

Exploration of Methodological and Participant-Related Influences on the Number of Artifacts in ERP Data

Stephanie M. Shields¹, Caitlin E. Morse¹, and David F. Nichols¹

¹*Roanoke College, Salem, Virginia 24153*

Event-related potential (ERP) data has low signal-to-noise ratio, requiring the conduction of a large number of trials in order to collect sufficient amounts of data for subsequent analysis. Therefore, it would be highly beneficial if researchers could minimize the number of artifacts that occur in the data, minimizing the number of discarded trials and the total number of trials needed. This study thus examined connections between the number of trials that have to be eliminated due to artifacts and a set of methodological variables, physical considerations, and individual differences. In half of the electroencephalography (EEG) data collection blocks, naïve undergraduate participants were asked not to blink for the duration of the block (approximately 2.5 minutes), but in the other half, the stimulus set included blinking cues to give participants a chance to blink during blocks. The number of artifacts did not differ based on whether participants were cued to blink during blocks nor which type of block participants experienced first. However, the first block had significantly more artifacts than other blocks, and the third block had significantly fewer. Participants who had previously known one or both investigators had significantly fewer artifacts in their data than participants who had not, but no significant relationship was found between the number of artifacts and any other individual difference or physical consideration examined. These results imply that researchers could preemptively reduce the number of artifacts in their EEG data by including practice blocks and recruiting friends or acquaintances for studies if possible. Based on subjective, unsolicited participant feedback, the authors also recommend having blink cues in data collection blocks in order to make the task more comfortable for participants. Future studies with similar aims could use different equipment setups, e.g. electrode caps, and experimental manipulation of individual difference factors, e.g. motivation and comfort.

Abbreviations: CI – confidence interval; EEG – electroencephalography; ERP – event-related potential; ISI – interstimulus interval

Keywords: electroencephalography, event-related potential, blinking, experimental task, individual differences, noise

Introduction

Event-related potentials (ERPs) are frequently used to noninvasively study brain activity. ERPs are brain responses to specific events that are measured through electroencephalography (EEG) and that can be used to determine the time course of activation after a stimulus event (Teplan, 2002). EEG records electrical activity in the brain by measuring the sum of voltage fluctuations in

neural tissue over time, with high temporal precision. Due to this high temporal resolution, EEG is useful for the examination of neural responses to stimuli, as researchers can pinpoint changes in neural activity at very specific time points after stimulus presentation. Unfortunately, EEG is highly susceptible to various sources of noise. Technical sources include cable movements, excess or dried

electrode paste, and nearby electrical devices (Teplan, 2002; Luck, 2005). Participant-related sources include sweating, eye movements, and muscle movements (Teplan, 2002; Goncharova et al., 2003; Ma et al., 2012). Noise obscures the weak electrical signal picked up at the scalp, reducing the likelihood that effects in the data resulting from neural activity can be detected. In order to reduce the influence of noise on the neural signal, researchers can take measures to reduce environmental noise, use data processing techniques (e.g., offline filtering and independent component analysis), and remove participant-caused artifacts (i.e., instances of non-neural electrical activity) from their data (Corby and Kopell, 1972; Gevins et al., 1977; Semlitsch et al., 1986; Luck, 2005).

The low signal-to-noise ratio of ERP data requires the conduction of a large number of trials (i.e., stimulus presentations) in order to collect sufficient amounts of data, both so that noise can be averaged out of the data and so that enough trials will remain after those with artifacts have been eliminated. As a consequence, collecting ERP data can be highly time consuming and can result in a great deal of wasted effort when large amounts of data must be discarded. Therefore, it would be highly beneficial for researchers if they could minimize the number of artifacts that occur in the data, therein minimizing the number of discarded trials and the total number of trials needed. Ideally, this minimizing would occur before or during data collection, reducing the need to rely on potentially flawed analysis methods to correct for the occurrence of artifacts (see Luck, 2005 and Cohen, 2014 for discussions of potential issues with these methods). Research on the reduction of artifacts in EEG, however, overwhelmingly focuses on methods of finding artifacts in collected data and removing them from the EEG waveform using data processing techniques that factor out noise, leaving an estimate of the pure neural electrical activity and avoiding the need to eliminate all artifact-containing trials from analysis (e.g., Gevins et al., 1977; Semlitsch et al., 1986; Goncharova et al., 2003; Viola et al., 2009; Mognon et al., 2011; Ma et al., 2012; Plöchl et al., 2012; for review, see Urigüen and Garcia-Zapirain, 2015; Radüntz et al., 2015, 2017; Kilicarslan, 2016).

