

Event-Related Potentials in Humans for Emotional Words versus Pictures

Lauren Kennedy¹, Sara Dorrance¹, Tyler Stoneham¹, Matt Bryant¹, Kendra Boyd¹, Kayla Flippen¹, David F. Nichols¹

¹*Roanoke College, Salem, Virginia 24153*

Along with self-reported emotional reactions, changes in brain activity occur when someone is exposed to emotionally-charged images related to pleasure and disgust, such as baby animals and flesh wounds. Previous studies have shown that event-related potentials (ERPs) are enhanced separately by emotional pictures and emotional words, but none have yet to consider related pictures and words in the same study. This study examined the effects that stimulus type and level of emotional valence had on brain activity. Electroencephalography (EEG) was used to record P300 and late positive potential (LPP) activity of 21 undergraduate students. Two stimulus types were used (pictures and words), and three levels of emotionality were also used (positive, neutral, and negative). Analysis showed that pictures generated higher magnitude P300 and LPP peaks as compared to words for all emotional states. In the channel involving the occipital lobe, which is responsible for basic visual processing, results showed an effect for stimulus type but not for emotionality. The second channel, involving an electrode over the parietal lobe (a brain area participating in emotional processing), revealed an effect associated with positive and negative stimuli compared to neutral stimuli, with the magnitude of the effect depending on stimulus type. It was found that pictures consistently led to a greater emotional response compared to related words. Further analysis took into consideration the possibility of individual differences in self-reported arousal in the LPP time range, though no significant effects based on individual differences were found on the magnitude of the LPP. The continued presence of main effects of stimulus type, however, was indicative of differential neural processing of pictures versus words. Overall the results indicated larger magnitude changes in brain activity when pictures were used as visual stimuli compared to related words, both in regards to basic visual processing and emotional processing.

Abbreviations: EEG – electroencephalography; ERP – event-related potential; LPP – late positive potential 700-900 ms after stimulus presentation; P300 – positive peak 340-500 ms after stimulus presentation

Keywords: EEG, ERP, P300, LPP, Valence, Arousal

Introduction

Over the course of biological evolution, the human race has developed and refined the most integral component of explicit communication: language. Using words to represent and symbolize specific written or spoken referents in the environment around us has become a reliable and long-standing way to

express ideas. Due to the importance and ubiquity of language in our everyday lives, which began in the form of spoken word and evolved into written word, it would be reasonable to assume that the human brain has structures devoted to the processing of words, much like other areas' roles in processing

frequently encountered stimuli (e.g., the fusiform face area's primary role in facial processing; Kanwisher et al., 1997). One such area that is known to play a primary role in reading ability and response to visual words, the Visual Word Form Area, is located in the left occipitotemporal sulcus, which borders along the fusiform gyrus (McCandliss et al., 2003). But with the even older tradition of graphically representing the world of naturally occurring visual images through manmade depictions (e.g., cave paintings), the question arises as to how the neural processing of these two forms of expression may differ on a physiological level. Is the magnitude comparable for the activation of brain areas involved in processing words and pictures?

With modern technology, researchers can measure physiological reactions to emotionally-charged images, indicative of changes in brain activity, when people are exposed to certain visual stimuli. These changes can be measured using electroencephalography (EEG), a method that records the temporal aspect of brain activity in response to a time-locked stimulus (Weinberg and Hajcak, 2010). Images of human and animal babies often evoke a positive affect, while photographs depicting human rights violations, mutilation, and threatening situations regularly elicit a negative response (Schienle et al., 2008). Though this trend has been established for pictures in terms of neural activity using EEG, fMRI, and PET technology (Weinberg and Hajcak, 2010; Kensinger and Schacter, 2006; Hajcak and Olvet, 2008; Mallan and Lipp, 2011; Schienle et al., 2008), a corresponding reaction to specific written words representative of these images has been studied to a lesser extent. The current study aims to explore any trends associated with related visual pictures and words of varying valence using EEG equipment.

Valence has been referred to by Kensinger and Schacter (2006) as how negative or positive a stimulus appears. Alternatively, arousal relates to how stimulating or calming an event is to the viewer (Kensinger and Schacter, 2006). In response to visual stimuli in the form of images, various event-related potentials (ERPs) have been shown to reliably indicate specific processes. Overall, the P300 peaks are

associated with automatic attention processes and appear in the early phase, while LPP peaks appear in the later phase and are representative of regulated attention and emotional processes (Nechvatal and Lyons, 2013). According to Schienle et al. (2008), P300 amplitudes correspond positively with emotionality; neutral stimuli would not be expected to elicit meaningful P300 amplitudes. Based on previous research, LPP amplitudes indicate an increase in arousal as a response to negative stimuli when compared to those of neutral stimuli (Weinberg and Hajcak, 2010; Mallan and Lipp, 2011; Schienle et al., 2008; Hajcak and Olvet, 2008).

After analyzing and applying learning concepts from B. F. Skinner, it is plausible that, from an evolutionary standpoint, our response to positive and negative stimuli could be necessary for survival. Our attention to negative stimuli arises from the fear of being harmed or killed, while our attention to positive stimuli arises from the need to seek companionship, have families, etc. (Skinner, 1984). This perspective supports humans' seemingly predisposed avoidance of threatening visual stimuli and inclination towards social or familial situations. Overall it complements what is currently known about physiological brain responses to visual stimuli.

Though research on neural responsiveness to words is limited compared to pictures, Citron (2012) found that non-verbal stimuli, such as pictures, elicited higher brain activity, evidenced by ERPs, related to emotions than for verbal stimuli, such as words. Skinner's (1984) concepts can be applied in understanding and explaining this idea as well. Humans have been able to see positive and negative visual stimuli for thousands of years, but have only recently begun reading. Therefore we would expect heightened neural responsiveness to pictures compared to written words.

Findings from Schupp et al. (2007) indicate that increased P300 and LPP amplitudes are observed corresponding to high-arousing stimuli (their categories included erotica and mutilation stimuli) when compared to low-arousing controls (neutral category). But researchers cannot reasonably expect all subjects

to have the same emotional response to the same items in reality, which emphasizes the need to standardize the subjective experience in some way. The magnitude of arousal for a particular image shown to a particular participant is still unknown. No two positive pictures will elicit the same physiological response; some pictures may elicit a powerful response in the form of a large amplitude, while others may be less arousing to a specific individual. Further, amplitudes corresponding to positive and negative stimuli may be significantly higher compared to those associated with neutral stimuli, but there's no way of knowing whether they are different from each other or not unless they are separated into high- and low-arousal subgroups, which is the aim of the current follow-up analysis of an initial study. It is also worth noting that, as Citron (2012) mentions, valence and arousal have been shown to be positively correlated to some degree, in that higher arousal may be indicative of more positive or negative valences compared to neutral.

