ARTICLE
A Cross-Course Experiment with a Word Recall Task and Salivary Cortisol Measurements

Michael G. Ruscio & John J. Widholm
Department of Psychology and Program in Neuroscience, College of Charleston, Charleston, SC 29424.

https://doi.org/10.59390/XIIC7680

We describe the experimental design and procedures for a word recall task in combination with positive (i.e. pleasant) or negative (i.e. unpleasant) valance images and salivary cortisol response. The word recall task was a component of a 200-level psychological statistics and methods course. Two groups of student subjects were presented with one of two sets of 25 word-image pairs: identical words with images of positive or negative emotional valence. Salivary cortisol was collected prior to word-image pair presentation and following word recall. Cortisol was then analyzed in a 400-level advanced behavioral neuroscience laboratory course, and by student researchers (independent studies). These students learned the basic procedures of an enzyme immunoassay including aspects of quality control. Data collected across four semesters demonstrated word recall was significantly greater in subjects who viewed the positive valence word-image pairs. Salivary cortisol was not different between the groups. This paradigm generated a novel shared data set across classes appropriate for exploration and statistical analysis in each class. Conceptually, this approach provided a gateway for the discussion of the neuroendocrinology of cortisol and memory. It produced greater student investment in the experiment and outcome. Assessment data revealed significantly improved performance on a pre- versus post-quiz of central concepts in the 200-level course and to a lesser degree in the 400-level course. This approach resulted in a greater breadth and depth of topics that otherwise could not be accomplished within a single class. Here, we present guidelines for executing this experiment in the classroom with possibilities for novel variations.

Key words: cross-course, word recall, neuroendocrinology, experimental design, statistics, neuroanatomy instruction, HPA axis, ELISA, salivary hormones, cortisol, memory

Limitations of time and laboratory resources are perpetual hurdles for neuroscience curricula. Here we present a relatively straightforward, cost-effective cross-course experiment, that has the potential to be adapted to various novel manipulations, using a similar set of tools and foundational principles. The two central courses involved were an undergraduate 200-level statistics and methods course comprised of psychology majors (enrollment ~22-25 per semester), and a 400-level advanced behavioral neuroscience laboratory course, comprised of mostly fourth year students pursuing a neuroscience minor with a psychology or biology major (enrollment ~12 per semester). Additional students (enrollment ~2 per semester) enrolled in a semester-long independent study contributed to the collection of data as well.

The basis of this experiment was a word recall task in combination with images intended to convey a positive (i.e. pleasant) or negative (i.e. unpleasant) emotional valence paired with measurements of a salivary cortisol response. The word recall task was conducted as a component of the 200-level course. Two groups of student subjects were presented with one of two sets of 25 word-image pairs. Salivary cortisol was collected prior to image presentation and following word recall (pre vs. post). It was then analyzed as a component of the 400-level laboratory, and by independent study students. The data collected was then shared among all courses for statistical analyses.

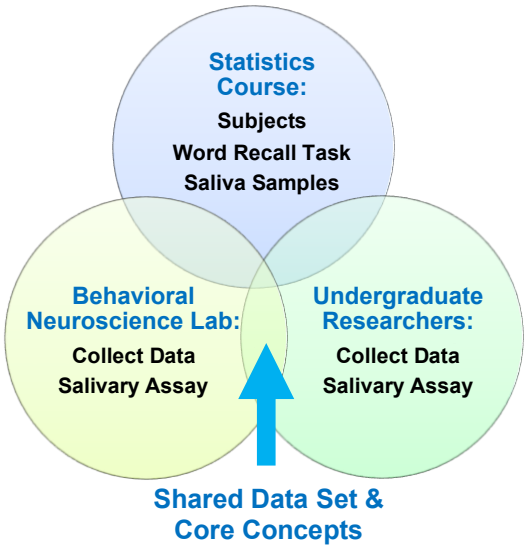

Figure 1. Cross-course design. Content and contributions from each course.

We chose to use salivary cortisol as a dependent measure for several reasons, including the fact that it is a widely used valid and reliable assay. Cortisol produced

from the adrenal gland can be readily detected in saliva at baseline levels. More pertinent for experimentation, is that changes in salivary cortisol (pre to post manipulation) can be used as an index of reaction to a psychological or cognitive stressor, as well as a physiological stressor. The time course for these types of responses is well established (10-30 minutes). Although correlating salivary cortisol response with experienced subjective stress or psychological state is a nuanced issue, numerous manipulations have demonstrated the salivary assay as an informative metric. The literature in this field is extensive, but there are several valuable meta-analyses of experimental manipulations and evaluations of the technique (e.g. Kirshbaum et al., 1993; Michaud et al., 2008; Pulopulos et al., 2020; Stoffel et al., 2021). Beyond the primary literature, there are also classic sources for a general scientific audience that effectively describe the behavioral neuroendocrinology of the stress response (e.g. Sapolsky, 2004)

Cross-course designs and assignments have been used effectively in various contexts and generally show positive learning outcomes. These designs have been used to enhance writing skills, as well as laboratory skills within and across various STEM disciplines (Keeler & Gotwals, 2021; Kjonass et al., 2017; Stiemsma, et al., 2020) including neuroscience (Branco & Chan, 2023). Beyond enhancing core concepts and laboratory skills there is the pragmatic advantage that cross-course collaboration facilitates access to limited resources such as lab equipment and materials. Our aim in the current cross-course design was to employ these same advantages.

