalivaBio oral swab (SOS). Salimetrics Technical Summary.* 1–2 (2018).
33. Salimetrics. *Collection methods: Passive drool using the saliva collection aid. Salimetrics Technical Summary*. 1–2 (2018).

34. Salimetrics. *Expanded range high sensitivity salivary cortisol enzyme immunoassay kit. Salimetrics Technical Summary*.1–21 (2016).
35. Salimetrics. *Salivary α-amylase kinetic enzyme assay kit. Salimetrics Technical Summary*. 1–17 (2016).
36. Moore, D. Innovative Hormone Testing: Saliva Test Specifications, ZRT Laboratory Reports, Available: http://www.zrtlab.com/resources-and-data (2014).
37. Call, B., Goodridge, W., Villanueva, I., Wan, N., Jordan, K. Utilizing electroencephalography measurements for comparison of task-specific neural efficiencies: spatial intelligence tasks. *Journal of Visualized Experiments*. (114), e53327 (2016).
38. Ruel, E. E., Wagner III, W. E., Gillespie, B. J. *The practice of survey research: theory and applications*. SAGE Publications, Thousand Oaks: CA (2016).
39. Barrett, P. Euclidean distance: raw, normalized, and double-spaced coefficients. *The Technical Whitepaper Series*. **6**, 1–26 (2005).
40. Groeneveld, R. A. Influence functions for the coefficient of variation, its inverse, and CV comparisons. *Communications in Statistics- Theory and Methods*. **40** (23), 4139–4150 (2011).
41. Tronstad, C., Staal, O. M., Sælid, S., Martinsen, Ø. G. Model-based filtering for artifact and noise suppression with state estimation for electrodermal activity measurements in real time. *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2750-2753 (2015).
42. Routray, A., Pradhan, A. K., Rao, K. P. A novel Kalman filter for frequency estimation of distorted signals in power systems. *IEEE Transactions on Instrumentation and Measurement*. **51** (3), 469–479 (2002).
43. Benedek, M., Kaernbach, C. A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*. **190**, 80–91 (2010).
44. Taylor, S. et al. Automatic Identification of Artifacts in Electrodermal Activity Data. *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 1934–1937 (2015).
45. Andreasson, U. et al. A practical guide to immunoassay method validation. *Frontiers in Neurology*. **6** (179), 1–8 (2015).
46. Adam, E. K., Kumari, M. Assessing salivary cortisol in large-scale, epidemiological research. *Psychoneuroendocrinology*. **34** (10), 1423–1436 (2009)
47. Pruessner, J. C., Kirschbaum, C., Meinlschmid, G., & Hellhammer, D. H. Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. *Psychoneuroendocrinology*. **28** (7), 916–931 (2003).
48. Girden, E. R. *ANOVA: Repeated measures.* Thousand Oaks, CA: Sage (1992).
49. Raudenbush, S. W., Bryk, A. S. *Hierarchical linear models: Applications and data analysis methods (Vol. 1).* Thousand Oaks, CA: Sage (2002).
50. Duncan, T. E., Duncan, S. C., Strycker, L. A. *An introduction to latent variable growth curve modeling: Concepts, issues, and application.* Abingdon, United Kingdom: Routledge (2013).
51. Mehta, P. D., West, S. G. Putting the individual back into individual growth curves. *Psychological Methods*. **5**(1), 23–43 (2000).