Thus, there is an apparent need for research investigating factors that may impact the number of artifacts in a participant's data.

Physical factors affecting the eyes could subsequently affect the quality of EEG data as both blinks and large ocular movements cause interference in EEG signals (Plöchl et al., 2012). Participants are usually asked to fixate on a screen and blink as little as possible to reduce the amount of ocular movements and blinks that occur during data collection, but participants with dry or fatigued eyes, who are wearing contacts, or who have had less sleep or caffeine may have increased difficulty with maintaining fixation on a screen and refraining from blinking and thus have more artifacts in their data. Prior research has demonstrated that participants with dry eyes do blink more frequently (Himebaugh et al., 2009), as do participants who have been deprived of sleep (Crevits et al., 2003). These physical factors could potentially be accounted and corrected for through adjustments to study methodology that would make it easier for participants to avoid blinking during data collection, such as specifying opportunities for participants to blink during blocks of stimulus presentation trials.

Individual differences between participants could also play a role in differences in data quality. For example, participants who have no prior experience with EEG may be distracted from the task at hand by the novelty of the setup. Also, participants who are less invested in participating in the study may make less of an effort to keep the number of artifacts in their data low. Prior research on individual differences in spontaneous blink rate has found great variability between individuals but has found no differences in relation to age, gender (except perhaps during reading), eye color, or the wearing of eyeglasses (Bentivoglio et al., 1997; Doughty, 2002). Blink rates have, however, been found to differ depending on current cognitive processes (for a review, see Cruz et al., 2011). For example, Bentivoglio et al. (1997) found that participants blinked less frequently when reading than when resting, and Viggiano and Mecacci (2000) found that increased need for visual attention and increased mental load decrease the rate at which participants blink. Neurological and

psychological diseases are also thought to potentially impact blink rate. Kojima et al. (2002), for instance, found that participants diagnosed with panic disorder blinked more often, suggesting that anxiety and blink rate may be connected. Thus, there may be more artifacts in the EEG data of participants with higher levels of anxiety.

In an attempt to discern what factors are connected to EEG data quality, this study examined connections between the number of trials that would have had to be eliminated from the dataset in a typical ERP study due to artifacts and a set of methodological variables, physical considerations, and individual differences. Knowledge about what may be linked to the amount of artifacts in a participant's data would hopefully allow researchers to develop research-supported ways to minimize the amount of data that must be discarded in ERP studies.

Based on prior research, we expected to find that participants with dry eyes, who had slept less, and who were more anxious would have lower quality data with more artifacts. More speculatively, we reasoned in light of previous findings that participants who felt more tired, who were more stressed, and who had spent more time looking at a screen already that day would have lower quality data with more artifacts and that participants who felt more able to focus and those forced to be more engaged in focusing on the stimuli presented to them would have higher quality with fewer artifacts. We also suspected that participants who had consumed caffeine and thus may be less tired, who had prior EEG experience, and who knew one or both investigators would have higher quality data.

Material and Methods

Participants

112 participants were recruited from Roanoke College in Salem, VA. Participants whose data had an unusually high number of artifacts were eliminated, meaning only 66 datasets were used for analysis. The 66 participants whose data was used were between 18 and 22 years old ($M = 19.15$, $SD = 1.04$). 45

were females, and 21 were males. Data was collected over the course of three semesters – more specifically during the latter half of the first semester, the whole of the second, and the first half of the third. The only weeks during which data was not collected were the first and last weeks of the semesters. Time slots were posted on Sona, the Psychology Department's sign-up system, where potential participants were able to view information about the study and register for time slots. An email was sent out to all Sona account holders each time timeslots were added to help encourage sign-ups, and the researchers asked people they knew to sign-up to recruit even more participants. The study was conducted in accordance with the guidelines of the Roanoke College Institutional Review Board. Subjects were pre-screened via questioning to ensure that they were not susceptible to seizures, and anyone who was susceptible was ineligible to participate. Subjects provided informed consent and received 1 unit of class participation credit for their time.