The intention of our design was to provide specific overlap between the classes related to the experiment, while simultaneously providing enough flexibility as to not encroach on the independent learning outcomes. Therefore, within each class, instructors tailored their specific course content to attributes most fitting for their respective classes. For example, in the advanced lab course, this experiment was preceded by a module on the comparative neuroanatomy of the hippocampus (Grisham et al., 2018). This allowed students to develop a working knowledge of the neuroanatomy of the hippocampus and its well-established role in memory functions. There was related content regarding the interaction between limbic and forebrain structures during social behavior and stress related responses. The content of this preceding module set the stage for a more detailed discussion of neuroendocrinology of the cortisol response when the current experiment was undertaken later in the semester. In the statistics and methods course, this exercise was paired with discussions of the cognitive processes of memory encoding, storage and recall. The independent study students were simultaneously designing a related experiment intended to become a component of a senior thesis project. This experiment gave them experience with

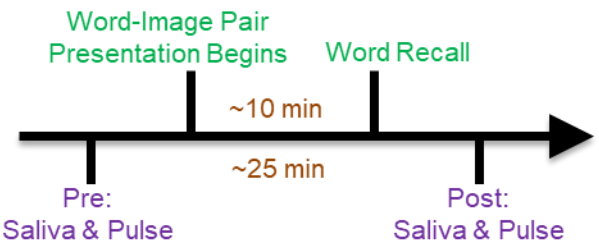

Figure 2. Procedural Timeline.

data collection from subjects and the administration of cognitive tests such as our memory recall task.

MATERIALS AND METHODS

Subjects

Subjects were 88 undergraduate students (85 females and 3 males) enrolled in a 200-level statistics and methods course at the College of Charleston during the Fall 2022, Spring 2023, Fall 2023 or Spring 2024 semesters. Approximately 2 weeks prior to experimentation, each class was told there would be a voluntary in-class experiment which would require them to take a cognitive task and provide a saliva sample. Participation was optional and not associated with any grade. No students declined to participate. Subjects were randomly assigned (using block randomization) to the positive (i.e. pleasant) or negative (i.e. unpleasant) image valance condition but were blind to their actual condition (e.g. assigned to group 1 or 2). The first group was instructed to arrive at the classroom during the regularly scheduled start of the class period (10:00 AM) to begin the experiment. The second group was instructed to arrive 60 min later (11:00 AM). Group order was randomly determined each semester, the time of day was the same each semester. *Figure 2* provides the sequence of the procedures.

General Instructions Provided to Subjects

Student experimenters (from the 400-level advanced behavioral neuroscience lab course, or 400-level independent studies) provided the initial instructions to the student subjects (from the 200-level statistics and methods course). Upon arrival, student subjects were given an instruction-response sheet. The student experimenters were tasked with creating and writing this instruction-response sheet. Student subjects were coded and identified by assigned subject numbers on the top of each sheet. Anonymity was preserved and student experimenters were blinded to subject names. No personally identifying information was kept beyond the master coding list retained by the course instructors.

The instruction-response sheet informed the subjects that pulse and saliva measures would be taken. Pre-test pulse readings (bpm) were taken using commercially available fingertip pulse oximeters (Aleshon # FPO-0218). Pulse oximeters provide a low cost, accurate and

potentially corroborative measure of changes in cortisol. Note that acrylic nails interfere with readings on fingertip pulse oximeters, so for those students with acrylics, they were instructed to measure pulse rate manually.

Subjects were required to confirm ("yes or no") their adherence to the pre-test conditions of "refraining from drinking coffee, smoking, vaping, chewing gum, brushing teeth, using mouthwash, drinking alcohol, or exercising one hour prior to the experiment." These pre-test conditions were specified by the instructor in the prior class meeting. Subjects were also asked to acknowledge ("yes or no") if they were currently taking any allergy medication or medication that they believed might affect their heart rate ("yes or no").

Prior to collecting the "pre" samples, subjects were then shown a short video with instructions on how to collect their own salivary hormones (collection details below). Subjects were then asked to rank their current level of stress at the beginning of the experiment, compared to their typical level of stress from 1 (lower stress) to 7 (higher stress). This question was asked again following the word recall task.

