The physiologic correlates of learning in the classroom environment

Frustace, Bruno Salvatore

http://hdl.handle.net/2144/14940
Boston University
BOSTON UNIVERSITY
SCHOOL OF MEDICINE

Thesis

THE PHYSIOLOGIC CORRELATES OF LEARNING IN THE CLASSROOM ENVIRONMENT

by

BRUNO SALVATORE FRUSTACE
B.A., Swarthmore College, 2010

Submitted in partial fulfillment of the requirements for the degree of
Master of Science
2014
THE PHYSIOLOGIC CORRELATES OF LEARNING IN THE CLASSROOM ENVIRONMENT

BRUNO SALVATORE FRUSTACE

ABSTRACT

This study served to further investigate learning and memory, and to offer a potential tool to support educational interventions. More specifically, this was accomplished by an investigation of the physiologic changes in the brain that occurred while students learned medical anatomy. A group of 29 students taking the Gross Anatomy course at Boston University School of Medicine participated in the study. Testing occurred in two sessions: prior to the course and at the completion of the course. For each session, scalp EEG was recorded while participants were shown 176 anatomical terms (132 relevant to the course and 44 obscure) and asked to respond with “Can Define”, “Familiar”, or “Don’t Know”. Behavioral results indicated a positive correlation between participants’ course grades and performance on the experimental tasks. EEG results were analyzed for event-related potential (ERP) components related to two memory components: familiarity and recollection. Results had a number of indications. For Don’t Know responses, a stronger early frontal, late parietal, and late frontal effect occurred more so for terms of Session 1 compared to Session 2. For an analysis of just Session 2 data, results indicated increased activity of the early frontal, late parietal, and late frontal effects for Can Define responses only. Session 2 Can Define responses elicited a stronger early frontal ERP, occurring between 300 and 500 milliseconds yet, the most post-retrieval processing and monitoring appeared for Can
Define terms of Session 2. Ultimately, we focused on investigating two points: 1) the effect of classroom learning on memory, and 2) the examination of ERPs as a tool to guide education interventions. Specifically, ERPs would potentially indicate markers to predict whether students would retain materials long before behavioral measures indicate these results. This has potential to determine whether long-lasting or transient learning will occur; as well as the potential to support early intervention strategies for not just students, but also individuals with learning disabilities or memory impairments.
# TABLE OF CONTENTS

TITLE ......................................................................................................................i

COPYRIGHT PAGE ...............................................................................................ii

READER APPROVAL PAGE ...............................................................................iii

ABSTRACT ............................................................................................................... iv

TABLE OF CONTENTS .......................................................................................... vi

LIST OF FIGURES ................................................................................................. viii

LIST OF ABBREVIATIONS ..................................................................................... ix

INTRODUCTION ...................................................................................................... 1

METHODS ............................................................................................................. 8

Participants .......................................................................................................... 8

Stimuli and Design ............................................................................................. 8

Experimental Procedure .................................................................................... 9

ERP Procedure ................................................................................................... 11

RESULTS ............................................................................................................. 14

Behavioral Performance ................................................................................... 14

ERP Results ...................................................................................................... 15

DISCUSSION ....................................................................................................... 22

REFERENCES ..................................................................................................... 28
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ERP ROI Map</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>Response Type By Session</td>
<td>15</td>
</tr>
</tbody>
</table>
LIST OF ABBREVIATIONS

AD .................................................................................................................. Alzheimer’s Disease
AgCl ............................................................................................................. Silver Chloride
CAI ............................................................................................................. Central Anterior Inferior
CPI ............................................................................................................. Central Posterior Inferior
EEG ............................................................................................................. Electroencephalogram
ERP ............................................................................................................. Event-Related Potential
fMRI .......................................................................................................... Functional Magnetic Resonance Imaging
ITI ............................................................................................................. Inter-Trial Interval
LAI ............................................................................................................. Left Anterior Inferior
LAS ............................................................................................................. Left Anterior Superior
LPC ............................................................................................................. Late Parietal Effect
LPI ............................................................................................................. Left Posterior Inferior
LPS ............................................................................................................. Left Posterior Superior
ms ............................................................................................................. millisecond
N400 .......................................................................................................... Early Frontal Effect
NaCl ......................................................................................................... Sodium Chloride
RAI ............................................................................................................. Right Anterior Inferior
RAS ............................................................................................................. Right Anterior Superior
ROI ............................................................................................................ Region of Interest
RPI ............................................................................................................. Right Posterior Inferior
RPS ............................................................................................................. Right Posterior Superior
INTRODUCTION

A primary goal of education is to help students store and recollect information. In a medical school setting, for instance, the goal is for significant amounts of knowledge to later be applied in a clinical setting. Thousands of experiments have been performed on the topic of learning and memory to improve education (Roediger, 2013). Many of these studies focused on behavioral tools, such as specific learning techniques (Dunlosky, 2013). For example, behavioral studies have shown that trying to actively recollect (i.e., active, self-generated attempts to retrieve information from memory) directly enhances learning (Glover, 1989; Roediger & Karpicke, 2006). While this research is insightful, little research has focused on the use of physiologic markers as a tool to help improve learning. Of the research focusing on physiologic markers, fMRI studies have begun to illustrate the process of human learning (Rosen, 2009; Karuza, Emberson, & Aslin, 2014). Further, Key (2006) utilized electrophysiological recordings in an experimental setup to examine indicators of learning in adults. In this study, participants learned simple rules for how to subsequently answer a set of test questions correctly. Results indicated that amplitudes of ERPs to learned stimuli were generally more positive than brain waves to novel stimuli. Yet, to our knowledge, there is little or no research that examines physiologic markers resulting from learning in a classroom setting. Our current research focused on investigating changes in electrophysiologic neural markers as a result of learning in a formal academic setting. More specifically, we investigated a students’ ability to recollect classroom information (anatomical terms taught during a medical
gross anatomy course) and then we measured their memory performance using
electroencephalography (EEG) and event-related potentials (ERPs) related to memory.