Figure 1

Click here to access/download;Figure;Figure 1_JOVE_final_5.25.19.pdf

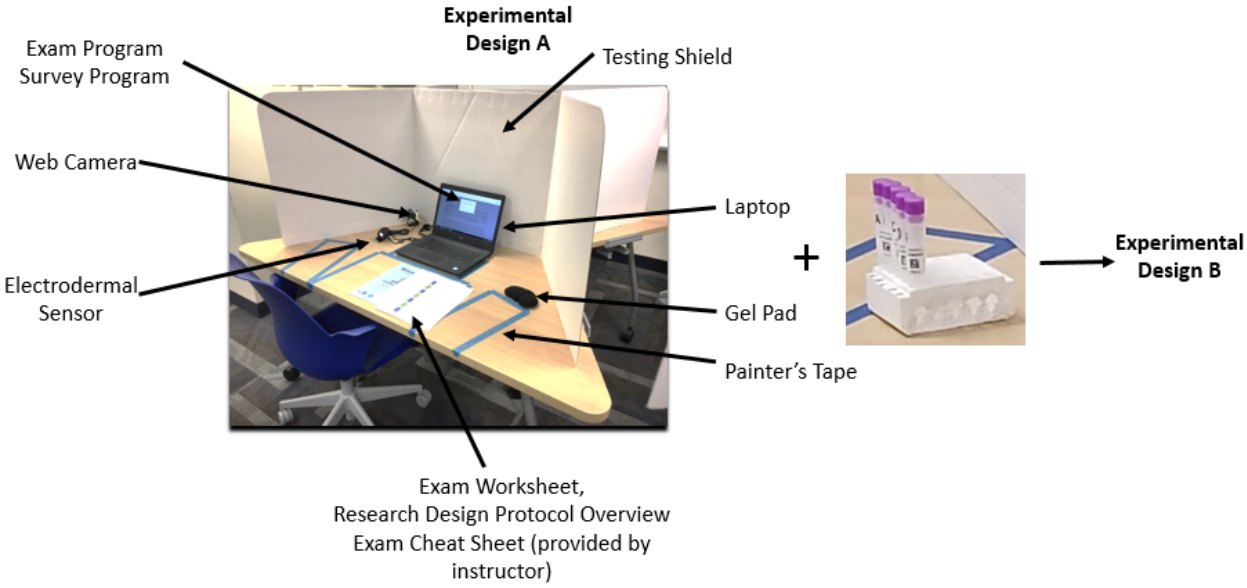


Figure 2

Click here to access/download;Figure;Figure 2_JOVE_5.25.19.pdf

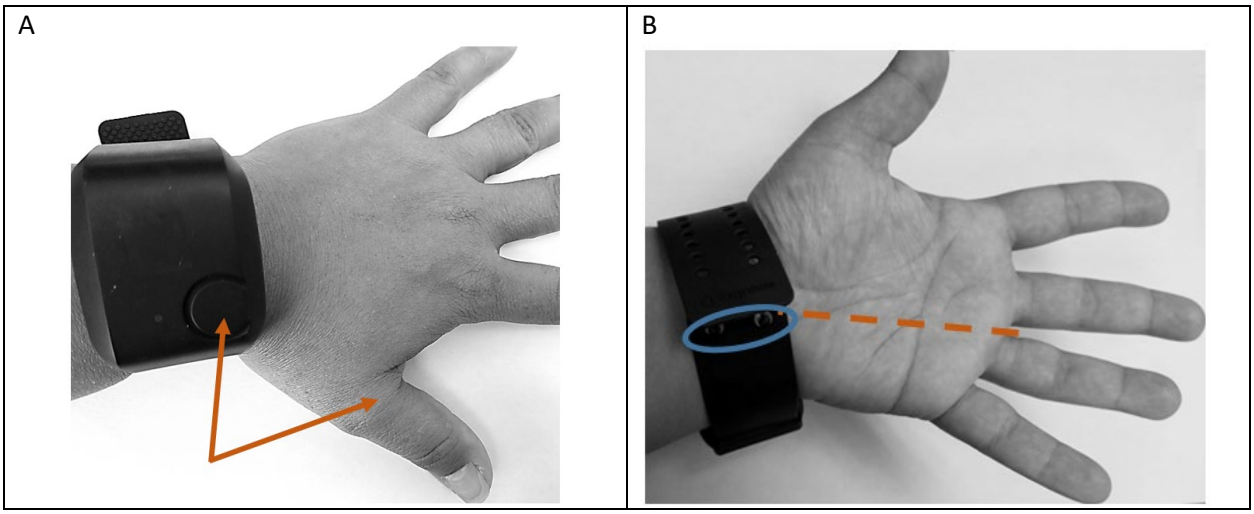


Figure 3