Word Recall Task

The order of presentation (positive vs negative images starting at 10:00 or 11:00 am) was randomly determined each semester. Each group saw 25 word-image pairs. The order in which the words were presented was identical across groups. Words were generated with an online random word generator specifying two syllable nouns. In the positive (i.e. pleasant) condition, images were selected to include human-human interaction. In the negative (i.e. unpleasant) condition, images were selected to elicit disgust (e.g. *Figure 3*). There was no purposeful linkage between the words and images and no evident association that might enhance recall. Word-images pairs on power point slides were projected onto a large screen and shown to all subjects within each group while seated in the classroom. Upon initiation of each slide, the image was presented alone for 3 s followed by the presentation of the word (alongside the image) for 4 s. After 7 s, the screen went blank (white) for 2 s and then next slide was initiated. This cycle continued for all 25 word-image pairs for approximately 3 min, 45 s. At that point a 5 s blank screen was presented (light blue) and the presentation process was repeated. Total time for the presentation was 7.5 min. At the end of the presentation there was a 1 min pause, then the subjects were given 2 min to recall as many of the words they could by writing them down on the sheet provided on the instructions-response sheet.

Figure 3. Example of word-image pair presentation with positive (i.e. pleasant, left) or negative (i.e. unpleasant, right) emotional valence.

Saliva Collection

Saliva samples were collected prior to the presentation of the word-recall task (pre) and again following (post). The duration between the samples was approximately 25 min. The duration of the word recall task (presentation of image and timed recall) was approximately 8-10 min. The time from recall to the post sample was approximately 2-4 min.

Student subjects were provided with saliva collection kits which contained 4 sterile cotton balls, and two 1.5 ml Eppendorf tubes. Subjects were instructed verbally and with a video created by the student experimenters (1 min 11 s) on how to collect saliva. Subjects inserted a single cotton ball in their mouths and were told to move it around gently for 3 min. Chewing vigorously or keeping the cotton ball stationary between the cheek and gum is not effective. Subjects were then instructed to remove the cotton ball and squeeze its contents into the Eppendorf tube marked "pre". Different color Eppendorf tubes were used (blue vs orange) to distinguish pre vs post samples. Almost all subjects produced enough saliva for a cortisol assay (approximately 50 µl per sample is necessary). As a rough index, experimenters noted subjects who filled up to the 1 ml mark on the Eppendorf tubes, which was most subjects. This process was repeated for the "post" sample. All saliva samples were placed in a laboratory freezer (-20°C) immediately following collection.

It should be noted that this method of collection works effectively for salivary cortisol but may be less effective for other salivary hormones (e.g. oxytocin) or for clinically reliable data. There are alternative and more efficacious methods for collecting saliva through passive drool techniques. Kits can be purchased (e.g. Salimetrics # 5016.04), but this increases the cost and for our pedagogical focus this was not necessary. For more elaborate designs or if multiple or more sensitive assays are planned, passive drool kits should be seriously considered by experimenters.

Salivary Cortisol Assay:

We used Salimetrics Salivary Cortisol Assay Kit (#1-3002). We have found this kit to be reliable and affordable. Our procedure directly followed the instructions in the kit (https://salimetrics.com/assay-kit/salivary-cortisol-elisa-kit/). Preparing and running the assay takes several hours, so it does not neatly fit into single class or laboratory period. For this reason, student experimenters from the lab class were assigned a time to participate in the assay. Independent study students generally had more availability as to when they could assist. Although running the assay is very straightforward, at all times students had instructor supervision.

Samples were removed from the freezer and thawed. It is critical that the samples are completely thawed and centrifuged (Salimetrics kit recommendations at 1500 x g for 15 minutes) to ensure pipetting from the samples contains minimal particulate matter and mucins. Multiple freeze thaw cycles for the samples should be avoided. If there are intentions to run multiple assays from the same samples, separate aliquots of the samples prior to freezing are recommended. The assay kit was removed from the refrigerator to come to room temperature before use.

Each Salimetrics kit contains a single 96-well plate. We ran 16 subjects plus standards within a single plate run. Each subject requires 4 wells: pre and post samples, each sample is run in duplicate. An equal number of subjects from each group were run each semester. 64 total samples were run across all 4 semesters. Therefore, saliva samples were not run from every single subject. This was primarily a cost limitation (otherwise requiring multiple assay kits each semester). Selection criteria for salivary samples included subjects who produced enough saliva (1ml) in their pre and post samples and who clearly completed all responses on the instruction-response sheet. Subjects who did not meet the pre-screening criteria (refraining from eating, etc. 1 hr. prior) were not selected. Samples from the 3 male subjects were not included in our salivary cortisol analysis. There are sex differences in salivary cortisol, and we did not have enough males within each run (1 or less) to draw any conclusions about variance or significant differences. After applying these exclusion criteria samples were randomly selected from each group.