The initial group experiment of the current study was designed to determine how pictures and words affect arousal/valence based upon ERPs. It was also designed to investigate if the positivity or negativity of the words and pictures would affect the arousal/valence based upon ERPs. Specific ERP components that were investigated included P300 and LPP amplitudes. Schienle et al. (2008) used the following time windows in their design, as a general reference: 0.34-0.50s after the stimulus was presented for P300 and 0.55-0.77s after the stimulus was presented for LPP in a previous study. In the initial experiment of the current study, we expected positive and negative stimuli to cause larger amplitudes in the P300 and LPP peaks in comparison to neutral stimuli because positive and negative stimuli were expected to cause more emotionality (especially in the P300 component) and arousal (primarily in the LPP component). Neutral stimuli were not expected to show significant deviation from the baseline for the P300 or LPP peaks. In accordance with the biological preparedness theory, which states that our arousal to certain stimuli is based upon experience and inheritance patterns (Citron, 2012), we expected to see greater brain activity

for picture stimuli compared to that of word stimuli, which would be reflected by higher P300 and LPP amplitudes.

Despite the clear trends we expected to see as a result of this study based on previous literature, we also needed to foresee the inevitability of confounds. Potential confounds that were likely to emerge included varying degrees of emotional impact on an individual basis and familiarity with certain stimuli due to previous experiences, which would directly impact individuals' interpretations of and responses to them. Certain stimuli that researchers deem to belong in the positive category might, in fact, induce fear in any number of participants, skewing the group averages. In effect, a wide range of emotions could be triggered based on previous exposure and experience with certain images due to individual differences.

Hence, we intended, through a planned follow-up analysis, to take the findings of Schupp et al. (2007) and Citron (2012) into consideration in the further analysis of the results obtained from the initial experiment. The categories approximated the categories established by Schupp et al. (2007), but stayed away from the extremes of erotica and mutilation. In place of using these more specific and predetermined "high-arousing" categories, this analysis used individuals' ratings to identify what the participants considered to be relatively high-arousing or low-arousing. This is an important distinction that should be made in order to account for individual differences. Since stimuli were placed into categories prior to the study, we could not expect that each participant would view the selected stimuli as fitting into the categories to which they are designated. Due to this restructuring, the "high-arousal" category comprised those stimuli that each participant considered to be the most positive and most negative. Similarly, the "low-arousal" category was determined by participants' ratings of least positive and least negative stimuli.

Through this continued analysis, we hoped to reveal a relationship between High-Arousal and Low-Arousal Stimuli, based upon participants' subjective ratings, to brain activity. After determining how individual participants

rated each of the positive and negative stimuli, we identified, on an individual basis and on a group basis, those stimuli that were rated most positive and most negative compared to those that were rated least positive and least negative. In light of existing literature responding to similar questions, and the efforts made to ameliorate any confounds, we expected to see a distinction in the LPP component, especially in Channel 2, which centers over the parietal lobe. Specifically, we expected those items falling into the High-Arousal Stimuli category to elicit significantly higher amplitudes in the LPP component when compared to those in the Low-Arousal Stimuli category. In terms of how stimuli are subjectively rated, we did not expect to see a significant difference between the numbers of stimuli rated high or low within the categories established. In order to get a better idea of how the number of valid ERPs across participants compared between High-Arousal and Low-Arousal Stimuli, analysis was run on the number of valid EEGs, as well. We expected there to be minimal statistical significance, since a significant difference between valid trials for High-Arousal versus Low-Arousal Stimuli would skew the data due to a lack of reliable data from which to draw.

Materials and Methods

Group ERP Experiment

Participants

A total of 21 participants (9 males, 12 females) were recruited from the student body at Roanoke College in Salem, VA, including the six student authors. The participants included students from ages 18-23 years. All participants were required to have normal or corrected-to-normal vision and could not participate if susceptible to seizures induced by flashing stimuli. The experiments were approved by and conducted in accordance with the guidelines of the Roanoke College Institutional Review Board with subjects providing informed consent.

Equipment

EEG signals were recorded using a PowerLab 26T device from AD Instruments, Inc. (Colorado Springs, CO). Five lead-shielded

electrodes transmitted voltage signals from the scalp of the participants. The ground (Fp2), Fp1, and right earlobe electrodes were disposable and simply stuck to the skin with their adhesive backing. The remaining electrodes, Oz and Pz, were attached to participants' scalps using electrode paste. All electrodes connected to a bio-amp specially designed to record signals in the biologically relevant range and to minimize artifacts from other electrical devices in the room. The room was designed in consideration of reducing artifacts. The analog input from the electrodes was converted by a PowerLab 26T into a digital time series output that was sent to a computer for additional processing by LabChart 7 software (AD Instruments, Inc). The temporal presentation of the stimuli was indicated by a signal sent from an external Cedrus StimTracker (San Pedro, CA) device to the same computer, which was also recorded by the LabChart 7 software. The LabChart 7 software was run on a Dell XPS 15z laptop computer (Round Rock, TX) and presented on an external 17" Dell monitor to be viewed by the experimenters only. The stimuli were presented on the internal 15" widescreen monitor of the laptop using SuperLab 4.5 (San Pedro, CA). The output of the channels was band pass filtered with an acceptable range from 0.5 to 50 Hz and a Mainz digital filter was applied.

Stimuli

The experimental stimuli in each block consisted of a total of 120 trials, comprising 60 pictures and 60 words, which were displayed in random order. Consensus reactions of all experimenters to each picture determined which grouping each would fall under: positive, neutral, or negative. The pictures, selected from the internet, consisted of 20 pictures that elicited positive emotions, 20 pictures that elicited neutral emotions, and 20 pictures that elicited negative emotions. After determining which pictures would be used, researchers chose one or two words to capture each picture, resulting in 20 positive words, 20 neutral words, and 20 negative words that corresponded directly with the pictures selected.

Additionally, task stimuli pictures of houses and the word 'house' appeared eight





Positive stimulus	Neutral stimulus	Negative stimulus	Task stimulus
Panda	Water	Snake	House
			

Table 1. Sample stimuli. These samples include one stimulus in the form of a word along with its picture counterpart from each of the categories. The task stimulus for which the participant was instructed to blink is illustrated as the picture of the house and the word “house.”

times each. One example from each stimulus category is displayed in Table 1. When a house stimulus was shown, the participant’s task was to blink. This additional instruction regarding task stimuli was included in order to control for excessive interference with ERP data due to artifacts. If participants could refrain from blinking until reaching the randomly inserted task stimuli, the likelihood of more accurate ERP data for those stimuli with which researchers were most concerned increased greatly. It was decided to use blinking specifically as the task response due to the recognizable electrical pattern of eye blinks within ERP data, which can easily be filtered out.

Each picture was displayed in grayscale in 250x250 pixels, while the words appeared in white font in size 24 against a gray background. There were a total of 136 trials per block. Each trial was presented for 0.50s and there was an interval of 1.00 to 1.20s between trials, with the time between trials randomly chosen to avoid a constant interval in relation to the ongoing EEG oscillations. On average there was 1.50 to 1.70s from the start of one trial to the start of the next trial, making the total duration of each run approximately four minutes. There were five blocks of the 136 trials, making the total experiment roughly 20 minutes. This resulted in 100 trials for each of the six stimulus conditions.