Equipment

This study used the same equipment setup as was used for prior EEG studies in the lab (e.g., Hurless et al., 2013; Shields et al., 2017). That is, EEG signals were recorded using a PowerLab 26T from AD Instruments, which is an instructional-level set of equipment that allows for data collection on various physiological measures, including EEG. EEG signals are transmitted from participants' scalps to the PowerLab unit using five lead shielded electrodes. Conductive paste was used to adhere electrodes to the scalp and to facilitate the transmission of electrical signals, and a headband was used to secure electrodes in place. Data collection occurred in a dark room with minimal possible distractions. Stimuli were presented using SuperLab 4.5 from Cedrus Corporation, and their temporal presentation was recorded using a StimTracker device also from Cedrus Corporation. Software was run on a Dell XPS 15z laptop, and subjects viewed all stimuli on the internal 15" widescreen monitor of the laptop with a resolution of 1366 x 768 pixels. Through LabChart 7 software from AD Instruments, the output of EEG signals was

displayed in real-time on an external 17" Dell monitor visible only to investigators.

Stimuli

A set of 80 neutral black and white images of faces from the FERET Facial Stimuli Database from the Defense Advanced Research Products Agency was used (Phillips et al., 1996; Shields et al., 2017). Face images were 384 pixels high and 256 pixels wide and were embedded in a gray (RGB value 127) rectangle. In the "blink" condition, 20 images of fingerprints were used as blink cues to give participants an opportunity to blink during a data collection block. The images of fingerprints were 383 pixels high and 254 pixels wide and had no grey border. The fingerprints appeared as black lines on a white background.

All images were presented on a grey background (RGB value 130) using SuperLab 4.5. A white cross appeared in the center of the screen for 500 ms at the beginning of each run. There was then a 600 ms pause before the presentation of the first stimulus. Images were presented for 300 ms each and were centered on the same zero point at which the cross appeared. This center point of the screen was near the faces' noses. The interstimulus interval (ISI), i.e., the time gap between the end of the presentation of one stimulus and the presentation of the next stimulus, was 1300, 1400, or 1500 ms. The ISI provides a buffer in time between stimulus presentations, reducing the likelihood a participant will still be processing a stimulus when the next is presented. The duration of each individual ISI was randomly set using the SuperLab software so that ERPs would not incorrectly reflect a constant phase on brain activity occurring on a repetitive time cycle.

Survey

Participants filled out a self-administered survey that contained questions about individual difference variables. Most of the survey was closed response. The aim of the survey was to find which factors affect performance during EEG studies. Variables participants were asked about included year in college, reason for participating, course for which credit would be awarded if applicable, use

of corrective lenses (glasses and/or contacts), preexisting eye conditions, and experience with EEGs (see Appendix for the full survey). Participants were also asked to rate the level of anxiety, motivation, tiredness, stress, and ability to focus that they experienced at the beginning and at the end of data collection on a scale of 1, "Not at all," to 7, "Extremely." They were asked to retroactively rate how they felt at the beginning of data collection, as we were interested in how they were feeling at that time, but did not want to introduce a possible confound by giving them a survey asking them to think about their current state of being before the EEG data was collected. We instead elected to more closely follow the procedures of prior EEG studies in the lab (e.g., Hurless et al., 2013; Shields et al., 2017), which did not involve giving surveys to participants before data collection, and instead only give them a survey after EEG data collection was complete.

The physical considerations selected for analysis were use of corrective lenses, whether participants' eyes were dry, whether participants had consumed caffeine that day, hours of sleep, and hours of screen time. The individual difference variables selected for analysis were experience with EEG, whether participants knew one or both investigators, and the levels of the five affective measures that they reported feeling at the beginning and end of data collection.

Procedure

Prior to the beginning of the experiment, participants were randomly assigned to one of two block orders to avoid confounds resulting from individual differences. Avoiding such confounds was particularly important for this study, as we sought to investigate possible connections between individual differences and data quality, which would not have been possible if there were systematic differences between the participants assigned to one block order and those assigned to the other. The block order assignment determined whether participants completed the "no blink" or "blink" portion of the experiment first. In the "no blink" condition, participants viewed the 80 images of faces and were asked not to blink for the duration of the block (approximately 2.5 minutes). In the "blink" condition, participants

were shown the images of faces and 20 blink cues, and they were asked to blink only when they saw a blink cue. Each “blink” block lasted approximately 3 minutes.