When running the assay, proper pipetting technique is essential. For many students this was the first time using a pipettor. Students were given directions on pipetting technique by the course instructor and practiced using water into weigh boats prior to actual samples. In most cases the instructor pipetted the standards included with the kit. The accuracy of this portion of the assay is critical to creating the standard curve by which the unknown subject samples will be compared. Student experimenters pipetted the majority of the subject samples. Below (results section) we note the differences in intra-assay variability found with the standards compared to samples. Single pipettors were used for standards and samples. A multichannel pipettor and reservoirs were used for the remainder of the assay procedures. It would be exponentially more laborious and decrease accuracy if single pipettors were used for this portion. A reliable multichannel pipettor is essential.

Upon completion of the assay, we used a Bio Tek ELx 800 plate reader. Samples were read at 450 nm. A plate reader is the largest investment for this procedure. However, it is a fairly standard piece of lab equipment available on many campuses often used in biology or neuroscience labs. For the purposes of this assay most readers that work effectively at the 450 nm range would suffice. The plate reader was used in tandem with a Bio Tek Gen 5 microplate reader software on a MS Windows operating system (note: Gen 6 software is currently available). Software that can create a standard curve and convert the 'read' values (density) to concentration (µg/dL) is essential for this laboratory exercise context (and for most research applications). Additionally, the software calculates the coefficient of variability (CV) for each sample based upon the duplicate reads. Settings within the software provide options for formatting the output and the details of the standard curve values and calculations. Manually calculating (e.g. via Excel) the standard curve to convert the 'read' values to concentrations could be a useful exercise for students, but it would be laborious and was beyond our goals of this exercise. However, student experimenters were instructed regarding the interpretation of the standard curve, and how a competitive binding assay works.

Statistics

The output generated from the software was entered into IBM SPSS Statistics v27 or v28 (relative to the most recent version available on our campus each semester). For the advanced lab class this provided an important exercise in creating a data file. Data from the word recall task, and instruction-response forms were entered by the independent study students or advanced lab students depending upon the semester. Generally, with these data there are straightforward t-tests such as comparing positive and negative groups. Likewise, the data lend themselves to evident correlations (e.g. cortisol with words recalled).

For the salivary cortisol samples, it is important to note that for most comparisons we used the percentage change from pre to post samples. This is a well-established metric for physical or psychological manipulations to assess individual cortisol response (Pulopulos et al., 2020; Stoffel et al., 2021). The generally wide range in individual variability makes taking samples at a single time-point an ineffective method for determining changes that result from a particular manipulation. Throughout this experiment we found that a common misconception among students is inferring cognitive stress levels from single or baseline measure of cortisol. This experiment therefore provides a tangible way to make this critical point of understanding cortisol response relative to a single or baseline measure.

In the advanced lab class, we also had discussions regarding variability in individual samples run in duplicate and when a sample value might be considered an outlier ($\pm$3 SD units of the mean). The results presented below include our combined data across all four semesters and are representative of the statistical analyses performed within each semester (i.e. Data were combined progressively across each semester). The results are not intended to be an exhaustive analysis of our data, but rather analyses we found most straightforward and useful in our instruction.

RESULTS

Intra- and Inter-Assay Coefficients of Variability

In the advanced lab class, we discussed the importance of intra- and inter-assay variability. Inter-assay variability (plate to plate variability) was calculated per the guidelines as described in the Salimetrics kit. Inter-assay variability was low across our four separate runs of the samples (1 plate run each semester) at 1.32%.

Intra-assay coefficients of variability (variance between duplicates of the standards and samples within the same run) were low, but also a critical variable to consider for future experiments. They were 15.6%, 6.17%, 10.25% and 15.54% in chronological order across semesters. Generally, intra-assay CVs of less than 10% are considered acceptable. However, considering that students with minimal pipetting experience were able to achieve these levels speaks to the robust and forgiving nature of the assay. Importantly, it should be noted that the range for the intra-assay CV among the kit standards (used for generating the standard curve) was considerably lower: 4.69%, 2.13%, 2.28% and 8.25% in chronological order across the semesters. Throughout our study we provided instructor oversight for the standards portion of the assay and in most cases the instructor pipetted all the standards as the "examples" of how to pipette. If care is not taken at these stages, the data may not be accurate or usable in an instructive way. Creating low intra-assay variability among the standards and contrasting it with the greater variability among the samples, can be an instructive way to broach the concepts of technique and repeatability.

T-test Comparisons

Given the nature of our design the most evident and accessible statistical comparisons for students were *t*-tests. The number of words recalled in positive valance group was significantly greater than negative valance group t (86) = 2.14, $p < 0.05$ (*Figure 4*). The mean difference was 1.794 with a 95% CI [0.133, 3.455]. This includes all viable data across all semesters. For implementation of experimental paradigm, instructors should consider that when our data are broken down the same significant difference exists in 3 of the 4 semesters in which the experiment was conducted, however the direction of the difference is the same in all 4 semesters. So, the significance demonstrated here may be difficult to replicate with smaller classes (i.e. sample sizes).