Click here to access/download;Figure;Figure 3_JOVE_final_5.25.19.pdf

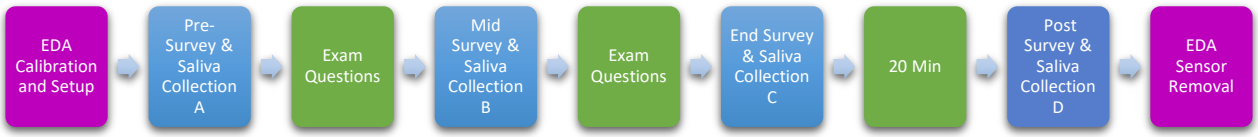


Figure 4

Click here to access/download;Figure;Figure 4_JOVE_final_5.24.19.pdf

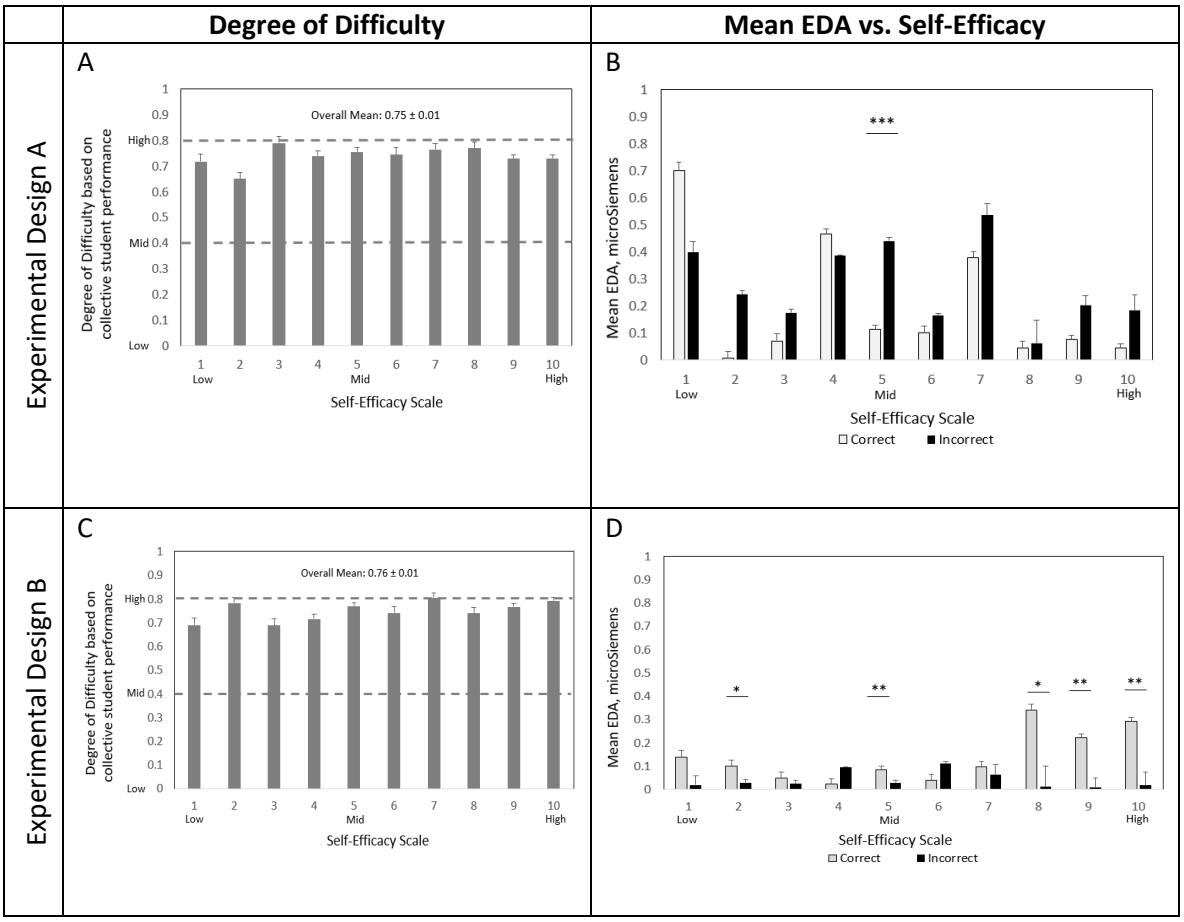
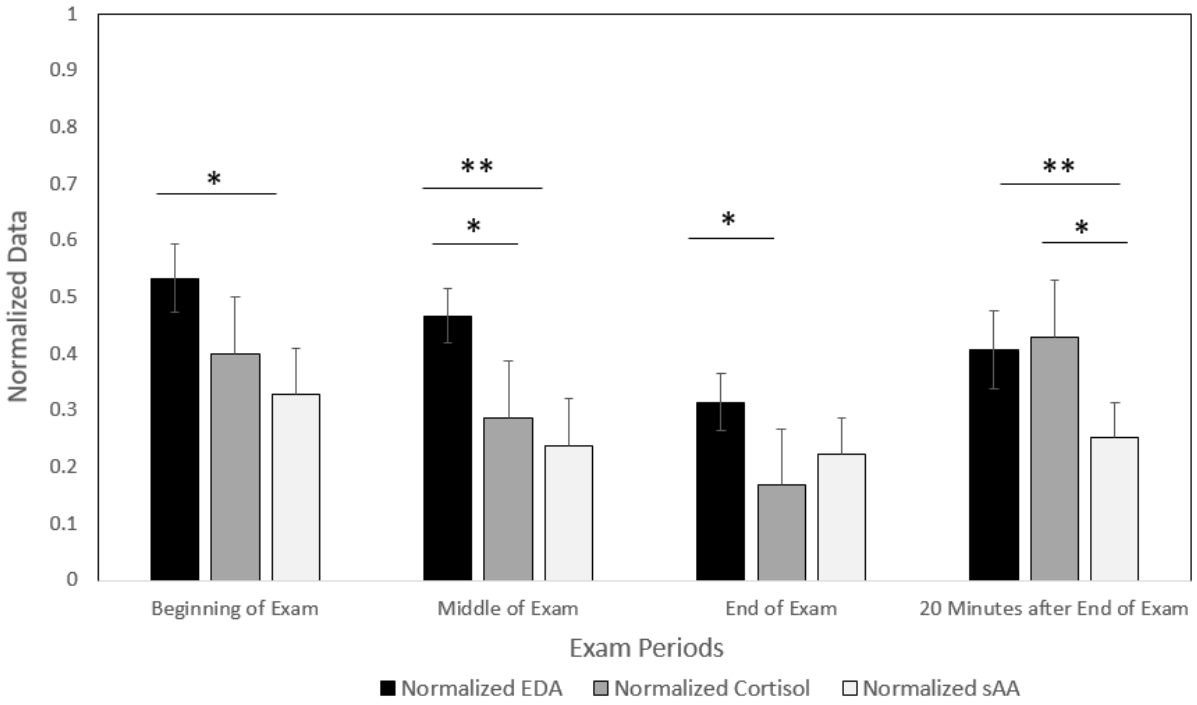