Learning and memory are intimately related, and numerous ERP studies have
examined episodic memory in great detail. The investigation of episodic memory has
been supported from a dual-process perspective that distinguishes two processes that are
used for recognition memory: familiarity and recollection (Woodruff, Hayanna, & Rugg,
2006; Yonelinas, 2002). Yonelinas (1994) described familiarity and recollection as
independent bases for recognition, with familiarity being a continuous measure, while
recollection being a measure with a threshold-like character. Further, familiarity has
been described as a fast and acontextual sense that an item has been seen before, whereas
recollection has been described as a slower retrieval of contextual information about an
item (Mandler, 1980; Jacoby, 1991; Woodruff et al., 2006). For example, being able to
remember the outfit, name, and additional details of a cashier from the grocery store
would signify an ability to recollect the cashier. On the other hand, familiarity would
represent not being able to specify the contextual details and only having a general sense
of knowing the cashier. The use of high-density ERPs has been shown to support the
dual-process perspective (Rugg & Curran, 2007). However, these studies have examined
recognition memory through paradigms that involved both the study and the test of
stimuli under experimental conditions (e.g., Donaldson & Rugg, 1998; Holamon, Morris,
& Retzlaff, 1995; Noldy, Stelmack, & Campbell, 1990; Mäntysalo & Gaillard, 1986;
knowledge, the current study is the first to use ERPs in a recognition paradigm that
examines learning and memory from a more naturalistic perspective. More specifically, it involves terms learned directly, over the course of a semester, from a classroom setting. The goal of this thesis was to investigate the effect of long-lasting learning in a classroom setting by examining neural correlates that are traditionally associated with recognition memory. We hoped to determine whether knowledge from the classroom can be distinguished via physiologic indications of recollection, or by other memorial processes, such as familiarity or post-retrieval verification and monitoring.

ERPs are used largely due to their precise temporal resolution that allow for distinguishing familiarity, recollection, and post-retrieval verification and monitoring in recognition memory paradigms (Duarte et al., 2006; Li, Morcom, & Rugg, 2004; Woodruff et al., 2006). Researchers suggest that a neural correlate of familiarity peaks from 300-500 milliseconds (ms) at bilateral frontal electrode sites. This early frontal effect (also referred to as the “N400” or “FN400”) precedes the individuals’ recollection of information (Curran, 2000; Friedman & Johnson, 2000; Rugg et al., 1998). Previous research has shown that familiar test items elicited a larger early frontal effect than unfamiliar test items (Curran, 2000; Goldmann et al., 2003). Further, Duzel, Yonelinas, Mangun, Heinze, and Tulving (1997) asked subjects to make either a “remember” judgment if they remembered specific details of an item’s presentation at study, or a “know” judgment if they simply had a feeling of “knowing” that an item was shown at study but could not recollect details of the item’s presentation. The “remember” and “know” judgments were described to reflect the definitions of recollection and
familiarity, respectively. They found that responses associated with “know” judgments elicited a greater early frontal effect than “remember” judgments.

In addition to familiarity, the literature describes correlates of recollection as occurring maximally at left parietal regions from 500-800 ms after test stimulus onset. This late parietal effect (also referred to as the “LPC”) is greater for items correctly identified as previously studied and is relatively insensitive to alterations in the strength of familiarity (Woodruff et al., 2006). Some researchers have argued that the parietal effect indexes the amount of information retrieved (Fjell, Walhovd, & Reinvang, 2005; Vilberg, Moosavi, & Rugg, 2006). Although it remains unclear exactly what role the parietal cortex is playing in recognition memory, the parietal activity may reflect the reactivation of the stored memory representation or the actual matching of representations stored in memory with perceptual representations of the test items (Addis & McAndrews, 2006; Ally & Budson, 2007; Schnyer, Nicholls, & Verfaellie, 2005; Wagner, Shannon, Kahn, & Buckner, 2005).

Lastly, the late frontal effect (occurring between 1000-1800 ms) is thought to reflect post-retrieval monitoring and verification processes. It supports the idea that individuals hold information in working memory during an evaluation of stimuli (Achim & Lepage, 2005). It is associated with the ongoing evaluation and monitoring of the memory retrieval attempt. Investigators have reported evidence of this late frontal effect despite unsuccessful recollection, and suggested that in addition to ongoing evaluation, this activity may reflect an executive search function of the frontal lobes directing
subsequent retrieval attempts (Ally & Budson, 2007; Budson et al., 2005; Goldmann et al., 2003; Li et al., 2004).