The percentage change in cortisol (pre to post) was not significantly different between groups t (61) =0.341, p =0.735 (*Figure 5*). The mean difference was -2.830 with a 95% CI [-19.450, 13.79]. As noted above (methods) our subject number is different for cortisol comparisons relative to the word recall comparison because of the limitations of the number of subjects that can be run in a single salivary cortisol kit**.** It should be noted that our average cortisol levels (pre M = 0.450 µg/dL, $SE \pm$ 0.038, and post M = 0.367 µg/dL, $SE \pm$ 0.303) were within the reported typical ranges per the Salimetrics instructions (Adult females 21-30, 0.272-1.348 µg/dL, AM range).

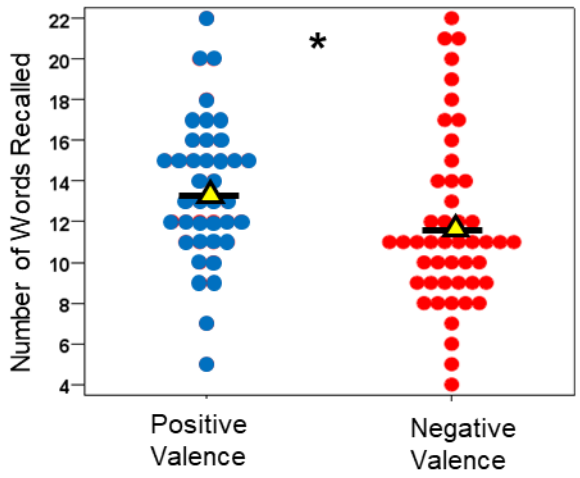

Figure 4. Subjects that viewed positive (i.e. pleasant) valence images had significantly greater word recall, $p < 0.05$. Triangles and bars indicate mean recall out of 25 words. Positive $M \pm SE$ = 13.59 ± 0.562. n=39. Negative $M \pm SE$ = 11.80 ± 0.596. n = 49.

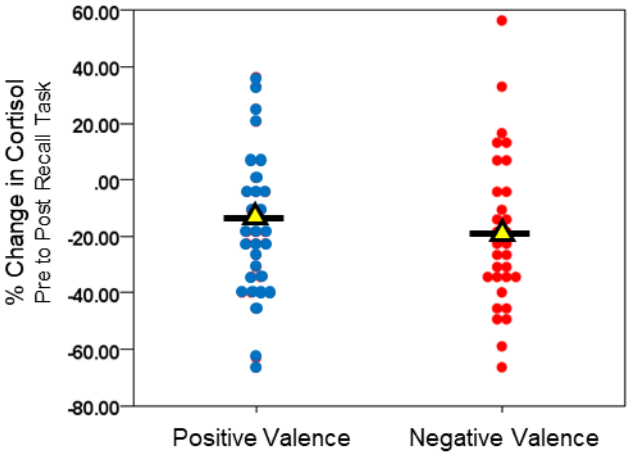

Figure 5. Percentage Change in Salivary Cortisol (pre to post word recall task). There was no significant difference between the groups. p = 0.735. Triangles indicate group means. Positive $M \pm SE$ = -16.93 ± 4.54. Negative $M \pm SE$ = -19.08 ± 4.91. n = 31 per group.

There was no significant difference between groups in the change of anxiety-stress (1-7) on Likert scale, pre to post, t (87) = 0.836, p = 0.405. The mean difference was 0.197 and a 95% CI [-0.271, 0.665]. There was no significant difference between groups in change in pulse rate, pre to post t (85) = 0.080, p = 0.936. The mean difference was 0.259 and 95% CI [-6.197, 6.717]

Correlations

Correlations are an informative but straightforward statistical test for students to grasp and provide an additional statistical method alongside the group mean comparisons described above. The most evident and

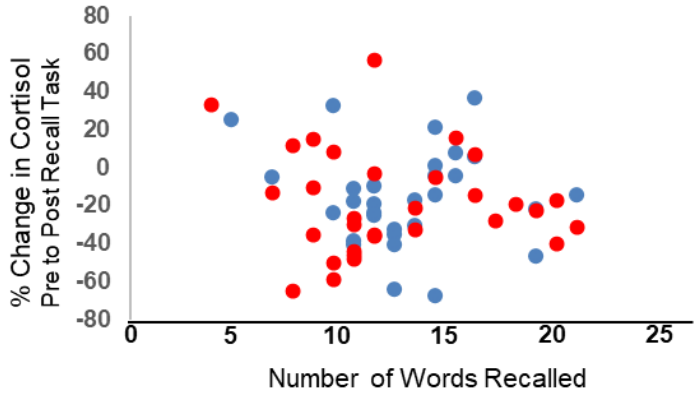

Figure 6. No significant correlation was found between words recalled and the percentage change in salivary cortisol. r (62) = -0.091, p = 0.488. Blue dots = Positive valence, Red dots = Negative valence.

relevant correlative comparison was the number of words recalled relative to the percentage change in cortisol, which was non-significant, r (62) = -0.091., p = 0.488 (*Figure 6*). This is not entirely surprising given the lack of difference in percentage change in cortisol between groups. Similarly, there was no significant correlation in the percentage cortisol relative to percentage change in pulse, r (60) = 0.047, p = 0.724. We also ran additional correlations within each group making the same comparisons, but no significance was found.