Figure 5

Click here to access/download;Figure;Figure 5_JOVE_5.24.19.pdf



Name of Material/Equipment	Company	Catalog Number	Comments/Description
1.1 cu ft medical freezer	Compact Compliance	# bci2801863	They can use any freezer as long as it can go below -20 degrees Celsius; these can be used to store salivary samples for longer periods of time (~4 months) before running salivary assays.
Camping Cooler	Amazon	(any size/type)	Can be used to store salivary samples during data collection
E4 sensor	Empatica Inc	E4 Wristband Rev2	You can use any EDA sensor or company as long as it records EDA and accelerometry
EDA Explorer	https://eda-explorer.media.mit.edu/	(open-source)	Can be used to identify potential sources of noise that are not necessarily due to movement
Laptops	Dell	Latitude 3480	They can use any desktop or laptop
Ledalab	http://www.ledalab.de/	(open-source)	Can be used to separate tonic and phasic EDA signals after following filtration steps

MATLAB	https://www.mathworks.com/products/m	(version varies according to updates)	To be used for Ledalab, EDA Explorer, and to create customized time-stamping programs.
Salivary Alpha Amylase Enzymatic Kit	Salimetrics	# 1-1902	For the salivary kits, you should plan t
Salivary Cortisol ELISA Kit	Salimetrics	# 1-3002	For the salivary kits, you should plan t
Testing Divider (Privacy Shields)	Amazon	#60005	They can use any brand of testing shield as long as they cover the workspace
Web Camera	Amazon	Logitech c920	They can use any web camera as long as it is HD and 1080p or greater

o either order the company to analyze your samples and/or go to a molecular biology lab for processing
o either order the company to analyze your samples and/or go to a molecular biology lab for processing

1 Alewife Center #200
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT

Title of Article:	Considerations for studying near-real-time authentic examination experiences: A cross-disciplinary and multi-modal experimental design
Author(s):	Idalis Villanueva, Jenefer Husman, Darcie Christensen, Kate Youmans, Md Tarique Hasan Khan, Paul Vicioso, Shawn Lampkins, Matt Graham

Item 1: The Author elects to have the Materials be made available (as described at http://www.jove.com/publish) via:

☑ Standard Access ☐ Open Access

Item 2: Please select one of the following items:

☑ The Author is **NOT** a United States government employee.

☐ The Author is a United States government employee and the Materials were prepared in the course of his or her duties as a United States government employee.

☐ The Author is a United States government employee but the Materials were NOT prepared in the course of his or her duties as a United States government employee.