To better understand the effect of classroom learning for anatomical terms and their later storage in memory, we devised a paradigm and compared behavioral performance and ERP memory components of students at two time points: prior to the start of a course and within two weeks after the course ended. Participants were 1st year medical students taking a 16-week long course in medical anatomy. It was expected that students would learn the content of the course successfully (as indicated by a passing course grade). Hence, behaviorally, we expected that students would be able to indicate that they can define the majority of anatomical terms on the experimental paradigm by the end of the course. This is expected because the students will be exposed to lectures, labs, and reading materials in order to increase their knowledge of medical gross anatomy through the 16 week course. For the ERP results, it was hypothesized that there would be little indication of learning during the baseline testing session (prior to the start of the course) because, at that time point, the participants would have had scant knowledge of the medical gross anatomy material. Instead, we expected that a late frontal effect would be present, as this would represent the participants’ mental search for knowledge related to the stimuli. During the second session, we expected that the early frontal effect would be apparent for terms that the participants indicate that they are familiar with. This is hypothesized because it reflects participants’ use of memorial familiarity (i.e., an acontextual and general recognition of the stimuli) of recognition memory when making memorial judgments. “Can define” responses may be indicative of recollection because
it assumes participants can specifically define the term. We expected that both the late parietal and early frontal components would be apparent, but only for terms that students indicated that they could define because 1) this response represents the memory component of recollection (i.e., a contextual and detailed recognition of the stimuli) and 2) the late parietal component is suggest to be an all or none effect that occurs in the presence of the continuous characteristics of the early frontal component (Yonelinas 2002). Lastly, we expect a late frontal effect for terms that students indicate that they are familiar with or that they do not know. This is hypothesized because it would reflect the students’ use of working memory to determine if in fact they can recognize the term.

The second goal of the current study was to examine ERPs as a potential tool to guide educational interventions. To be helpful, ERPs must show a clear benefit that would not otherwise be provided by other more cost-effective measures such as behavioral testing. The risks/benefits of ERPs as a physiologic tool in educational interventions are much like those of Alzheimer’s disease (AD). Specifically, early characterization of brain integrity in AD has the potential for improved treatment and outcomes. An understanding of neural mechanisms (by imaging or physiologic recordings) plays an important role in offering the potential to see neurological/brain changes before behavioral manifestations (Mueller et al., 2005; Rosen, 2009). Similarly, the current study offered a baseline measure for the type of memory held by students immediately after their anatomy course had finished. These results will subsequently be used in a 6-month follow-up study that will determine how much memory for the learned information has degraded over time. It is expected that, if these hypotheses are correct,
we will have a physiologic tool that can predict whether certain education interventions will engender long-lasting or transient learning. Furthermore, it opens the possibility to tailor these learning strategies to patients who may suffer from memory impairments.
METHODS

Participants

Thirty-one adults (13 males), age 20-29 years ($M$ age = 22.97 years, $SD = 2.34$ years). Thirty were 1st year Boston University medical students; one was a 1st year PhD student at Boston University School of Medicine. All participants were enrolled in the Medical Gross Anatomy course. Participants were excluded if left-handed, had non-corrected hearing or vision problems, were non-native English speakers, or had previously taken a formal Anatomy course. For behavioral analysis, data from 5 participants were omitted due to dropout (n=2) or computer error (n=3). Additionally, post-test analysis data of 3 participants were omitted because they failed to complete the post-test. For ERP analysis, 7 subjects’ data were missing from Session 1 with an additional 4 from Session 2, either due to data recording issues, subject attrition across sessions, or computer error.

Recruitment methods included two emails to all students registered for Medical Gross Anatomy. The email notifications occurred at 1 month and 1 week prior to start of experimenting. Additionally, flyers were posted around Boston University Medical Campus. Participants were reimbursed with $30 for each session.

Stimuli and Design

A total of 220 anatomical terms were used. Three-fifths of the terms were derived from the three sections of the Medical Gross Anatomy course: 44 Back & Limbs; 44
Thorax, Abdomen, and Pelvis; and 44 Head & Neck. These 132 terms were chosen from key-terms in the Gross Anatomy syllabus. Examples include: olecranon (Back & Limbs section); omentum (Thorax, Abdomen, and Pelvis section); and buccinators (Head & Neck section). The same 132 terms were used for each experimental session.

The remaining 88 terms were labeled “Obscure”. Obscure terms were derived from Fonahn, 1922 and were out-dated anatomical terms of Latin and Arabic roots. Examples include: lisan (tongue), natis (greater trochanter), and poples (popliteal fossa). These 88 terms were divided into thirds so that each of the 3 experimental sessions had 44 different obscure terms. Thus, each session consisted of 176 terms (132 relevant terms and 44 obscure terms).

**Experimental Procedure**

The experiment consisted of two test sessions. The first session occurred during the week before the Anatomy course began. The second session occurred within 12 days after the course ended ($M = 6.58$ days, $SD = 3.81$ days). The course was 16 weeks in length. Anatomical terms associated with the Back & Limbs; Thorax, Abdomen, and Pelvis; and Head & Neck were taught in the first, second, and third parts of the course, respectively. Both test sessions were identical, except that a different set of obscure terms were used for each session.

Participants were tested individually with each session lasting approximately 2 hours. Stimuli were presented on a 22 inch computer via E-Prime software, version 2.0 (Psychology Software Tools Inc.; [http://www.pst-net.com/eprime](http://www.pst-net.com/eprime)). For each session,
participants completed three separate tasks. Only the first task (recognition of terms while utilizing EEG) applies to the study discussed. The second (“Task 2”) involved recognition of images while utilizing Gaze-tracking. The third (“Task 3”) involved recognition of images while using a combination of EEG and gaze-tracking. The images used with these latter two tasks were not identical to all of the terms used in the first experiment; however some terms/images did overlap. This paper focuses on Task 1.