Efficacy and Assessment

Our primary intention with this manuscript is to provide an example of innovations in classroom and laboratory teaching. To that end, during select semesters, we gathered data regarding assessment of learning objectives in the 200-level statistics course and 400-level laboratory course. (Undergraduate researchers enrolled in independent studies were not included due to the small number). Assessment involved an 8-question pre and post multiple-choice quiz regarding basic statistical concepts and cortisol function. The pre quiz was administered by the end of the second week of the semester, and the post quiz was administered within the last two weeks of the semester (15-week semester). Both quizzes were administered using the online course platform D2L.

Overall performance in the 200-level course improved from pre-quiz M = 58.3% (n = 21) to post-quiz M = 84.1% (n = 15). In a paired sample t-test of the 15 students who completed both the pre- and post-assessment quizzes there was significant improvement t (14) = -5.850, p<.001. This improvement was notable on questions that reflected knowledge of cortisol function. For example, in the pre-quiz only 48% of students were able to identify the endocrine gland that produces cortisol (adrenal) compared with 93% in the post-quiz. They also showed improvement in their understanding of statistical concepts such as distinguishing a between-subjects design from a within-subjects design (52% pre-quiz versus 93% post-quiz). We attribute these findings to the fact that at the beginning of the course students were generally naïve to several statistical procedures and developed a working knowledge of them by the end of the semester as might be expected. We interpret their increased understanding of cortisol function as a positive indication of the goals of our cross-course approach. The students in this class were debriefed following their role as subjects regarding the design and hypotheses of the experiment. Additionally, the professor teaching the 400-level class provided a 60-minute guest presentation with discussion to the 200-level class. Therefore, these relatively limited interactions between the courses provided notable additional knowledge to the students. Understanding the basics of neuroendocrinology and cortisol function will ideally be a useful tool through various avenues of interest these students might explore in psychology or neuroscience. This is knowledge that would otherwise not typically be included in their 200-level course.

By comparison, we did not find significant differences in the pre- and post-quiz scores in the 400-level class although there was improvement (pre-quiz M =78.5%; post-quiz M = 88.1%). One proximate explanation for this pattern is that the pre-test average was higher in the 400-level course (78.5%) compared to the 200-level course (58.3%) leaving less room for significant improvement. This was not entirely surprising as the 200-level statistics course is a pre-requisite for the 400-level lab. Additionally, students in the 400-level lab course would have enrolled in additional courses applying these statistical principles. Also, many students in the 400-level lab had previously enrolled in Hormones and Behavior (300-level elective) taught by the same instructor as the 400-level lab. Consequently, their enrollment in the 400-level lab may have been purposeful and a reflection of an interest in behavioral endocrinology. As a result, many students accurately answered general questions of cortisol function in the pre-test. In the future, we plan on making the pre- and post- assessments more targeted at the expected knowledge in the lab class. For example, students in the 400-level course would be expected to have a deeper understanding of the ELISA technique, standard curve, assay variability. We could further assess how the analysis of these results interfaces with specific statistical procedures. Additionally, we also plan to gather data on students' impressions on the development of their own understanding of the concepts and confidence with laboratory techniques.

Importantly, we do not conclude from these data that this approach was without merit or utility for our 400-level students. The interactive and realistic nature of this design appeared to provide greater engagement for the laboratory students. Pragmatically, this intertwined with the development of laboratory skills associated with running human subjects and analyzing salivary hormone assays.

DISCUSSION

Overall, our procedure generated valid results, using techniques that were applied in a consistent manner across four semesters of our classes. Our initial success with this approach in the first semester it was utilized, inspired us to continue it on a regular basis. We have also considered variations in this approach, as the manipulation itself failed to generate a significant change in cortisol. However, the lack of a difference was still informative regarding the function of cortisol and prompted consideration of how the timing of the procedure could be optimized, or how an entirely different manipulation might prove more efficacious. Importantly, the data we generated were accurate. The word-image manipulation is easily executed and quantified. The salivary cortisol assay reliably produced values within previously established parameters. These cornerstone techniques make this approach an effective learning tool.