Procedures

The study recorded amplitudes using five electrodes, measuring two channels. Channel 1 consisted of two electrodes: one at the anterior pole (Fp1) and one at the posterior

pole (Oz) of the fronto-occipital axis. Channel 2 comprised an electrode on the right earlobe to serve as a reference and one above the centro-parietal area (Pz), together measuring signals from the centro-parietal axis. The fifth electrode was the ground (Fp2), placed on the forehead next to the Fp1 electrode.

Potential participants were presented with an informed consent sheet and pre-screened when they entered the experiment room. The pre-screening ensured that each participant was not susceptible to seizures, had normal or corrected-to-normal vision, and did not have metal in their head. They were also warned about the possibility of a migraine headache as a result of viewing the moderately-rapidly changing stimuli.

After giving informed consent, participants abraded their skin where the electrodes were to be placed. Once participants cleaned the abraded areas with alcohol swabs, experimenters attached electrodes to the participants’ scalps. All electrodes were held in place by elastic headbands and, additionally, the anterior, ground, and earlobe electrodes were kept in place by disposable electrodes. The locations for the anterior, posterior, and ground were chosen based on anatomical locations. The anterior and ground electrodes were placed symmetrically around the center of the forehead. The posterior electrode was placed one inch above the inion (the bump on the back of the head). A headband was wrapped around the forehead and back of the head to hold in place the anterior and posterior electrodes. The earlobe and medial electrodes were placed in a practical fashion in relation to the other

electrodes, though constrained by anatomical features. The medial electrode was placed dorsal and anterior on the scalp in relation to the placement of Oz, beneath a headband that was wrapped below the chin and over the top of the head. The remaining electrode was placed on the right earlobe with a disposable fastener.

Prior to the start of data collection, the electrode connections were tested by viewing the output of the channels on the computer screen. It was verified that the EEG signal for each channel remained within the range $\pm 60 \mu\text{V}$ when the participant held their head and eyes still and did not blink. During data collection, the signal was sampled at a rate of 400/s. A pillow was placed on a stack of books for a chin rest to aid the participant in keeping their head still. The size of the stack of books varied, ensuring that the participant was comfortable.

Each participant completed five blocks, with each block lasting approximately four minutes. The participant had the chance to take a break in between each block; the blocks were continued when the participant was ready to resume. After the completion of all blocks, the electrodes were removed, cleaned, and the participant was debriefed.

Following completion of the main experiment, the participant was given an electronic questionnaire in which they rated their general emotions toward the stimuli that they saw. The rating scale was a Likert-type scale, ranging from 1 to 9 in which 1 was the least pleasant, 5 was neutral, and 9 was the most pleasant. Stimuli also included one image and one word from the task-oriented stimuli. Therefore, there were 122 stimuli that participants rated. This took the participants on average 3-5 minutes to complete, and was run on the same laptop using a similar SuperLab program. The purpose of the experiment and the expected results were explained to the participant after the completion of the questionnaire. For most participants the total duration was approximately 40 minutes, including the set-up, body of the study, time for breaks, and questionnaire.

The data was collected as continuous EEG recordings. Any signals outside of the $\pm 60 \mu\text{V}$ were considered non-psychological and were excluded from the data. A specifically-labeled

comment marker appeared in the channel that correlated with a particular stimulus when that stimulus appeared.

Statistical Analysis

The EEG data collected was subsequently analyzed using MatLab (Natick, MA) and SPSS (Chicago, IL). Matlab was used to exclude those trials that had fallen outside the acceptable range of ERP amplitudes. The P300 peak was defined in both Channel 1 and Channel 2 as occurring between the time range 0.29s and 0.39s, and the LPP peak was defined in Channel 1 and 2 as occurring between the time range 0.70s and 0.90s. These values were chosen as the location of visible deflections in Figure 1 when ERPs were averaged across people as they fit the current data better, though they are in a slightly different range than used by Schienle et al. (2008). For each of these ranges, the average amplitude of the waveform for each condition was collected, which allowed data for each variable at each time to be summarized as a single point. With SPSS, analysis was performed using a 3x2 (valence: positive, neutral, negative; stimulus: pictures, words). Repeated Measures ANOVA in the SPSS Statistics program. An alpha level of 0.05 was used throughout.

Based on Mauchly's Test of Sphericity, there generally was no evidence for any violations in the sphericity of the data. The one exception was for the interaction effect of stimulus by variance in the P300 effect in Channel 1, which was a non-significant effect. It is not possible to violate the assumption of sphericity in the Individual ERP analysis, because there were only two levels of each of the independent variables. Therefore, all ANOVA results will be reported with sphericity assumed.

Individual ERP Analysis

Participants

For further analysis, data from 10 participants, whose EEG recordings yielded large amounts of background noise with indiscernible ERPs, were discarded. As a result, the current analysis consisted of data from 11 participants (four males, seven females) between

the ages of 18-23 years from the Roanoke College student body. These participants in the further analysis were a subset of the previously described data.

Equipment

Analysis was conducted using SPSS software and Matlab programming on the same laptops used to conduct the experiment.

Stimuli

All of the stimuli used in this analysis were derived from the original study, but all of the original stimuli were not necessarily analyzed. Regardless of how many were employed in the analysis, all of the stimuli came from the images and words from the positive and negative subgroups only; all neutral stimuli and task stimuli were excluded from additional analysis. Task stimuli were eliminated in this analysis due to the irrelevance of the data they yielded in the experimental stage. Since they existed solely to control for interference of ERP data, they would therefore have little accuracy in the analysis stage. Specific ERPs analyzed

varied across participants, depending on which stimuli they rated as most positive, most negative, least positive, and least negative.

In determining which stimuli would be analyzed for each participant, two groups were formed: Group grouping and Individual grouping. In the Group grouping condition, anything that was rated as either a 7, 8, or 9 within positive stimuli and anything rated as a 1, 2, or 3 within negative stimuli were combined to create the High-Arousal Stimuli subgroup. Those stimuli rated as a 4, 5, or 6 within positive and negative stimuli comprised the Low-Arousal Stimuli subgroup within the Group groupings condition. The size of each subgroup, then, was dependent upon the variability amongst individuals' ratings. If, for example, an individual rated every negative stimulus as a 1, 2, or 3 and every positive stimulus as a 7, 8, or 9, then that participant would have no stimuli in the Low-Arousal Stimuli subgroup. This is a uniform approach ensuring that every stimulus in the High-Arousal group had the same range of ratings (of either 1, 2, 3, 7, 8, or 9) and that every stimulus in the Low-arousal group also