ARTICLE AND VIDEO LICENSE AGREEMENT

1. **Defined Terms.** As used in this Article and Video License Agreement, the following terms shall have the following meanings: "**Agreement**" means this Article and Video License Agreement; "**Article**" means the article specified on the last page of this Agreement, including any associated materials such as texts, figures, tables, artwork, abstracts, or summaries contained therein; "**Author**" means the author who is a signatory to this Agreement; "**Collective Work**" means a work, such as a periodical issue, anthology or encyclopedia, in which the Materials in their entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole; "**CRC License**" means the Creative Commons Attribution-Non Commercial-No Derivs 3.0 Unported Agreement, the terms and conditions of which can be found at: http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode; "**Derivative Work**" means a work based upon the Materials or upon the Materials and other pre-existing works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Materials may be recast, transformed, or adapted; "**Institution**" means the institution, listed on the last page of this Agreement, by which the Author was employed at the time of the creation of the Materials; "**JoVE**" means MyJove Corporation, a Massachusetts corporation and the publisher of The Journal of Visualized Experiments; "**Materials**" means the Article and / or the Video; "**Parties**" means the Author and JoVE; "**Video**" means any video(s) made by the Author, alone or in conjunction with any other parties, or by JoVE or its affiliates or agents, individually or in collaboration with the Author or any other parties, incorporating all or any portion of the Article, and in which the Author may or may not appear.

2. **Background.** The Author, who is the author of the Article, in order to ensure the dissemination and protection of the Article, desires to have the JoVE publish the Article and create and transmit videos based on the Article. In furtherance of such goals, the Parties desire to memorialize in this Agreement the respective rights of each Party in and to the Article and the Video.

3. **Grant of Rights in Article.** In consideration of JoVE agreeing to publish the Article, the Author hereby grants to JoVE, subject to **Sections 4** and **7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Article in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Article into other languages, create adaptations, summaries or extracts of the Article or other Derivative Works (including, without limitation, the Video) or Collective Works based on all or any portion of the Article and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and(c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. If the "Open Access" box has been checked in **Item 1** above, JoVE and the Author hereby grant to the public all such rights in the Article as provided in, but subject to all limitations and requirements set forth in, the CRC License.

4. **Retention of Rights in Article.** Notwithstanding the exclusive license granted to JoVE in **Section 3** above, the Author shall, with respect to the Article, retain the non-exclusive right to use all or part of the Article for the non-commercial purpose of giving lectures, presentations or teaching classes, and to post a copy of the Article on the Institution's website or the Author's personal website, in each case provided that a link to the Article on the JoVE website is provided and notice of JoVE's copyright in the Article is included. All non-copyright intellectual property rights in and to the Article, such as patent rights, shall remain with the Author.

5. **Grant of Rights in Video – Standard Access.** This **Section 5** applies if the "Standard Access" box has been checked in **Item 1** above or if no box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby acknowledges and agrees that, Subject to **Section 7** below, JoVE is and shall be the sole and exclusive owner of all rights of any nature, including, without limitation, all copyrights, in and to the Video. To the extent that, by law, the Author is deemed, now or at any time in the future, to have any rights of any nature in or to the Video, the Author hereby disclaims all such rights and transfers all such rights to JoVE.

6. **Grant of Rights in Video – Open Access.** This **Section 6** applies only if the "Open Access" box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby grants to JoVE, subject to **Section 7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Video in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Video into other languages, create adaptations, summaries or extracts of the Video or other Derivative Works or Collective Works based on all or any portion of the Video and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. For any Video to which this **Section 6** is applicable, JoVE and the Author hereby grant to the public all such rights in the Video as provided in, but subject to all limitations and requirements set forth in, the CRC License.

7. **Government Employees.** If the Author is a United States government employee and the Article was prepared in the course of his or her duties as a United States government employee, as indicated in **Item 2** above, and any of the licenses or grants granted by the Author hereunder exceed the scope of the 17 U.S.C. 403, then the rights granted hereunder shall be limited to the maximum rights permitted under such statute. In such case, all provisions contained herein that are not in conflict with such statute shall remain in full force and effect, and all provisions contained herein that do so conflict shall be deemed to be amended so as to provide to JoVE the maximum rights permissible within such statute.

8. **Protection of the Work.** The Author(s) authorize JoVE to take steps in the Author(s) name and on their behalf if JoVE believes some third party could be infringing or might infringe the copyright of either the Author's Article and/or Video.

9. **Likeness, Privacy, Personality.** The Author hereby grants JoVE the right to use the Author's name, voice, likeness, picture, photograph, image, biography and performance in any way, commercial or otherwise, in connection with the Materials and the sale, promotion and distribution thereof. The Author hereby waives any and all rights he or she may have, relating to his or her appearance in the Video or otherwise relating to the Materials, under all applicable privacy, likeness, personality or similar laws.