Participants began Task 1 immediately after EEG/ERP setup completed. In the computer task, they were presented with 176 anatomical terms and asked to press, on a keyboard, “1”, “2”, or “3” to indicate “Can Define”, “Familiar”, or “Don’t Know”, respectively. Of the items, 44 were taught in the first-third of the Gross Anatomy course, 44 were taught in the second-third of the course, and 44 were taught in the final third of the course. In addition to these 132 items, there were 44 obscure terms that were not part of the Anatomy course curriculum to serve as control items. All 176 terms were presented in random order for each session. These stimuli numbers were used to ensure adequate bins sizes (> 15) for analyses.

Each trial began with a 1000 ms fixation character (“+”) prior to stimulus presentation. Test stimuli remained on-screen until the participant responded. Following their response, an inter-trial interval (ITI) of 1500 ms was displayed. After the ITI, the next trial would begin. The task lasted approximately 20 minutes. Participants then began Tasks 2 and 3. These tasks lasted approximately 40 minutes.

Once the experiments were completed, the EEG equipment was removed and the participant began a post-test test. This test allowed for confirming the accuracy of the
“Can Define” and “Familiar” responses during the ERP computer test. The test was taken on a laptop and presented with Microsoft Office Excel. On one column, definitions were presented for all of the terms that the participant indicated that they “Can Define” or are “Familiar” with. Participants were asked to write the term that best describes the definition. Definitions were derived from one of two sources: Stedman’s medical dictionary (http://www.stedmansonline.com/) or Merriam-Webster’s medical dictionary (http://www.merriam-webster.com/medical/).

**ERP Procedure**

Subjects were seated and fitted with an Active Two electrode cap (Behavioral Brain Sciences Center, Birmingham, UK). A full array of 128 Ag-AgCl BioSemi (Amsterdam, The Netherlands) “active” electrodes were connected to the cap in a pre-configured montage which places each electrode in equidistant concentric circles (Fig. 1). Active electrodes are amplified through the electrode at the source and do not require abrading of the skin or measuring skin-electrode impedance levels. Prior to placement, electrodes were soaked in warm sodium chloride solution (NaCl) that served as a conductor for electrical currents from the scalp to the electrodes. In addition to the 128 scalp electrodes, two mini-biopotential electrodes were placed on each mastoid process. Finally, vertical and horizontal electrooculogram (EOG) activity was recorded from bipolar electrodes placed below the left eye and on the outer canthus of the left and right eye. EEG and EOG activity was amplified with a bandwidth of 0.03–30 hertz (3 decibel points) and digitized at a sampling rate of 2048 hertz. Recordings were referenced to a
vertex reference point, but were later re-referenced to a common average reference to
minimize the effects of reference-site activity and accurately estimate the scalp
topography of the measured electrical fields (Curran, DeBuse, Woroch, & Hirshman,
2006; Dien, 1998).

The sampling epoch for each test trial lasted for a total of 2000 ms, which
included a 200 ms pre-stimulus baseline period. This pre-stimulus period was used to
baseline correct averaged ERP epochs lasting 1800 ms. ERPs were averaged and
corrected using the EMSE Software Suite (Source Signal Imaging, San Diego, CA,
USA). Trials were corrected for excessive EOG activity using the empirical EMSE
Ocular Artifact Correction Tool, in which artifact data are manually identified from the
clean data by the investigator. The Ocular Artifact Tool then produces a logarithmic ratio
of artifact data versus clean data and subtracts artifact data from all channels where it is
detected. Individual channels with poor recording were corrected with the EMSE spatial
interpolation filter.
Figure 1. Positions of the 128 electrodes on the bio-semi active two headcap with the eight regions of interest shown.
RESULTS

Behavioral Performance

Behavioral data represent the total number of a single Response Type (Can Define, Familiar, Don’t Know) divided by the total number of responses (176). These data were entered into a 2 (Session: 1, 2) X 3 (Response Type: Can Define, Familiar, Don't Know) repeated measures analysis of variance (ANOVA). Post-hoc t-tests were conducted as necessary. Only significant effects are reported.

A main effect of Response Type was observed, $F(2, 50) = 198.82, p < .001$, partial-eta squared = .888. A significant Session X Response Type interaction was also observed, $F(2, 50) = 491.06, p < .001$, partial-eta squared = .952. Bonferroni corrected -tests ($\alpha = .017$) showed that Can Define responses were significantly higher in Session 2, $t(25) = 27.74, p < .001$, and Don't Know Responses were significantly higher in Session 1, $t(25) = 25.237, p < .001$. See Figure 2 for the proportion of responses by Response Type and Session.

Post-hoc Pearson bivariate correlations were conducted to examine if the ability to accurately define terms was associated with course grades. Course grades were correlated with Experimental Can Define scores, and Post-Test Scores. This yielded significant positive correlations for Session 2 Can Define Experimental scores ($r = .49, p = .01$) and Session 2 Post-Test Scores ($r = .556, p = .006$). Both these correlations suggested that the higher the scores in the experiment and in the post-test, the more likely participants would receive higher grades in the course.
Figure 2. Mean recognition percentage for Can Define, Familiar, and Don’t Know responses by Session 1 (pre-course) and Session 2 (post-course) of the experimental task. Error bars represent one standard error of the mean standard errors. Significant within-group differences: for Can Define and Don’t Know response types (p<0.05).