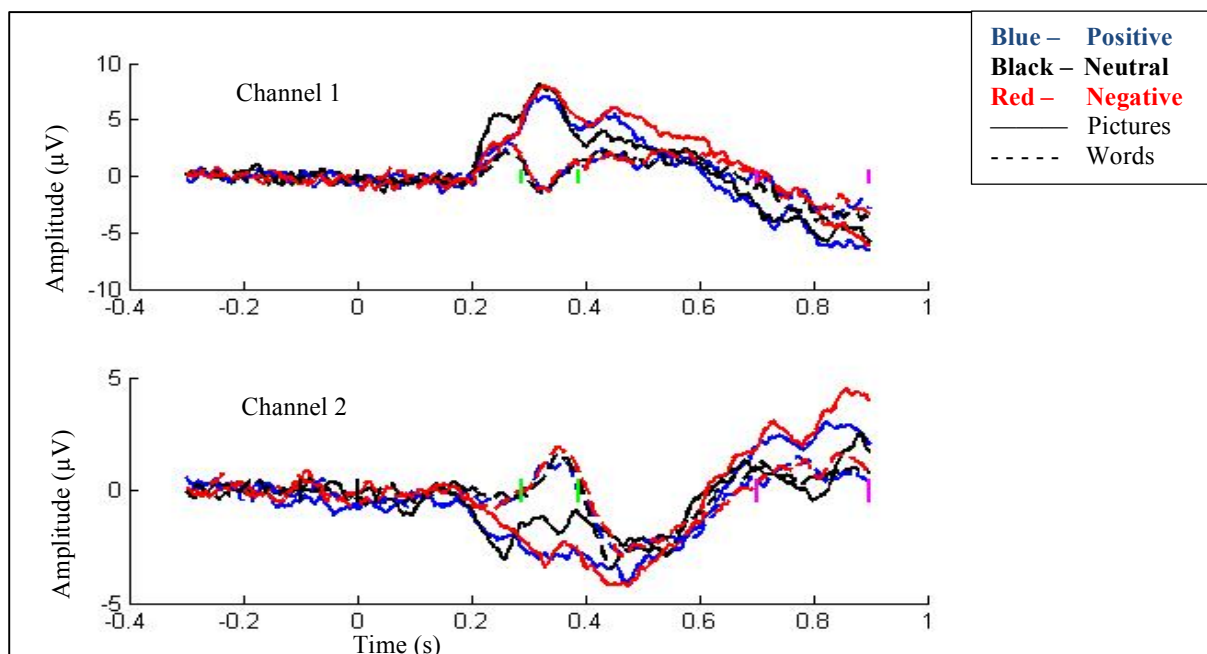


Figure 1. The different ERPs across time for each combination of stimuli and valence (e.g., positive word is a blue dash). The two sets of black vertical bars show the baseline. Further, the green vertical bars indicate the time range in which the P300 component was analyzed (0.29s to 0.39s), and the pink vertical bars indicate the time range in which the LPP component was analyzed across channels (0.70s to 0.90s).

had the same range of ratings (of either 4, 5, or 6). The tradeoff with using this classification system is that each individual does not necessarily have the same amount of data in each condition as the others.

In the Individual groupings condition, the five most positive ratings in the positive stimuli group comprised part of the High-Arousal Stimuli, along with the five most negative ratings in the negative stimuli group. Each subset could therefore include more than five stimuli and their corresponding ERPs. If, for example, an individual's most negative ratings include two ones, two threes, and four fours, the number of ERPs analyzed for that participant in the subcategory would be eight, before determining the number of positive ratings that would combine to form the High-Arousal Stimuli subgroup. Although the first five of those would only include the two ones, two threes, and one of the fours, it would be impractical to assume that only one of the stimuli given the rating of four was any less negative than the remaining stimuli with the same rating. The same methodology determined which stimuli fell into the category of least positive and least negative, which comprised the Low-Arousal Stimuli group. This individualistic approach ensures that each participants' subjective opinion was considered in relation to the rest of their ratings (i.e., some participants didn't use one to rate any stimuli, but a three would still be the most negative in relation to their other ratings). The drawback, however, is that some data had to be thrown out (i.e., if the participant rated five of the twenty items as a 1, five of them as a 5, and the remaining ten as anything in between, those middle ten items were not used in compiling that participants' data). Table 2 shows one individual's rating of stimuli 1-20 in random order from the positive pictures subcategory.

Due to the nature of participants' subjective ratings, some of the participants included in the follow-up analysis may have had as few as 10 or as many as 170 valid trials per condition with which to analyze. Beyond this initial check for validity, no single participant ended up with 100% of valid ERPs that corresponded with their ratings; the highest

Stim	Rating	Group grouping	Individual grouping
9	6	1	1
4	6	1	1
16	6	1	1
17	6	1	1
19	6	1	1
7	6	1	1
18	7	2	0
12	7	2	0
13	7	2	0
20	7	2	0
14	7	2	0
10	7	2	0
8	7	2	0
3	7	2	0
5	8	2	2
2	8	2	2
6	8	2	2
1	8	2	2
11	8	2	2
15	9	2	2
Total Low		6	6
Total High		14	6

Table 2. An example of how stimuli were coded according to the Group groupings and Individual groupings classification systems. This sample comes from one participant's ratings of the positive pictures for all 20 stimuli in that category. This particular participant relied solely on ratings 6, 7, 8, and 9. The bottom two rows of the table show the total number of stimuli falling into the Low category and the High category before factoring in the individual's ratings of negative pictures. The "1" indicates that the particular stimulus falls into the High-arousal Stimuli subcategory and "2" indicates that it falls into the Low-arousal Stimuli subcategory. While under the Group groupings system there were more than twice as many ratings coded as High compared to Low, the number of each was equal in the Individual groupings system.

percentage of ERPs among participants was 94.55% while the lowest was 25.00%.

Procedures

Further analysis was conducted using text documents recorded by the SuperLab

program during the experiment. Using these text documents, we were able to determine how participants rated stimuli in the questionnaire, as well as in which order the stimuli were presented for each of the five blocks for every participant. The setup of the SuperLab program enabled us to match comment markers to individual stimuli. Additionally, MatLab was used to average participants' ERPs across the five blocks for each category. SPSS software was integral in analyzing data and statistics in order to move forward and make conclusions. In addition to analyzing the P300 and LPP components, SPSS was used to discern whether or not significant differences between the frequency of high and low ratings existed, and further to determine if the percentage of valid trials varied significantly from expectations.

Results

Group ERP Experiment

For each participant, ERPs were determined by taking the time-locked average of the EEG across all five blocks for each event of interest, while keeping each condition separate. The amplitudes of these ERPs were then averaged across participants within different time ranges for the different dependent variables (Fig. 1).