10. **Author Warranties.** The Author represents and warrants that the Article is original, that it has not been published, that the copyright interest is owned by the Author (or, if more than one author is listed at the beginning of this Agreement, by such authors collectively) and has not been assigned, licensed, or otherwise transferred to any other party. The Author represents and warrants that the author(s) listed at the top of this Agreement are the only authors of the Materials. If more than one author is listed at the top of this Agreement and if any such author has not entered into a separate Article and Video License Agreement with JoVE relating to the Materials, the Author represents and warrants that the Author has been authorized by each of the other such authors to execute this Agreement on his or her behalf and to bind him or her with respect to the terms of this Agreement as if each of them had been a party hereto as an Author. The Author warrants that the use, reproduction, distribution, public or private performance or display, and/or modification of all or any portion of the Materials does not and will not violate, infringe and/or misappropriate the patent, trademark, intellectual property or other rights of any third party. The Author represents and warrants that it has and will continue to comply with all government, institutional and other regulations, including, without limitation all institutional, laboratory, hospital, ethical, human and animal treatment, privacy, and all other rules, regulations, laws, procedures or guidelines, applicable to the Materials, and that all research involving human and animal subjects has been approved by the Author's relevant institutional review board.

11. **JoVE Discretion.** If the Author requests the assistance of JoVE in producing the Video in the Author's facility, the Author shall ensure that the presence of JoVE employees, agents or independent contractors is in accordance with the relevant regulations of the Author's institution. If more than one author is listed at the beginning of this Agreement, JoVE may, in its sole

jove JOURNAL OF VISUALIZED EXPERIMENTS
1 Alewife Center #200
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT

discretion, elect not take any action with respect to the Article until such time as it has received complete, executed Article and Video License Agreements from each such author. JoVE reserves the right, in its absolute and sole discretion and without giving any reason therefore, to accept or decline any work submitted to JoVE. JoVE and its employees, agents and independent contractors shall have full, unfettered access to the facilities of the Author or of the Author's institution as necessary to make the Video, whether actually published or not. JoVE has sole discretion as to the method of making and publishing the Materials, including, without limitation, to all decisions regarding editing, lighting, filming, timing of publication, if any, length, quality, content and the like.

12. **Indemnification.** The Author agrees to indemnify JoVE and/or its successors and assigns from and against any and all claims, costs, and expenses, including attorney's fees, arising out of any breach of any warranty or other representations contained herein. The Author further agrees to indemnify and hold harmless JoVE from and against any and all claims, costs, and expenses, including attorney's fees, resulting from the breach by the Author of any representation or warranty contained herein or from allegations or instances of violation of intellectual property rights, damage to the Author's or the Author's institution's facilities, fraud, libel, defamation, research, equipment, experiments, property damage, personal injury, violations of institutional, laboratory, hospital, ethical, human and animal treatment, privacy or other rules, regulations, laws, procedures or guidelines, liabilities and other losses or damages related in any way to the submission of work to JoVE, making of videos by JoVE, or publication in JoVE or elsewhere by JoVE. The Author shall be responsible for, and shall hold JoVE harmless from, damages caused by lack of sterilization, lack of cleanliness or by contamination due to the making of a video by JoVE its employees, agents or independent contractors. All sterilization, cleanliness or decontamination procedures shall be solely the responsibility of the Author and shall be undertaken at the Author's expense. All indemnifications provided herein shall include JoVE's attorney's fees and costs related to said losses or damages. Such indemnification and holding harmless shall include such losses or damages incurred by, or in connection with, acts or omissions of JoVE, its employees, agents or independent contractors.

13. **Fees.** To cover the cost incurred for publication, JoVE must receive payment before production and publication of the Materials. Payment is due in 21 days of invoice. Should the Materials not be published due to an editorial or production decision, these funds will be returned to the Author. Withdrawal by the Author of any submitted Materials after final peer review approval will result in a US$1,200 fee to cover pre-production expenses incurred by JoVE. If payment is not received by the completion of filming, production and publication of the Materials will be suspended until payment is received.

14. **Transfer, Governing Law.** This Agreement may be assigned by JoVE and shall inure to the benefits of any of JoVE's successors and assignees. This Agreement shall be governed and construed by the internal laws of the Commonwealth of Massachusetts without giving effect to any conflict of law provision thereunder. This Agreement may be executed in counterparts, each of which shall be deemed an original, but all of which together shall be deemed to me one and the same agreement. A signed copy of this Agreement delivered by facsimile, e-mail or other means of electronic transmission shall be deemed to have the same legal effect as delivery of an original signed copy of this Agreement.

A signed copy of this document must be sent with all new submissions. Only one Agreement is required per submission.