Participants were prescreened and were not eligible to participate if they were susceptible to seizures, had metal in their heads, or did not have normal or corrected-to-normal vision. They were given an informed consent form and were warned that the study’s procedures could potentially cause migraines. After giving informed consent, participants abraded the skin on their foreheads with abrasive gel on a cotton round and cleaned their skin with alcohol swabs before the electrodes were secured in place.

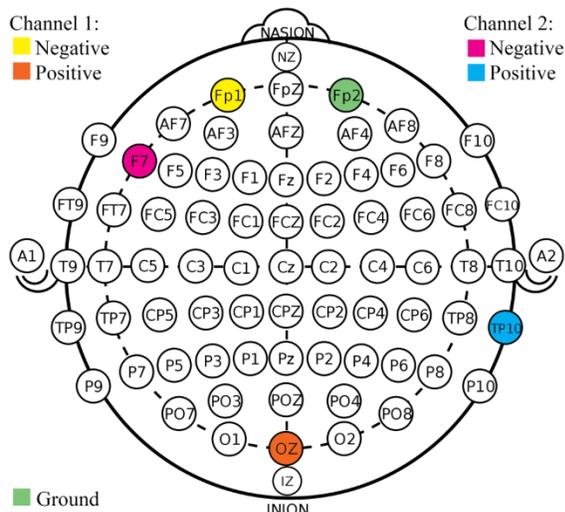


Figure 1. Diagram of electrode placement.

The ground electrode was placed on the right side of the forehead about halfway between the eyebrows and hairline (FP2; Figure 1). The channel 1 electrodes were placed on the left side of the forehead about halfway between the eyebrows and hairline (FP1) and approximately an inch above the inion (the bump on the back of the head; OZ). The channel 2 electrodes were placed on the left temple (F7) and about an inch behind and above the right ear (TP10), and a headband was wrapped around the forehead and back of the head to hold the electrodes in place. These locations were chosen to record neural activity from specific brain areas relevant for Shields et al. (2017). FP1 and FP2 are located over the frontalis muscle, F7 and TP10 are

located over the left and right temporalis muscles, respectively, and OZ is located between the left and right occipitalis muscles. While all electrode locations are susceptible to interference from muscular activity, research has shown that muscular interference is most likely to occur at peripheral locations over or adjacent to the contracting muscle (Goncharova et al., 2003; Ma et al., 2012), meaning that all locations used for this study are particularly susceptible to muscular interference.

Participants were asked to sit in a relaxed position with their chins on a pillow placed on a stack of books in front of them, and all lights were turned off. The investigator opened the LabChart 7 software and evaluated the incoming electrical signals to check that equipment was properly set up and functioning. Investigators checked to make sure that signals seemed to be generally oscillating between ± 50 microvolts with peaks moving outside of the ± 60 microvolt range when participants blinked, which indicates proper recording of eye muscle activity. It was also confirmed that there were no obvious abnormalities in the incoming signal, as abnormalities could indicate a problem with the conductivity between the scalp and the electrodes.

Once the investigator had tested the electrode connections, participants were instructed to remain as still as possible during data collection blocks to reduce participant-caused artifacts. Investigators gave specific examples of potential sources of artifacts, asking participants to refrain from fidgeting, clenching the jaw, raising the eyebrows, making facial expressions, moving ears, drumming fingers, tapping feet, tugging on the wires to the electrodes, etc. Four blocks were conducted with the investigators in the room monitoring incoming data and participant behavior. The first two blocks were of the initial condition to which participants were assigned, i.e. with blinks cues or without, and the final two were of the other condition. Participants were always given a chance to rest their eyes between blocks if they needed to do so. Some participants stated that they did not need a break and simply wanted to continue, whereas others needed a chance to rest their eyes because of eye fatigue or dryness. If any major issues arose during a block (e.g., the

SuperLab 4.5 software crashed, an eyelash fell into the participant's eye, an electrode fell off), data collection was stopped, the data for that block was deleted, and data collection for that block was restarted. Blocks had to be restarted for less than 10% of participants.

After four complete blocks of data collection, electrodes were removed, and electrode paste was wiped away. Participants were given the survey to complete while investigators cleaned the equipment and ensured all data was properly saved. They were then debriefed and told about the objectives for the current study. All in all, this study took 25-35 minutes for each participant: approximately 5 minutes for initial set-up, about 15-20 minutes to run through the stimuli and conduct the EEG, and about 5-10 minutes for removal of electrodes and electrode paste, completion of the survey, and debriefing.