ERP Results

ERP analysis used ANOVAs and focused on three epochs (300-500, 500-800, and 1000-1800 ms post stimulus onset). These time intervals have been previously associated with the three components of the old/new ERP effect (early frontal, parietal, and late frontal effect, respectively) as described above (Curran et al., 2006; Curran, 2000, 2004). Mean amplitudes were calculated for these three time intervals, which were then averaged across groups of electrodes that formed ten separate regions of interest (ROIs). These ROIs are the Left Anterior Inferior (LAI), Central Anterior Inferior (CAI), Right Anterior Inferior (RAI), Left Anterior Superior (LAS), Right Anterior Superior (RAS),
Left Posterior Inferior (LPI), Central Posterior Inferior (CPI), Right Posterior Inferior (RPI), Left Posterior Superior (LPS), and Right Posterior Superior (RPS). Each ROI consisted of a seven-electrode cluster. See Fig. 1 for scalp topography of the ROIs.

Frontal ROIs were of most interest for the early frontal effect and were the only ROIs used in that analysis (LAI, CAI, RAI, LAS, RAS). More posterior ROIs were used for the Parietal Effect (LPI, CPI, RPI). For the late frontal effect, frontal ROIs were again used for the subsequent analyses. Only significant effects are reported in the subsequent analyses. Post-hoc t-tests are reported as necessary.

Session 1 & Session 2 Don’t Know Analyses:

This analysis compared early frontal (300-500 ms), parietal (500-800 ms), and late frontal effects (1000-1800 ms) for Don’t Know Responses across sessions. As participants rarely responded with Can Define and Familiar responses in Session 1, ERP waveforms could not be generated for these two types of responses. However, Session 1 Don’t Know responses served as a suitable baseline with bin/event sizes of 15 or more. In order to adequately analyze data across sessions, a maximum likelihood algorithm was used to estimate missing data and preserve the variance across sessions. Maximum likelihood method for missing data is thought to induce less statistical bias compared to other methods and improve overall statistical power.

Analyses for the early front effect (N400) Don’t Know responses across sessions were entered into a 2 (Session: 1, 2) X 5 (ROI: LAI, CAI, RAI, LAS, RAS) repeated measures ANOVA. There was a main effect ROI, $F(4, 122) = 11.81, p < .001$, partial-eta
squared = .297, indicating that anterior inferior ROIs tended to be more negative compared to anterior superior ROIs. There was also a significant Session X ROI interaction, $F(4, 112) = 16.224, p < .001$, partial-eta squared = .231. A post-hoc t-test showed that the RAI was significantly more negative in Session 1 compared to Session 2, $t(28) = 3.90, p = .001$. No other significant effects were observed.

For the late parietal component, a 2 (Session) X 3 (ROI: LPI, CPI, RPI) repeated measures ANOVA was conducted. This yielded a significant main effect of ROI, $F(2, 56) = 7.37, p = .001$, partial-eta square = .208. This indicated that the right posterior region was more positive than the central or left regions. A significant Session X ROI interaction was observed, $F(2, 56) = 14.02, p < .001$, partial-eta square = .334. Post-hoc t-tests showed that the LPS was more positive in Session 1 than Session 2, $t(28) = .035, p = .035$, and the RPS was more positive in Session 2 than Session 1, $t(28) = 3.52, p = .001$. No other significant main effects or interactions were observed.

Finally, for the late frontal effect, a 2 (Session) X 3 (ROI: LAI, CAI, RAI, LAS, RAS) repeated measures ANOVA was conducted. There was a main effect of Session, $F(1, 28) = 7.64, p = .01$, partial-eta square = .215, indicating that frontal sites were more negative in Session 1 compared to Session 2. There was a main effect of ROI, $F(4, 112) = 5.26, p = .001$, partial-eta squared = .158, indicating that central frontal sites tended to be more positive than all other sites. A significant Session X ROI interaction was observed, $F(4, 112) = 4.03, p = .004$, partial-eta square = .126. Post-hoc t-tests revealed that the CAI was more negative in Session 1 versus Session 2, $t(28) = 2.59, p = .015$ and the RAS was more negative in Session 1 versus Session 2, $t(28) = 3.33, p = .002$. 17
Session 2 Analyses:

Analyses were performed on the three selected time intervals (300-500, 500-800, 1000-1800 ms). For each, three separate ANOVAs are reported: a comparison of Can Define versus Don’t Know responses, Can Define versus Familiar responses, and Familiar versus Don’t Know responses. Data for the early frontal effect (300-500 ms) focused on frontal ROIs (LAI, CAI, RAI, LAS, RAS). Data for the late parietal effect (500-800 ms) focused on posterior parietal ROIs (LPS, CPS, RPS). Data for the late frontal effect (1000-1800 ms) focused on frontal ROIs (LAI, CAI, RAI, LAS, RAS). Only significant effects are reported, and post-hoc t-tests are reported as necessary.