After running the 3x2 Repeated Measures ANOVA for the subjective ratings, we found that there was an interaction effect present between valence and stimuli, such that how extreme the ratings of positive and negative were depended on whether the stimuli were in the form of pictures or words ($F(2,40)=22.929$, $p<0.001$) (Fig. 2). While positive pictures were rated more positively than positive words, negative pictures were rated more negatively than negative words. Although there was no significant main effect of stimulus ($F(1,20)=2.812$, $p=0.110$), there was a significant main effect of emotion, in which positive was rated higher than neutral, which was rated higher than negative ($F(2,40)=249.012$, $p<0.001$). Overall, it is evident that the valence rating survey validated the category of the stimuli used for each

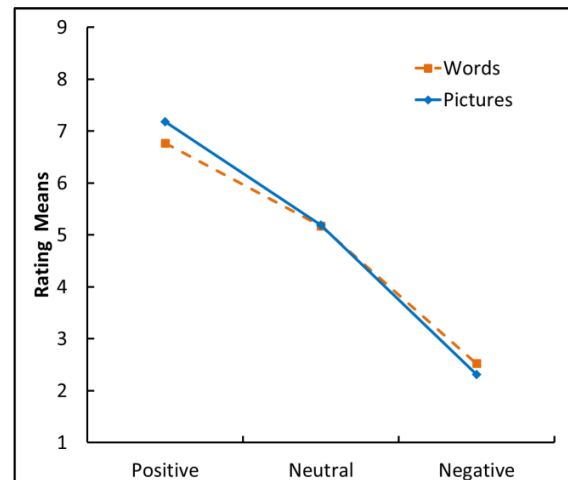


Figure 2. Mean subjective ratings of each stimulus. Subjects found words and pictures that were supposed to be positive as predominately positive, neutral as being neutral, and negative pictures and words were seen as negative.

condition (positive, neutral, and negative) for both pictures and words.

Upon analyzing the P300 amplitude in Channel 1 with Repeated Measures ANOVA, we found that there was no significant interaction effect ($F(2,40)=0.634$, $p=0.536$) (Fig. 3). There was a significant main effect of stimuli in which pictures had greater amplitudes than words ($F(1,20)=72.466$, $p<0.001$). A

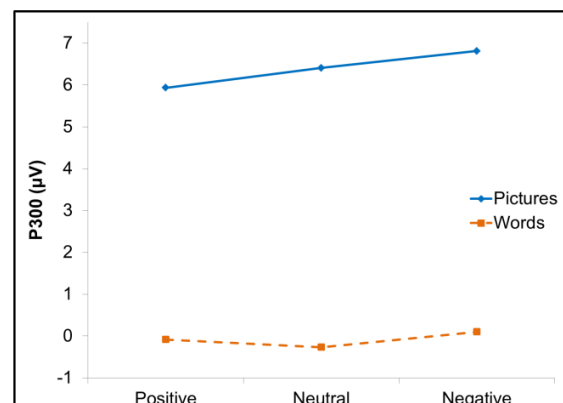


Figure 3. P300 output for Channel 1. This graph shows a main effect of stimuli where pictures show greater brain activity than words.

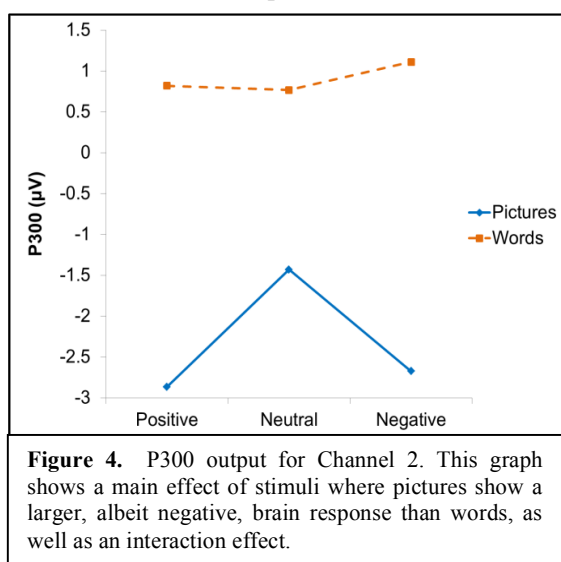
significant main effect of valence was not present ($F(2,40)=1.391$, $p=0.261$).

Performing a Repeated Measures ANOVA on the P300 amplitudes in Channel 2 revealed that an interaction effect was present

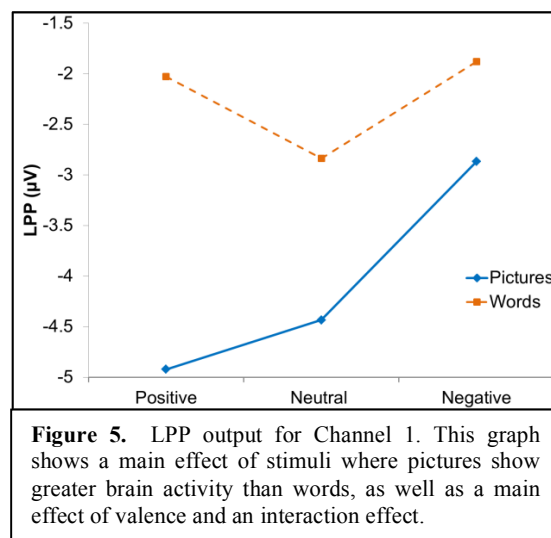
($F(2,40)=5.254$, $p<0.05$) (Fig. 4). This interaction effect suggests the brain activity observed in response to positive and negative stimuli depends on the form of the stimulus: there was a greater effect of emotionality for pictures compared to words. There was also a significant main effect of stimuli in which pictures had greater amplitude than words ($F(1,20)=23.321$, $p<0.001$). However, there was not a significant effect of valence, indicating no consistent difference between positive, negative, or neutral stimuli ($F(2,40)=3.125$, $p=0.055$).

Analysis of the LPP component using Repeated Measures ANOVA revealed significant results across channels. In Channel 1 (fronto-occipital axis), an interaction effect was found, which suggests that the response towards the valence of the stimulus is dependent upon the nature of the stimulus (picture or word) ($F(2,40)=6.114$, $p<0.05$) (Fig. 5). Additionally, there was a significant main effect of stimuli, in that pictures produced higher amplitudes than words ($F(1,20)=22.201$, $p<0.001$). Valence also displayed a significant main effect ($F(2,40)=4.806$, $p<0.05$), indicating that positive and negative stimuli evoke more brain activity than neutral stimuli.

The same pattern of results was



obtained through Repeated Measures ANOVA in the LPP time range of Channel 2 (centro-parietal lobe) (Fig. 6). There was an interaction effect ($F(2,40)=7.016$, $p<0.05$), which was such that whether valence (positive or negative



compared to neutral) affected the physiological response depended on the form of the stimulus. We see a greater response for positive and negative pictures compared to neutral, but little difference between positive and negative words compared to neutral. A main effect of stimulus was also present ($F(1,20)=27.484$, $p<0.001$), suggesting that pictures elicit greater brain activity than words again, as well as a main effect of valence ($F(2,40)=8.386$, $p<0.005$).