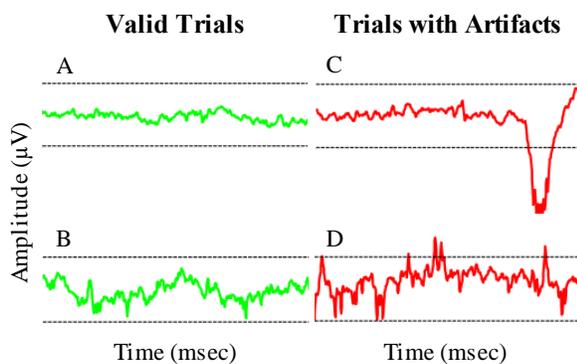


Figure 2. Example ERPs. A) Clean data with little noise. B) Clean data with medium noise. C) Artifact due to a blink. D) Artifact due to too much noise. A and B are examples of valid trials that would have been used in ERP analysis. C and D are examples of trials that would have been eliminated from ERP analysis.

Analysis

The number of trials in participants' data containing artifacts in channel 1 was used as a measure of data quality, following the precedent of Shields et al. (2017) which focused on analyzing ERPs in channel 1 data. Though these numbers are not exactly equal to the number of artifacts since multiple artifacts can occur within one trial, the number of trials that would need to be eliminated from further analysis in typical ERP studies is a more relevant measure of data quality than the exact

number of artifacts. A trial was marked as containing an artifact if the voltage in the trial went beyond ± 60 microvolts away from baseline at any point in time up to 1200 ms after stimulus onset, which was the minimum ISI. Baseline voltages were set as the average voltage of the 300 ms before stimulus presentation. Figure 2 displays examples of trials marked as valid (i.e., trials that did not exceed the ± 60 microvolt cutoff range and thus likely did not contain any artifacts; A and B) and trials containing artifacts (C and D).

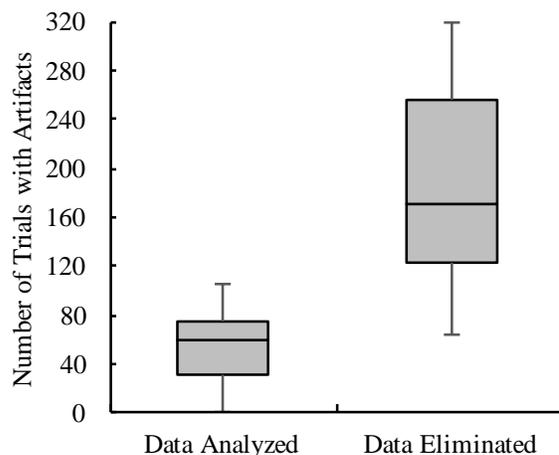


Figure 3. Box plot depicting the distribution of the total number of trials containing artifacts in participant data included in analyses ($n = 66$) and data eliminated from analyses ($n = 46$). Error bars indicate the minimum and maximum values. The boxes represent the middle 50% of the data, with the lines through the boxes denoting the medians.

After the number of trials with artifacts was calculated, participants with extremely high numbers of trials containing artifacts were eliminated from subsequent analyses as it could not be confirmed whether the high number of artifacts was due to excessive blinking or movement, to improper setup of the equipment, to equipment malfunctions, or some other unknown extraneous factor. Therefore, sound conclusions about potential factors that led to those participants having more trials excluded than other participants could not confidently be drawn. Participants were excluded if more than half of their trials across blocks (i.e., more than 160 trials out of 320 total trials) or all trials in a single block contained artifacts. Overall, there was a noticeable gap in the total number of

artifacts per participant in data included in analyses ($M = 54.85$, $SD = 26.29$, range: 1-105) and eliminated from analyses ($M = 188.33$, $SD = 81.61$, range: 65-320; Figure 3). 84.78% of eliminated datasets had more trials containing artifacts than any of the analyzed datasets. The discrepancy between the number of trials with artifacts in the analyzed and eliminated datasets supports the possibility that an extraneous factor, such as an issue with the equipment, rather than one of the factors of interest in this study contributed to the low data quality, justifying the use of the stated exclusion criteria. A total of 66 participants were used in subsequent analyses, 33 who completed the “blink” condition first and 33 who completed the “no blink” condition first.