CORRESPONDING AUTHOR

Name:	Idalis Villanueva
Department:	Engineering Education, College of Engineering
Institution:	Utah State University
Title:	Assistant Professor of Engineering Education
Signature:	Idalis Villanueva
Date:	March 20, 2019

Please submit a **signed** and **dated** copy of this license by one of the following three methods:

1. Upload an electronic version on the JoVE submission site
2. Fax the document to +1.866.381.2236
3. Mail the document to JoVE / Attn: JoVE Editorial / 1 Alewife Center #200 / Cambridge, MA 02140

Editorial comments:

General:

1. Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammar issues.
2. Please ensure that the manuscript is formatted according to JoVE guidelines–letter (8.5" x 11") page size, 1-inch margins, 12 pt Calibri font throughout, all text aligned to the left margin, single spacing within paragraphs, and spaces between all paragraphs and protocol steps/substeps.
3. Please avoid the use of personal pronouns (you, your, we), in particular within the protocol.

Protocol:

1. There is a 10 page limit for the Protocol, but there is a 2.75 page limit for filmable content. Please highlight 2.75 pages or less of the Protocol (including headers and spacing) that identifies the essential steps of the protocol for the video, i.e., the steps that should be visualized to tell the most cohesive story of the Protocol. Remember that non-highlighted Protocol steps will remain in the manuscript, and therefore will still be available to the reader.
2. Please ensure that all text in the protocol section is written in the imperative tense. The actions should be described in the imperative tense in complete sentences wherever possible. Notes should be concise and used sparingly.
3. For each protocol step, please ensure you answer the "how" question, i.e., how is the step performed? Alternatively, add references to published material specifying how to perform the protocol action. If revisions cause a step to have more than 2-3 actions and 4 sentences per step, please split into separate steps or substeps.

Figures:

1. Please remove 'Figure 1.' etc. from the Figures.
2. Figures 4, 5: Is anything being marked as statistically significant here? Why is it mentioned in the legends?
3. Figure 6: Please explicitly explain the asterisks in the legend.

Discussion:

1. As we are a methods journal, please revise the Discussion to explicitly cover the following in detail in 3–6 paragraphs with citations:

a) Critical steps within the protocol
b) Any modifications and troubleshooting of the technique
c) Any limitations of the technique
d) The significance with respect to existing methods
e) Any future applications of the technique

References:

1. Please ensure references have a consistent format.

Table of Materials:

1. Please ensure the Table of Materials has information on all materials and equipment used, especially those mentioned in the Protocol.

Response to Editor Comments:

Thank you for the input. We have highlighted what we deem are important steps in this protocol. In coordination with the videographer, if there are items that require additional highlighting of steps, we are happy to discuss further for the script development. All "Figure" labels were removed and asterisks were explained in all of the Figure Captions. The discussion was revised according to the editor recommendations. References were verified once more for formatting. Table of Materials were verified for consistency to the Protocol.

Reviewer #1:
Manuscript Summary:
an important and novel technique

Major Concerns:
introduction is long
discussion is very short and lack comparison with other studies
lack conclusion

Minor Concerns:
figure 6 need title

Response to Reviewer #1 Comments:

Thank you for the input. We have reduced the introduction and followed the Editor's recommendations to the Discussion, primarily focusing on the protocol as suggested by the editor. According to the JOVE manuscript submission guidelines, a Conclusion section is not required. However, we did include a discussion about the protocol according to the editor's recommendations. We include all figure captions and legends according to the JOVE guidelines.

Reviewer #2:
Manuscript Summary:
The authors could rethink the keywords (near-real-time does not seem to be a common keyword; ecological valid is questionable (see below)) and should be more careful with using strong expressions like "objective measures of emotional experiences"

Response to Comment:

We have eliminated the "near real-time" as the multi-model keyword captures the nature of the study. We have also eliminated "ecologically valid" and added "performance". Although, as

we argue below, the study is an authentic testing environment, and is far closer to the real world test than prior work. However, the descriptor may not fit as a keyword.

Major Concerns:
The experimental setup does not really represent an ecological valid situation; it is neither an authentic exam nor does it take place in an actual classroom.The experimental setup even seems to be intrusive and/or artificial with experimenters walking around collecting salivary samples, probably while having a loud freezer in the same room, advising participants not to move too much. I can see that some of these concerns are unavoidable when collecting physiological data - but I do not understand why the authors repeatedly state that they use an ecological valid situation, collecting objective data; this is simply not the case.