Individual ERP Analysis

Ratings

As was predicted, a 2x2 (valence: High- and Low-Arousal; stimuli: pictures, words) Repeated Measures ANOVA design on the frequency of stimuli falling into the different categories yielded few significant results. Within the Group groupings category, there was no interaction effect ($F(1,10)=3.447$, $p=0.093$), nor was there a main effect of stimuli ($F(1,10)=3.750$, $p=0.082$). There was a main effect of valence ($F(1,10)=20.093$, $p<0.05$) such that pictures less frequently fell into the Low-Arousal category and more frequently fell into the High-Arousal category when compared to words. On average, there were 11.27 stimuli falling into the Low-arousal Stimuli for pictures, compared to 28.00 stimuli in the High-arousal Stimuli for pictures. The large difference in number of stimuli was consistent in words as well, with 15.18 stimuli in the Low-arousal

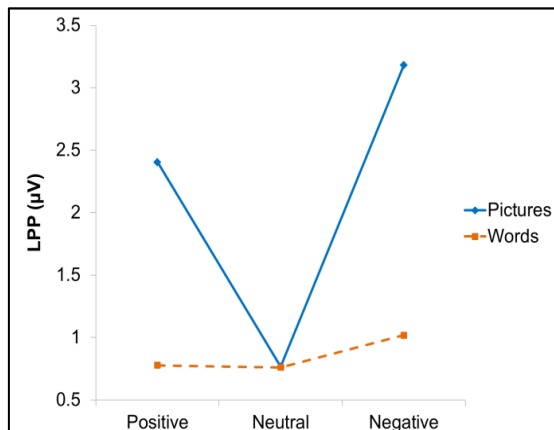


Figure 6. LPP output for Channel 2. This graph shows an interaction effect in that emotionality matters for pictures but not for words, along with a main effect of stimuli and of valence.

Stimuli category and 24.36 stimuli in the High-arousal Stimuli category for words (Table 3).

Within the Individual groupings category, there was no interaction effect ($F(1,10)=4.479$, $p=0.060$), no significant main effect of stimulus ($F(1,10)=0.707$, $p=0.420$), and no main effect of valence ($F(1,10)=0.050$, $p=0.827$) present. The overall lack of significance suggests that the number of trials in each condition was more evenly balanced than in the Group groupings classification. The average number of stimuli within each category is reflective of this (Low-Arousal Stimuli, pictures: 14.55; High-Arousal Stimuli, pictures: 17.82; Low-Arousal Stimuli, words: 17.91; High-Arousal Stimuli, words: 15.73).

Valid ERPs

Repeated Measures ANOVA analysis confirmed the hypothesis that there was no significant difference between the percentages of valid ERPs across participants in the High-Arousal and Low-Arousal Stimuli subgroups. Within the Group groupings classification, data revealed no interaction effect ($F(1,10)=2.173$, $p=0.171$) and no significant main effects of stimuli ($F(1,10)=3.564$, $p=0.088$) or valence ($F(1,10)=0.003$, $p=0.960$). Channel 2 revealed a similar pattern (interaction effect: $F(1,10)=2.926$, $p=0.118$; main effect of stimuli: $F(1,10)=1.753$, $p=0.215$; main effect of valence:

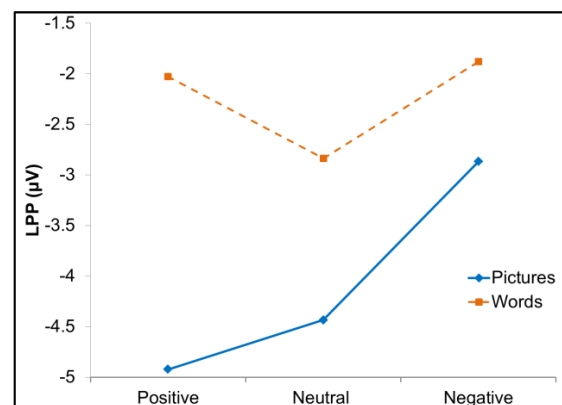


Figure 5. LPP output for Channel 1. This graph shows a main effect of stimuli where pictures show greater brain activity than words, as well as a main effect of valence and an interaction effect.

$F(1,10)=0.775$, $p=0.399$), indicating no significant results.

Analysis of the Individual groupings classification similarly yielded no significant results for Channel 1 (interaction effect: $F(1,10)=1.474$, $p=0.253$; main effect of stimuli: $F(1,10)=0.037$, $p=0.852$; main effect of valence: $F(1,10)=0.020$, $p=0.891$) or for Channel 2 (interaction effect: $F(1,10)=0.173$, $p=0.686$; main effect of stimuli: $F(1,10)=0.052$, $p=0.824$; main effect of valence: $F(1,10)=1.804$, $p=0.209$).

P300 Analysis

Upon analyzing the P300 amplitude with a 2x2 (valence: High-Arousal and Low-Arousal; stimuli: pictures and words) Repeated Measures ANOVA, data revealed few significant results. In the Group groupings classification in Channel 1, an interaction effect was observed ($F(1,10)=5.448$, $p<0.05$) suggesting that arousal increased when comparing Low-Arousal and High-Arousal categories (High-Arousal showed a greater P300 amplitude), but only when the stimulus was in the form of a word. As will be discussed later, this interaction effect is not believed to be meaningful. Data also revealed a main effect of stimulus ($F(1,10)=43.002$, $p<0.001$), in that pictures consistently elicited higher amplitudes than words. There was, however, no main effect of valence ($F(1,10)=0.561$, $p=0.471$).

Channel 2 within the Group groupings categorization revealed no interaction effect ($F(1,10)=2.380$, $p=0.154$). There was a main effect of stimulus presented ($F(1,10)=13.535$, $p<0.05$), but not of valence ($F(1,10)=0.392$, $p=0.545$). In Channel 2, the main effect of stimulus indicated the same trend in which pictures consistently elicited higher amplitudes than words (Fig. 7).

The Individual groupings classification yielded similar results across both Channels. Neither Channel revealed a significant interaction effect (Channel 1: $F(1,10)=1.346$, $p=0.273$; Channel 2: $F(1,10)=0.091$, $p=0.769$). In both Channel 1 and 2, a main effect was observed for stimuli (Channel 1: $F(1,10)=31.144$, $p<0.001$; Channel 2: $F(1,10)=14.686$, $p<0.05$). Similar to results of main effect of stimulus found in the Group groupings system, both Channels in the Individual groupings revealed consistently higher amplitudes associated with pictures over words. However, neither Channel showed a significant main effect of valence (Channel 1: $F(1,10)=0.380$, $p=0.551$; Channel 2: $F(1,10)=0.003$, $p=0.955$).