Considerations for Future Implementations of this Paradigm

Our current manipulation did not produce a change in cortisol. To this end, there are several modifications to consider. Regardless of the manipulation, the time of day and the timing of the cortisol response need to be carefully calibrated. In our current experiment the time of day was consistent across semesters, with one group tested beginning at 10:00 AM, the other 11:00 AM (order was randomized across semesters). Having the subjects engage in the experiment during their regularly scheduled class period provided a convenient control for time of day and the well-established diurnal pattern of cortisol secretion (Edwards et al., 2001; Stoffel et al., 2021). However, future researchers should also consider using a morning-evening questionnaire (MEQ) (Bailey & Heitkemper, 1991). In our experiment we did not collect any data on the students' circadian rhythms or sleep-wake cycle. In a college age population, there is likely greater variability in this relative to starting a procedure during the late-morning hours. More precisely, relative to their wake time, some students may be near their peak daily cortisol levels, others may be hours past it. In our study using percentage change in salivary cortisol for our statistical comparisons addresses this concern to a degree. However, inclusion of these types of MEQ metrics would improve accuracy and allow for the introduction of additional statistical considerations, such as including wake time or MEQ score as a co-variate.

The timing of our pre vs post sample could have been more effective. Cortisol generally rises 10-30 minutes after the experience of emotional states (Kirshbaum et al., 1993. Michaud et al., 2008). In our experiment the post sample itself came less than 10 minutes after the last image and recall period. Future iterations should allow for more time following the onset and completion of the stimuli, with the intention of confidently collecting samples within the 10-30 min range. For novel experimental protocols it would not be unusual to collect samples at timed intervals following the response. But for a lab class procedure, this would significantly increase the investment in time and money with increased samples per subject. For a more comprehensive review on the collection of samples relative to timing and other parameters see Stoffel et al., 2021.

Relatedly, a longer acclimation period (20-30 min) prior to the subjects starting the manipulation would be ideal (as is typical for most behavioral manipulations that aim to detect a cortisol response). In our data, the majority of subjects in both groups showed a decrease in salivary cortisol pre to post (*Figures 5 & 6*). This may be in part because our procedure began only minutes after students arrived in the classroom. Cortisol may have increased simply from driving to campus, parking or walking hurriedly across campus to the classroom therefore, making a change related to our manipulation more difficult to detect. Similar considerations apply to changes in pulse rate. In the current experiment we were constrained by trying to run two groups within the regularly scheduled class period. Providing an acclimation period would have required time outside of the regular class period.

Another consideration for future iterations of this procedure is increasing the precision of heart rate though measurements of heart rate variability (HRV), rather than using a fingertip pulse oximeter. A fingertip pulse oximeter is easy and affordable. However, more compelling data exists regarding the relationship between changes in salivary cortisol and HRV. Commercially available HRV detectors, when paired with the appropriate software, can be used to provide the root mean square of successive difference (RMSSD), a more accurate index of successful emotion and stress regulation (Pulopulos et al., 2020; Stoffel et al., 2021). Multiple wearable personal health monitors exist. However, the variability in the output of these devices paired with the variability in student access to them (e.g. affordability) made them a less practical choice for our laboratory procedure.

Specific to our manipulation, the construction of our particular word-image pair stimuli could be altered. The words were generated from a random word generator. Most of our images came from an internet search for images that would potentially generate disgust or reflect a positive human-human interaction, or nature scene. As an alternative, databases with previously assessed images, such as the International Affective Picture System (IAPS), or Open Affective Standardized Image Set (OASIS; Kurdi et al., 2017) which ranks images based on valance and arousal, could provide greater responses.

Quantifying the perception of stress relative to our manipulation could be more precise. In the current experiment, we asked subjects to rate their levels of stress pre- and post-manipulation from a single question on Likert scale 1-7. This was in part the result of students' discussion and decisions as to what to include, and somewhat less central to our initial considerations of this

experiment. However, there are multiple different established and verified scales that could provide a more accurate measurement. For example, the perceived stress scale PSS (Cohen et al., 1983, Pulopulos et al., 2020) is a potential index, of baseline stress measures, but would not determine if stress increased relative to a particular manipulation, without customization of the scale. For a focus on psychological stress response, visual analogue scales (McCormack et al.,1988) could be employed. In the current experiment our questionnaire asked for a numerical response 1-7. Presenting this question (and potentially additional questions) in a manner consistent with the traditional VAS format could be considered. Similarly, the state-trait anxiety inventory (STAI) (Knowles and Olatunji, 2020) or the positive and negative affect schedule (PANAS) (Crawford and Henry, 2004) are other viable options. Using any of these scales comes with their own caveats and considerations. Some are optimized for clinical populations, whereas others are designed to detect changes across broader periods of time (e.g. week-to-week). So, instructors should consider that exploring various scales within the context of a laboratory class, could require substantial time and effort. (For a more detailed review of these options, also see Pulopulos et al., 2020).

Beyond modifying the current manipulation, instructors should consider the general ease and versatility of the salivary cortisol assay for entirely novel manipulations. There are established manipulations to elicit cortisol responses (e.g. hand in ice water). Instructors might also challenge a given class to come up with their own manipulation for detecting a cortisol response. For example, a previous advanced behavioral neuroscience laboratory course developed a protocol around float-therapy/ sensory deprivation at a local spa. (Interestingly, it decreased cortisol in subjects who ranked the experience as highly relaxing and increased cortisol in those that found it stressful).