The Fisher z -transformation was applied to observed correlation coefficients in order to calculate 95% confidence intervals (CIs) so that judgements about differences in correlations and the strength of correlations could be made. For correlation coefficients, common standards are $r = \pm 0.1$ for small effect sizes, $r = \pm 0.3$ for medium, and $r = \pm 0.5$ for large (Cohen, 1988). r^2 is the amount of variance explained, so for this study, $|r| = 0.316 - 0.5$ was defined as the range for moderate correlation values because ± 0.316 is the Pearson correlation coefficient for which 10% of variance is explained and ± 0.5 is the r value for which 25% of variance is explained. Thus, correlations can be classified as less than moderate if the upper and lower bounds of the 95% CI fall between ± 0.316 . A correlation can be classified as at least moderate if the upper and lower bounds fall outside of the ± 0.316 range (i.e., the upper bound is below -0.316 or the lower bound is above 0.316) and as greater than moderate if the upper and lower bounds fall outside of the ± 0.5 range.

Results

Task Structure

A mixed-design ANOVA with a within-subjects factor of block number and a between-subjects factor of block order (i.e., whether participants were assigned to first complete two “blink” blocks or two “no blink” blocks) was performed in order to test for the effect of giving

participants designated opportunities to blink during blocks on the number of artifacts detected in experimental trials. Mauchly's sphericity test was significant, $X^2(5, N = 66) = 30.359$, $p < 0.001$, so sphericity could not be assumed. The lower-bound correction was used to be conservative. There was a significant effect of block, $F(1,64) = 9.624$, $p = 0.003$, but there was no significant order effect, $F(1,64) = 1.102$, $p = 0.298$, nor was there a significant interaction between block number and block order, $F(1,64) = 0.858$, $p = 0.358$ (Table 1; Figure 4). Because there was no significant order effect, the number of artifacts across all four blocks and across both block orders was used as the dependent variable in subsequent analyses of factors affecting artifact rates.

Table 1. Average number of artifacts in a particular block.

	“Blink” 1st		“No Blink” 1st	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Block 1	26.67	17.53	30.58	14.62
Block 2	21.12	13.41	23.61	12.30
Block 3	18.91	11.62	17.55	9.18
Block 4	19.91	13.18	24.15	12.69

Cells in grey represent numbers for blocks of the “no blink” condition.

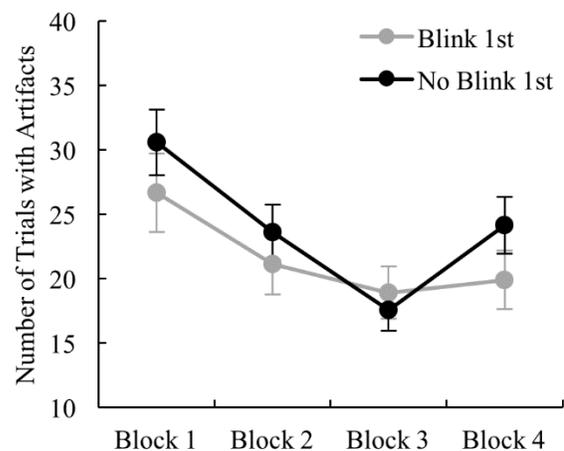


Figure 4. Number of trials containing artifacts for each of the four blocks separated by which condition participants started with. Error bars show one standard error of the mean.

Though the number of artifacts in a block did not significantly differ based on condition or the order of conditions, there were significantly more artifacts in block 1 ($M = 28.62$, $SD = 16.13$) than in the three other

blocks, all p 's ≤ 0.010 (Figure 4). Furthermore, there were significantly fewer artifacts in block 3 ($M = 18.23$, $SD = 10.42$) than in the other three blocks, all p 's ≤ 0.013 . The number of artifacts in block 2 ($M = 22.36$, $SD = 12.83$) and block 4 ($M = 22.03$, $SD = 13.01$) was not significantly different, $p = 0.860$.