LPP Analysis

After running the same Repeated Measures ANOVA analysis on the LPP component, relatively similar results ensued. No significance was revealed in the Group groupings category in Channel 1 for the

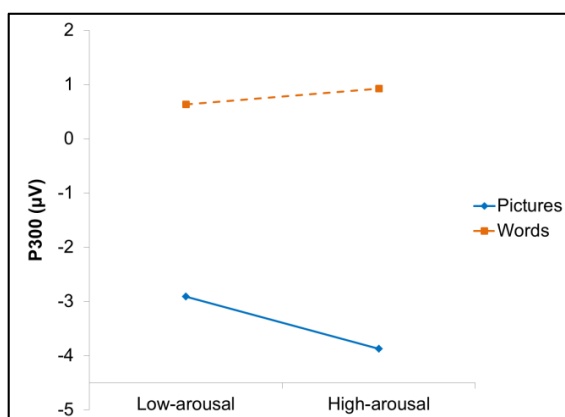


Figure 7. P300 output for Channel 2 within the Group groupings classification. This graph shows a main effect of stimuli where pictures show greater brain activity than words. This graph exemplifies the findings found in other areas of P300 results as well.

interaction effect ($F(1,10)=1.518$, $p=0.246$), main effect of stimulus ($F(1,10)=4.337$, $p=0.064$), or main effect of valence ($F(1,10)=4.197$, $p=0.068$) (Fig. 8). Note, however, that while the main effect of stimulus in Channel 1 of Group groupings in the LPP component was not significant according to our alpha threshold of 0.05, it is certainly approaching significance at a p-value of 0.064. Channel 2 within Group groupings revealed no significant interaction effect ($F(1,10)=0.007$, $p=0.935$). It did indicate a significant main effect of stimulus ($F(1,10)=55.408$, $p<0.001$) in which pictures were consistently more arousing than words. However, no main effect of valence ($F(1,10)=2.948$, $p=0.117$) was found.

Within the Individual groups classification for the LPP component, Channels 1 and 2 both displayed a lack of interaction effects (Channel 1: $F(1,10)=0.018$, $p=0.895$; Channel 2: $F(1,10)=2.207$, $p=0.168$). There was a significant main effect of stimulus across both channels (Channel 1: $F(1,10)=8.573$, $p<0.05$; Channel 2: $F(1,10)=52.919$, $p<0.001$). In both cases pictures consistently elicited more arousal than words. However, there was no main effect of valence observed in either Channel (Channel 1: $F(1,10)=1.956$, $p=0.192$; Channel 2: $F(1,10)=0.612$, $p=0.452$).

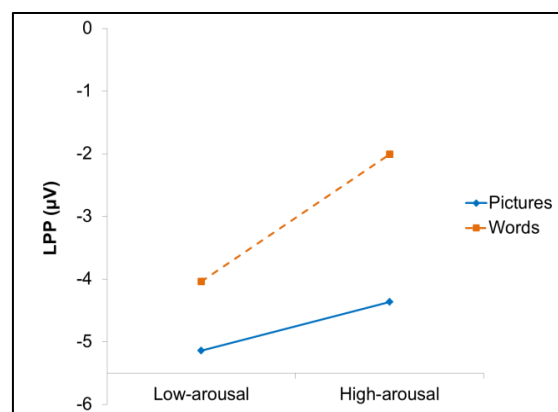


Figure 8. LPP output for Channel 1 within the Group groupings classification. Similarly to Figure 7, this graph shows a main effect of stimuli in that pictures consistently elicit higher amplitudes than words. This graph also serves as a representation of the findings found in other areas of LPP results.

Discussion

The results from the initial experiment showed that P300 amplitudes across channels were consistently higher for pictures than words. Unlike Channel 1, Channel 2 revealed higher P300 amplitudes for stimuli associated with emotionality, whether positive or negative, compared to neutral stimuli. With the current knowledge of the occipital lobe's primary role in basic visual processing, it is intuitive that Channel 1, with electrodes along the fronto-occipital axis, would reveal significant differences between the neural processing of pictures and words. Similarly, we would expect Channel 2 (an electrode over the centro-parietal lobe) to yield differences associated with arousal/valence due to the contribution of the parietal lobe in emotional processing, among other brain areas (Morris et al., 1998).

It is important to note that P300 amplitudes do not differentiate between positive and negative emotions, just valence (Schienle et al., 2008). This characteristic explains why we found a significant difference between positive and neutral stimuli, as well as between negative and neutral stimuli. As we would expect, therefore, there was no significant difference in P300 amplitudes between positive and negative stimuli.

The lack of interaction effect in Channel 1 indicates that one independent variable did not depend on the other. Specifically, responses to stimulus type had no effect on responses to valence, supporting the lack of emotional processing of the occipital lobe.

Data from Channel 2 shows a significant effect in arousal/valence in which positive and negative stimuli elicited greater P300 amplitudes than neutral stimuli. This evidence is supported by the activation of the amygdala during emotional exposure to the three conditions. The amygdala is the neurological area that evaluates the encoding of valence and arousal in addition to emotional memory (Phelps, 2006). The stimuli causing activation of the amygdala produce greater P300 amplitudes in comparison to neutral stimuli (Phelps, 2006).

Results from Channel 2 in the LPP time range indicate differential processing of emotional stimuli, such that positive and negative visual stimuli produced higher magnitude LPP peaks, but only for pictures and not for words. Also worth noting is the absolute lack of difference amongst the reactions to neutral stimuli, regardless of whether the neutral stimulus was a picture or a word. This interpretation supports the distinction among neural areas in information processing; unlike the occipital lobe (whose activity is measured in Channel 1), the parietal lobe (whose activity is measured in Channel 2) is not involved in basic visual processing, but rather in the distinction necessary for emotional processing.

Though the categories Schupp et al. (2007) delineated, i.e., High- and Low-Arousal Stimuli, do not coincide directly with those of this study, ours imitate theirs on a less extreme level. We expected to see similar results from the additional analysis based on these intrinsic similarities. While the current study had many subjects for a within-subject, repeated measures design, ERP data was not observed for all participants, which limits the accuracy of the results. The patterns observed in this study could not be validated for all subjects. Also, the reduction in the number of trials per condition by separating subsets of the positive and negative stimuli into High- and Low-Arousal conditions further limited the number of valid trials included in the ERP analysis.

After running analyses using the new subgroups of High-Arousal and Low-Arousal Stimuli in the follow-up analysis, no significant differences were found between the individually assigned subjective ratings of stimuli by participants, regardless of the Channel analyzed (Channel 1 or Channel 2), the classification system used (Group groupings or Individual groupings), or the component analyzed (P300 or LPP). Regardless of whether a stimulus was in the form of a picture or word, and regardless of whether it was identified as extremely negative or extremely positive, we expected to see higher P300 and LPP amplitudes in the High-Arousal Stimuli category when compared to those of the Low-Arousal Stimuli. The overall lack of main effects of valence was therefore not in line with

the predictions established, but there could be a number of reasons explaining this.

Once the original group of 21 participants was narrowed down to the final 11, those 11 participants were included in every level of analysis. The fewer valid ERPs each participant had to draw from, the less reliable their data was. It is worth noting, however, that the lack of significant main effects and interaction effects in the Repeated Measures ANOVA of percent valid EEG discounts this possibility as serving as a confound. While analysis run on the Group groupings coding system revealed a significant difference between the number of high and low ratings across participants, the Individual groupings system, and its lack of main effects, supports that confounds were minimal. These two analyses served as verifiers to ensure that, though they may have played a minor role in the lack of significant results, they were not significant enough to affect the statistics.