Response to Comment:

Based on the reviewers concerns we realize we needed to clarify aspects of the testing environment. See section 3.1 and particularly lines 289-291.

Prior work utilizing multi-model data to understand student learning have been typically conducted utilizing assessment and instructional tools, which are disconnected from students' area of study or coursework. The most common use is in the area of computerized learning settings. In our laboratory design and setting, the test is a match (in content, difficulty, and length) to a midterm exam the students are currently studying for. The test is an authentic exam. The students' performance does not determine their grade, and is therefore less anxiety producing then an actual exam. In addition to matching the practice test to the midterm, we also made an effort to match the exam experience. The experiment does not take place in a classroom, but neither does the exam that students take the following week. As is the case at many universities, the students take the test in a testing center. Our testing room is similar to the testing center. The testing center has proctors (as our setting does). See line 109-112, lines 216-219, section 3.1 (lines 273-291) and lines 490-493.

To respond to a prior concern, we kept the saliva samples in a cooler with dry ice; as such, there is no noise from a freezer. See lines 189-190, lines 268-269, and the Table of Materials descriptions.

We agree that the lack of movement required by the participants does make our testing environment less like a testing situation. We have altered our language to discuss the "increased ecological validity and authenticity" of the testing situation. See section 3 (lines 271-342).

Major Concerns:
The second major concern pertains to the theoretical section: The authors refer to research on emotions and especially research on test anxiety. The authors disregard more recent research that has used physiological data and/or real-time data (experience-sampling studies in actual exam situations) on emotions; their statements about the shortcomings of these studies are too strong and they could make their point stronger, why it is necessary to also collect physiological data (e.g. by referring to the theory that emotions consist of multiple components, including a physiological one).

Response to Comments:

Thank you for the input, particularly around making our statement of the use of physiology stronger. We have included more contexts in the introduction sections to this end, particularly in lines 49-66, 72-78, 80-81, and 98-100.

We also included more information about experience-sampling studies and can be found in lines 25-27 and 73-76. However, we argue that experience-sampling is a form of survey research and has the same limitations as prior studies. It is also not real-time in that participants cannot be actually taking an exam while simultaneously completing an experience-sampling survey. We have added more clarification on the value of our protocol in the introduction section (lines 73-76; 80-81; 98-100) and in the discussion section (lines 502-506).

Concerns:
Secondly, why don't they use self-reports on emotional experiences but only on self-efficacy? That doesn't really make sense to me. Self-efficacy is only briefly introduced and the merit of integrating physiological data is really vague (l.64-65)
The whole section (Introduction) needs to be restructured and rewritten in my point of view. For instance, in l.48 the paragraph starts with the statement that emotions are important - then the authors state that previous studies have not used appropriate methodology to assess emotions. The next paragraph (l.58) again starts with the importance of emotions (as they have consequences for learning behavior and so on) and then again refers to the methodological shortcomings of previous studies.

Response to Comments:

Our study does use self-reports on emotional experiences based on the AEQ, a validated instrument developed by Pekrun & colleagues (2014) to measure academic emotions on test related tasks. These were asked at the beginning and the end of the session. We have clarified this in our survey section. Self-efficacy is a known mediator of emotions (Bandura, 1997). Since emotions are multi-component (Pekrun & Linnenbrink-Garcia, 2014) and includes affective as well as physiological processes, including EDA alongside students' self-efficacy may allow us to better understand the associations between self-efficacy, emotional arousal responses, and performance to individual exam problems. We have clarified this in the introduction and methods section of the manuscript. For the purpose of this paper, which is primarily focused on the protocol, we simply showed exemplars of a sub-set of our data (see figures; particularly our design in Figure 2).

Minor Concerns:
Many typos (e.g. multi-model vs. multi-modal); sometimes missing references (e.g., definition of self-efficacy); sometimes use of different wordings (e.g. real-time vs. near-real-time); many repetitions in the protocol (e.g. avoid eating sugary food before the experiment; line 220 / 247; but also information on time stamping and calibration periods are repeated several times) - I would advise the authors to rethink the structure of the protocol and shorten it in some places.

Empatica is used as an example device at first (e.g. 367), but later on it seems to be expected that the reader uses an Empatica device (385)

Response to Comments:

Thank you for the note on grammatical mistakes. We have read through the manuscript and hope we identified and fixed all the errors in the text. We have shortened the protocol as well.

To clarify the concern provided, the protocol written in this manuscript is contextualized for use with an Empatica device based on the nature of the study. However, reviewers are welcome to use any other Bluetooth-enabled physiological sensor as has been stated in line 137 and lines 238-239.