Early Frontal Effect Analyses:

A 2 (Response Type: Can Define, Familiar) X 5 (ROI: LAI, CAI, RAI, LAS, RAS) repeated measures ANOVA was conducted on frontal ROIs for data between 300-500 ms of stimulus onset. This yielded a main effect of ROI, $F(4, 96) = 2.81, p = .03$, partial-eta square = .105. This indicated that right and central frontal sites (RAI and CAI) were trending to be more negative than all other ROIs. A significant Response Type X ROI interaction was observed, $F(4, 96) = 3.24, p = .015$, partial-eta squared = .119. Post-hoc t-tests revealed that the CAI was significantly more negative for Can Define responses compared to Familiar responses, $t(24) = -2.22, p = .036$.

A 2 (Response Type: Can Define, Don’t Know) X 5(ROI) repeated measure ANOVA was conducted for the 300-500 ms interval. This yielded a significant main
effect of ROI, \( F(4, 96) = 3.24, p = .015, \) partial-eta squared = .119, indicating that right and central sites (RAI and CAI) were trending to be more negative than all other ROIs. There was a significant Response Type X ROI interaction, \( F(4, 96) = 13.74, p < .001, \) partial-eta squared = .364. Post-hoc t-tests revealed that in the RAI, Can Define responses were significantly more negative than Don’t Know responses, \( t(24) = -4.38, p < .001. \) Conversely, Don’t Know responses were more positive than Can Define responses in the LAS, \( t(24) = 2.41, p = .024. \)

Finally, a 2 (Response Type: Familiar, Don’t Know) X 5(ROI) repeated measures ANOVA was conducted for the 300-500 ms interval to compare Familiar versus Don’t Know responses. No significant main effects or interactions were observed in this analysis.

Late Parietal Effect Analyses:

A 2 (Response Type: Can Define, Familiar) X 3 (ROI: LPS, CPS, RPS) repeated measures ANOVA was conducted posterior parietal ROIs for data between 500-800 ms of stimulus onset. This yielded a main effect of Response Type, \( F(1, 24) = 9.78, p = .005, \) partial-eta square = .290. This indicated that ERPs were significantly more positive for Can Define responses versus Familiar responses. A main effect of ROI was observed, \( F(2, 48) = 8.20, p = .001, \) partial-eta squared = .256. This indicated that right and central parietal electrodes (CPS and RPS) were trending to be more positive than all other ROIs. A significant Response Type X ROI interaction was observed, \( F(4, 96) = 3.24, p = .015, \) partial-eta squared = .119. Post-hoc t-tests revealed that the CAI was significantly more
negative for Can Define responses compared to Familiar responses, \( t(24) = -2.22, p = .036 \). Central and Right parietal ROIs (CPS, RPS) were significantly more positive overall.

A 2 (Response Type: Can Define, Don’t Know) X 3 (ROI) repeated measures ANOVA was conducted for the 500-800 ms interval, comparing Can Define and Don’t Know response types. This yielded a significant main effect Response Type, \( F(1, 24) = 6.34, p = .019 \), partial-eta squared = .290. This indicated that ERPs were significantly more positive for Can Define versus Don’t Know responses. A main effect of ROI was observed, \( F(2, 48) = 9.84, p < .001 \), partial-eta squared = .291. This indicated that right and central parietal electrodes (CPS and RPS) were trending to be more positive than all other ROIs. A significant Response Type X ROI interaction was observed, \( F(2, 48) = 12.01, p < .001 \), partial-eta squared = .334. Post-hoc t-tests showed that the LPS and CPS were significantly more positive for Can Define responses versus Don’t Know responses, \( t(24) = 3.47, p = .001 \) and \( t(24) = 3.15, p = .004 \) respectively.

Finally, a 2 (Response Type: Familiar, Don’t Know) X 3(ROI) repeated measures ANOVA was conducted between the 500-800 ms interval, comparing Familiar and Don’t Know responses. A significant main effect was observed for ROI, \( F(2, 48) = 13.19, p < .001 \), partial-eta squared = .355. This indicated that right and central electrode sites (CPS, RPS) were more positive than the left electrode site (LPS). No other effects were observed.

Late Frontal Effect Analyses:
A 2 (Response Type: Can Define, Familiar) X 5 (ROI: LAI, CAI, RAI, LAS, RAS) repeated measures ANOVA was conducted frontal ROIs for data between 1000-1800 ms of stimulus onset. There was a trend towards significance for a main effect of Response Type, $F(1, 24) = 3.83, p = .06$, partial-eta square = .138. Can Define responses were more negative than Familiar responses. There was a significant main effect of ROI, $F(4, 96) = 3.97, p = .005$, partial-eta squared = .142. This indicated that right and central frontal sites (RAI and CAI) were trending to be more negative than all other ROIs.

A 2 (Response Type: Can Define, Don’t Know) X 5 (ROI) repeated measure ANOVA was conducted for the 1000-1800 ms interval, comparing Can Define and Don’t Know response types. This yielded a significant main effect of ROI, $F(4, 96) = 2.70, p = .035$, partial-eta squared = .101. This indicated that right and central sites (RAI and CAI) were trending to be more negative than all other ROIs. There was a significant Response Type X ROI interaction, $F(4, 96) = 11.82, p < .001$, partial-eta squared = .330. Post-hoc t-tests revealed that in the RAI, Can Define responses were significantly more negative than Don’t Know responses, $t(24) = -4.34, p < .001$. There was a trend for Can Define responses to be more negative in the RAS compared to Don’t Know responses, $t(24) = -1.84, p = .077$.