While there was one instance among P300 analysis where we saw an interaction effect, in the Group groupings classification of Channel 1, this could potentially be attributed to the possibility of a Type I error. Type I errors emerge when we inappropriately reject the null hypothesis when it is in fact true. Having an alpha value of 0.05 means that there is a 5% chance that the significance revealed is not truly significant. Though it would be impossible to determine whether or not this is the case, it is mathematically very likely in light of the amount of analysis performed. As the number of analyses increases, so does the likelihood of making a Type I error. The formula for determining the probability of a Type I error is $1-(1-05)^n$, where n is equal to the number of analyses run (16 in this case). Thus, the likelihood that the significant finding here is reflective of a Type I error is roughly 56%. Since this finding doesn't align with our predictions (if anything, pictures should be changing, not words), we might be inclined to attribute it to a Type I error.

Based on these results, it would be possible to conclude that input that is self-reported as being very positive or very negative actually does not necessarily correspond directly with similar physiological responses, which we

would expect to be evidenced by increased amplitudes of the P300 and/or LPP components. Similarly, we can conclude that stimuli that are perceived as subjectively more arousing are not guaranteed to be physiologically more arousing, as well. These results would seem to imply that our perceptions and impressions do not quite mimic underlying functions of the brain. Due to the lack of significance, we cannot reject the null hypotheses and must accept that either the research design did not incorporate sufficiently arousing stimuli (i.e., more graphic images) or that the trends we expected to see are not truly present in the brain.

In almost all conditions analyzed with Repeated Measures ANOVA, the arousal evidenced by P300 and LPP amplitudes corresponding with verbal stimuli increased from Low-Arousal Stimuli to High-Arousal Stimuli. The LPP component in Channel 2—in both Group groupings and Individual groupings—indicate a decrease in arousal from Low- to High-Arousal Stimuli, but the remaining conditions reveal an increase in this arousal. These results may indicate deeper semantic processing associated with more emotionally-charged words.

Other studies have supported this idea. Chun-Tuan (2006) found that participants respond better to health information through advertisements when the information is accompanied by graphics (visual stimuli). Their findings indicate that pictures elicit higher P300 amplitudes across positive and negative stimuli, which can be interpreted as more processing of the visual stimuli (Chun-Tuan, 2006).

Educational approaches could greatly benefit from the knowledge that pictures evoke increased brain activity compared to words. For those who identify themselves as visual learners and may not function as well in the 'one-size-fits-all' school model the United States has largely adopted over the years, this evidence can be of huge support. Felder (1993) makes the claim that visual learners predominate science classrooms in the college setting of Western cultures. Despite this fact, the most prevalent teaching styles make use of verbal information and sparingly employ non-verbal representations (Felder, 1993). In order to more adequately increase retention rates in these settings, then,

Felder (1993) urges that information be taught in a way that coincides with the students' learning styles. This advice might incline professors in these disciplines to present information graphically and pictorially rather than verbally. Since many of the students who participated in the study were Psychology or Biochemistry majors, it may be that these students were already visual learners who thus responded more acutely to images rather than words.

Additionally, an inherent limitation when presenting familiar images and words is an individual's previous experience with those stimuli. Pollak et al. (1997) demonstrated this when looking at P300 amplitudes within maltreated children specifically. These children elicited higher P300 amplitudes in response to angry stimuli than controls. Likewise, different stimuli could have triggered unexpected emotions for participants, which limits the accuracy of the data.

Conroy and Polich (2007) suggested isolating the variables of valence and arousal in order to get a more accurate indication of P300 amplitudes corresponding directly with valence. Since P300 amplitudes are associated with increased processing and attention time, it follows that "subjects allocate more looking time to intense and arousing images regardless of valence" (Conroy and Polich, 2007), making it impossible to attribute P300 peaks to one measure over the other. Our observed findings of increased P300 amplitudes, therefore, could be due to higher arousal states elicited by certain stimuli, which would effectively skew what we perceived as attributable to valence. In light of the findings from their study and the subsequent uncertainty of ours as a result of this information, future research could involve isolating valence by ensuring equal levels of arousal across stimuli.

Acknowledgements

We would like to thank the Roanoke College Department of Psychology for supplying the EEG equipment and analysis materials.

Corresponding Authors

Lauren Kennedy
Roanoke College
221 College Lane
Salem, VA 24153
lrkennedy@mail.roanoke.edu

References

- Chun-Tuan C (2006) Is a picture worth a thousand words? Influence of graphic illustration on framed advertisements, *Adv Consum Res* 33:104-12.
- Citron FMM (2012) Neural correlates of written emotion word processing: A review of recent electrophysiological and hemodynamic neuroimaging studies, *Brain Lang* 122: 211-26.
- Conroy MA, Polich J (2007) Affective valence and P300 when stimulus arousal level is controlled, *Cogn Emot* 21:891-901.
- Felder R (1993) Learning and teaching styles in college science education, *J Coll Sci Teach* 23:286-90.
- Hajcak G, Olvet DM (2008) The persistence of attention to emotion: Brain potentials during and after picture presentation, *Emotion* 8:250-5.
- Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: A module in human extrastriate cortex specialized for face perception, *J Neurosci* 17:4302-11.
- Kensinger EA, Schacter DL (2006) Processing emotional pictures and words: Effects of valence and arousal, *Cogn Affect Behav Neurosci* 6:110-26.
- Mallan KM, Lipp OV (2011) The relationship between self-reported animal fear and ERP modulation: Evidence for enhanced processing and fear of harmless invertebrates in snake- and spider-fearful individuals, *Motiv Emot* 35:474-83.
- McCandliss BD, Cohen L, Dehaene S (2003) The visual word form area: Expertise for reading in the fusiform gyrus, *Trends Cogn Sci* 7:293-9.
- Morris JS, Friston KJ, Buchel C, Frith CD, Young AW, Calder AJ, Dolan RJ (1998) A neuromodulatory role for the human

- amygdala in processing emotional facial expressions, *Brain* 121:47-57
- Nechvatal JM, Lyons DM (2013) Coping changes the brain, *Front Behav Neurosci* 7:1-8.
- Phelps EA (2006) Emotion and cognition: Insights from studies of the human amygdala, *Annu Rev Psychol* 57:27-53.
- Pollak S, Cicchetti D, Klorman R, Burmaghim J (1997) Cognitive brain event-related potentials and emotion processing in maltreated children, *Child Dev* 68:773-87.
- Schienze A, Schäfer A, Naumann E (2008) Event-related brain potentials of spider phobics to disorder-relevant, generally disgust- and fear-inducing pictures, *J Psychophysiol* 22:5-13.
- Schupp HT, Stockburger J, Codispoti M, Junghöfer M, Weike AI, Hamm AO (2007) Selective visual attention to emotion, *J Neurosci* 27:1082-9.
- Skinner BF (1984) The evolution of behavior, *J Exp Anal Behav* 41:217-21.
- Weinberg A, Hajcak G (2010) Beyond good and evil: The time-course of neural activity elicited by specific picture content, *Emotion* 10:767-82.