In summation, we hope this example serves as gateway for instructors to consider cross-course collaborations and excises, using a versatile tool (salivary cortisol) that lends itself to a variety of lab class friendly manipulations. Ultimately, this can serve to increase student engagement by providing novel data to collect and analyze.

REFERENCES

Bailey SL, Heitkemper MM (1991). Morningness-eveningness and early-morning salivary cortisol levels. Biol Psychol, 32(2-3), 181-192. doi:10.1016/0301-0511(91)90009-6

Branco RC, Chan TMV (2023). Cross-course harmonized assignments in neuroscience. J Undergrad Neurosci Educ. 21, A85-A90. doi:10.59390/BOTF3129.

Cohen S, Kamarck T, Mermelstein R (1983). A global measure of perceived stress. J Health Soc Behav, 24(4), 385-396. https://www.ncbi.nlm.nih.gov/pubmed/6668417

Crawford JR, Henry JD (2004). The positive and negative affect schedule (PANAS): construct validity, measurement properties and normative data in a large non-clinical sample. Br J Clin Psychol,43(Pt3),245-265. doi:10.1348/0144665031752934

Edwards S, Evans P, Hucklebridge F, & Clow A (2001). Association between time of awakening and diurnal cortisol secretory activity. Psychoneuroendocrinology, 26(6), 613-622. doi:10.1016/s0306-4530(01)00015-4

Grisham W, Greta, S, Burre A, Tomita, W, Rostamian, D, Schottler, N, & Krull JL (2018). Using online images to teach quantitative skills via comparative neuroanatomy: Applying the directives of vision and change. J Undergrad Neurosci Educ, 16(3), A236-A243.

Keeler DT, Gotwals B (2021). Addressing nursing student understanding of health care transitions through a cross-course writing assignment. Nursing Educ Per, 42, E129-E130. doi:10.1097/01.nep.0000000000000658

Kirschbaum C, Pirke KM, Hellhammer DH (1993). The 'Trier Social Stress Test'--a tool for investigating psychobiological stress responses in a laboratory setting. Neuropsychobiology, *28*(1-2), 76-81. doi:10.1159/000119004

Kjonaas RA, Fitch RW, Noll RJ (2017). Cross-course collaboration in the undergraduate chemistry curriculum. J Chem Edu. 94 (9) 1334-1337. https://pubs.acs.org/doi/10.1021/acs.jchemed.2c00905.

Knowles KA, Olatunji BO (2020). Specificity of trait anxiety in anxiety and depression: Meta-analysis of the State-Trait Anxiety Inventory. Clin Psychol Rev, 82, 101928. doi:10.1016/j.cpr.2020.101928

Kurdi B, Lozano S., Banaji MR (2017). Introducing the open affective standardized image set (OASIS). Behav Res Methods, 49 (2). 457-470. doi:10.3758/s13428-016-0715-3

McCormack HM, Horne DJ, Sheather S (1988). Clinical applications of visual analogue scales: a critical review. Psychol Med,18(4),1007-1019. doi:10.1017/s0033291700009934

Michaud K, Matheson K, Kelly O, Anisman H (2008). Impact of stressors in a natural context on release of cortisol in healthy adult humans: a meta-analysis. Stress, 11(3), 177-197. doi:10.1080/10253890701727874

Pulopulos MM, Baeken C, De Raedt R (2020). Cortisol response to stress: The role of expectancy and anticipatory stress regulation. Horm Behav, 117, 104587. doi:10.1016/j.yhbeh.2019.104587

Stoffel M, Neubauer AB, Ditzen B (2021). How to assess and interpret everyday life salivary cortisol measures: A tutorial on practical and statistical considerations. Psychoneuroendocrin, 133, 105391. doi:10.1016/j.psyneuen.2021.105391

Sapolsky, R. (2004). *Why Zebras Don't Get Ulcers* 3rd Edition. Holt Publishing.

Stiemsma LT, Davis SD, Brewster JL (2020). Analysis of microbial water contamination, soil microbial community structure, and soil respiration in a collaborative first-year students as scholars program (SAS). Front Microbiol, 11. doi:10.3389/fmicb.2020.590035

Received August 16, 2024; revised January 9, 2025; accepted February 23, 2025.

Purchase of salivary hormone kits was made possible with College of Charleston Psychology Department R&D and the Undergraduate Research and Creative funds. Independent study student researchers Lydia Mahan, Caitlin Cody and Verina Salib provided crucial support. We are grateful for the participation of students subjects from a Psychology statistics course from Fall '22, through Spring '24 semesters and the analysis and contributions of students from an Advanced Psychology Laboratory Fall '22 and '23.

Address correspondence to: Dr. Michael Ruscio, Department of Psychology, 66 George St., Charleston, SC 29424.
e-mail: rusciom@cofc.edu