Finally, a 2 (Response Type: Familiar, Don’t Know) X 5 (ROI) repeated measures ANOVA was conducted between the 1000-1800 ms interval, comparing Familiar and Don’t Know responses. No significant main effects or interactions were observed for this analysis.
The present study investigated the physiologic correlates of learning from a classroom setting. We examined the electrical potentials through EEG while first year medical students made one of three types of responses to anatomical terms they saw on a computer screen. They either said they “Can Define” the term, were only “Familiar” with the term, or said they “Don’t Know” the term. Data was reviewed by analyzing results for Session 1 (baseline session occurring prior to the start of the course) and Session 2 (occurring immediately after the end of the 16-week long course). Session 1 and 2 data were also compared, yet only for Don’t Know responses. Furthermore, we measured different event-related potential components associated with memory across various predefined regions of interest on the scalp. This investigation served as an initial experiment that has supplied a baseline measure to subsequently (in a future experiment) determine whether a physiologic marker can be used to determine long-lasting or transient learning.

Behaviorally, the success of learning was measured in terms of the participants’ course grades, post-test scores, and number of terms indicated as “Can Define” or “Familiar”. It was our expectation that participants would improve in these three measures by the end of the course because the classroom setting should increase the participants’ anatomical knowledge. As suggested by the data, there was a significant correlation between course grade, post-test performance, and recognition of terms. Further, as expected, obscure words were largely indicated as “Don’t Know” terms for
both Session 1 and 2. This supplies further support that participants provided accurate judgment as to whether or not they recognized terms.

Regarding ERP results of Don’t Know responses for Session 1 vs. Session 2, it was expected that there would be no significant differences for familiarity and recollection between the two sessions. This was hypothesized because a Don’t Know term should theoretically produce similar electrical activity if the participant has no prior exposure to the term. On the other hand, it was expected that a late frontal component would be present for Session 2 Don’t Know responses because participants may be utilizing their acquired classroom knowledge to perform a deeper search of whether or not the term is recognizable.

For early frontal effect data, there was a greater negativity in Session 1 compared to Session 2 Don’t Know responses. (To note: negative and positive electrical activities have no distinguishing implications.) This unexpected result may be because the participant could have had prior exposure to the terms without objectively realizing. In other words, participants may have been familiar with the terms to some extent, yet still indicated that they did not know the term because of uncertainty in their response. On the other hand, during Session 2, there were an increased number of terms that participants indicated that they Can Define, and a decreased number of terms indicated as Don’t Know. Due to that, participants may have had more confident means of distinguishing between terms they can recollect and terms they do not know. Being able to better classify more terms as Familiar and Can Define allowed for increased certainty in their Don’t Know responses.
For the late parietal effect, it was observed that one ROI (RAI) showed greater negativity in Session 1 compared to Session 2. It is unclear what this may reflect. Prior literature suggests that the late parietal effect has maximum activity on the left hemisphere, as opposed to the right hemisphere (Rugg, 2007). Due to that, we do not expect that this result is the cause of mis-categorizing a response. Further investigation would be needed to explain this.

Lastly, data for the late frontal effect suggest that participants engage in stronger controlled memory retrieval in Session 1 versus Session 2. This is contrary to what was hypothesized. These results may be due to a reason similar to that for the early frontal effect Don’t Know data mentioned above. To reiterate, it may be due to the participants’ lack of confidence and certainty when deciding on a response. In other words, participants could not immediately dismiss terms unknown because, prior to learning anatomy, they have a poor conception of terms they certainly know or certainly don’t know. It would have been expected that for judgments that participants were more certain about, their response would be rapid and not require post-retrieval processing and monitoring.

Our predictions for the Session 2 early frontal effect was that only Familiar and Can Define terms would produce the early frontal effect patterns. This suggestion was guided by prior literature (Rugg, 2007). The current data suggested that ERP data associated with memorial familiarity were more evident for Can Define compared to Familiar or Don’t Know responses. This partially aligns with our predictions, as it provides evidence for the Can Define responses in the 300-500 ms interval, but not for
the Familiar responses. These indications of Can Define responses suggest a strong sense of familiarity and, thus, the potential to serve as a marker that could reflect long lasting learning. These results give further reason to investigate a 6-month follow-up experiment that examines participants’ memory for the anatomical terms previously tested on. If there is a relationship between the brain activity for a third session and the N400 patterns of Session 2 Can Define responses, then these results may serve as a physiologic measure to predict which types of memory will withstand long lasting learning. Further bolstering this idea, no effects were observed between ERPs associated with Familiar and Don’t Know responses. It could be that participants did not learn terms categorized as Familiar or Don’t Know.

The ERP data between the 500-800 ms interval supports our prediction. We expected increased activity for the late parietal component as an effect of Can Define responses. Results show stronger effects for Can Define responses versus Familiar and Don’t Know responses. Main effects of Response Types in each ANOVA revealed greater positivity for Can Define responses, relative to Familiar and Don’t Know responses. The Can Define responses seem to strongly reflect the recollection of learned information. This could be taken as evidence that the actual recollection of terms is actually occurring, and that participants did truly learn and retain these terms over the course of the semester. The lack of major effects between Familiar and Don’t Know suggest that terms in these categories were not learned and retained very well over the course of the semester.
Lastly, a late frontal effect was expected for Familiar and Don’t Know responses, as this would indicate executive function used in the search of a more certain answer (i.e., whether the participant can, in fact, define the term). However, data for the late frontal effect are not as clear as the early frontal and late parietal effects. ERPs for Can Define responses trended as more negative for the late frontal effect compared to Familiar and Don’t Know responses. This suggests that for these responses, participants may search memory and engage verification mechanisms for terms they recall. This is in contrast to Familiar and Don’t Know responses. For these response types, terms may have not been learned very well and information is stored weakly in memory. Verification processes may be engaged but there may be much less information to search for in memory overall, leading to a weaker late frontal effect. However, if this latter hypothesis is true, then we are left to re-evaluate the hypothesis for how the late frontal effect was present for Session 1 Don’t Know responses. In Session 1, Don’t Know responses were shown to have increased late frontal activity. One possibility to explain these results is that, in Session 2, participants may have better means to distinguish between Can Define terms and Familiar/Don’t Know terms. Therefore, for terms that they could not define, they were certain of this and, thus, did not require additionally verification and monitoring. Alternatively, there could indeed still be a late frontal effect that occurs for Session 2 Familiar and Don’t Know. However, since there is a much larger amount of information for Session 2 Can Define terms, the ERP results may simply indicate that the late frontal effect is much strong for Can Define terms than for Familiar and Don’t Know terms.
In conclusion, this study has illustrated components of learning and memory from a classroom setting. This illustration accomplished our first stated goal: to understand the effects of classroom learning on recognition memory. Moreover, it takes a first step towards accomplishing our second stated goal: to examine ERPs as a potential tool to guide educational interventions. The results of this experiment suggest ERP markers that may be utilized to improve understanding of long-lasting learning. The main question that remains is: What will results be if this study is followed with a, nearly identical, third session that occurs months after the course has ended? Our next planned step for this investigation is a 6-month follow-up third session. It is our hope that results from the current study will be able to predict specific results for the follow-up third session. In this respect, this procedure may offer a tool for identifying effective learning strategies long before behavioral measures indicate the outcomes. Furthermore, it is ultimately our hope that, if these results are clear, we would be able to extend these strategies to not just students, but also to groups with learning or memory impairments as a tool to support early intervention.
REFERENCES


CURRICULUM VITAE

Bruno Salvatore Frustace

13 Stonecrest Dr. • New Windsor, NY 12553 • cell: (845)-591-0272 • bfrusta2@bu.edu

Year of birth: 1988

Education

• Swarthmore College -- *B.A in Psychology*
  Sept. 2006 – May 2010
  o Swarthmore, PA

• Boston University School of Medicine -- *M.S. candidate in Anatomy & Neurobiology*
  Sept. 2012 – May 2014
  o Boston, MA

Work & Research

• Research Technician – Boston University School of Medicine
  o Performed clinical & translational research on Alzheimer’s disease
    ▪ Administered tests to healthy adults and Alzheimer’s patients
    ▪ Interacted in a one-on-one setting with participants
  o Worked in Memory Clinic at VA Medical Center in Jamaica Plain, MA
    ▪ Administered Neuropsychological tests to patients
    ▪ Worked approximately once per month on Mondays
    ▪ Interacted in one-on-one setting with patient and family member(s)

• Biostatistics Teaching Assistant – Boston University School of Medicine
  Jan. 2013 – May 2013
  o Held weekly review sessions for graduate students in the course
  o Graded exams
• **Masters Research Thesis – Boston University School of Medicine**  
  *June 2013 – Present*

  o Project title: *Physiologic correlates of learning in a classroom environment*

    ▪ Using event-related potentials, our lab explored the physiological changes in the brains of 1st year Boston University medical students from before they begin classes to when they complete their first semester

• **Dental Anatomy Prosector – Boston University School of Medicine**  
  *Mar. 2014 – Present*

  o Dissected head & neck sections of human cadaver

  o Lectured 1st year Boston University dental students on dissections performed

**Extracurricular Activities**

• **Volunteer tutor – Chester, PA**  
  *Feb. 2007 – May 2010*

  o Tutored weekly and organized field trips at an after-school program for under-privileged children

• **Volunteer for ACEing Autism – Weymouth, MA**  
  *Jan. 2012 – Present*

  o Mentor and teach children with Autism on how to play tennis

  o Sundays during the Fall and Spring

• **Volunteer coordinator – Anatomy & Neurobiology Dept. at Boston University School of Medicine**  
  *Sept. 2012 – Present*

  o Organize and lead activities such as assisting at a local soup kitchen

• **Homeless Outreach with Mosaic Church – Boston, MA**  
  *Dec. 2012 – Present*

  o Offer food, clothing, and social support to homeless individuals in Copley Square, Boston
• Wednesday evenings

• **Toastmasters – Boston University Medical Campus**  
  *Aug. 2013 – Present*

  • Practice and hone effective speaking skills

**Publications**

• Simmons-Stern NR, Deason RG, Brandler BJ, **Frustace BS**, O'Connor MK, Ally BA, Budson, AE. Music-based memory enhancement in Alzheimer's disease: Promise and limitations. *Neuropsychologia* 2012 50, 3295-3303


**Poster Presentations**


• Veterans Health Administration Research Poster Session – West Roxbury, MA – April 23rd, 2012