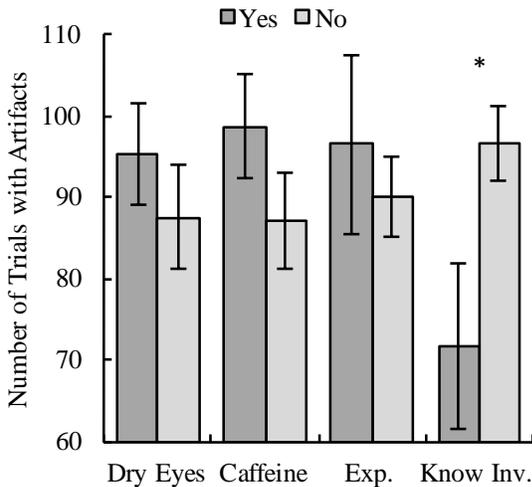


Figure 5. Total number of trials containing artifacts for each participant, separated by whether participants answered affirmatively or negatively when asked whether their eyes were dry, whether they had had caffeine that day, whether they had any prior experience with EEGs, and whether they knew one or both of the investigators. Error bars show one standard error of the mean. Asterisk indicates significant difference at the $p < 0.05$ level.

Physical Considerations

A one-way between-subjects ANOVA was run to test for differences in the amount of artifacts in participants' data based on whether they were wearing glasses, were wearing contacts, were not wearing corrective lenses at the time, or never wear corrective lenses. No significant difference was found, $F(3,62) = 1.232$, $p = 0.306$. Independent samples t -tests revealed that there were also no significant differences between participants who reported having dry eyes ($n = 32$, $M = 95.16$, $SD = 36.91$) and participants who reported that their eyes were not dry ($n = 34$, $M = 87.56$, $SD = 36.05$), $t(64) = -0.857$, $p = 0.395$, nor between participants who reported having consumed caffeine that day ($n = 24$, $M = 98.67$, $SD = 30.90$) and participants who had not consumed caffeine that day ($n = 42$, $M = 87.00$, $SD = 38.18$), $t(64) = -1.276$, $p = 0.207$ (Figure 5).

Furthermore, there were no significant correlations between the number of artifacts and reported hours of sleep, $r(64) = -0.026$, $p = 0.839$, or reported hours of screen time, $r(64) = 0.042$, $p = 0.737$ (Figure 6). 95% CIs show that both of these correlations are less than moderate.

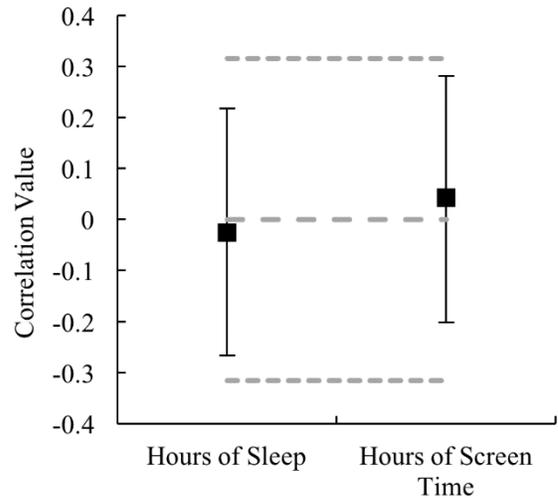


Figure 6. Correlations between the total number of trials containing artifacts per participant and their reported hours of sleep as well as their reported hours of screen time. Error bars represent the upper and lower bounds of the 95% CIs. The grey dashed line with wider dashes marks $r = 0$, and the grey dashed lines with thinner dashes mark $r = \pm 0.316$, used here as the cutoffs for classifying a correlation as less than moderate.

Individual Differences

Participants who reported having prior experience with EEG ($n = 13$, $M = 96.46$, $SD = 39.47$) did not have significantly more artifacts in their data than participants who reported having no prior experience ($n = 53$, $M = 89.96$, $SD = 35.27$), $t(64) = 0.582$, $p = 0.563$ (Figure 5). Participants who reported knowing one or both of the investigators ($n = 14$, $M = 71.64$, $SD = 38.29$) did have significantly fewer artifacts in their data than those who did not report knowing one of the investigators ($n = 52$, $M = 96.52$, $SD = 33.70$), $t(64) = 2.382$, $p = 0.020$ (Figure 5).

Participants rated the levels of anxiety, motivation, tiredness, stress, and ability to focus that they experienced at the beginning of data collection and at the end of data collection. On each of the five affective measures, the feelings participants reported they had at the beginning and end of data collection were highly

correlated, all r 's > 0.574 , p 's < 0.001 . None of these measures were significantly correlated with the number of artifacts, all p 's > 0.125 . Correlations with how participants reported they felt at the beginning of data collection are as follows: anxiety, $r(64) = -0.098$, $p = 0.432$; motivation, $r(64) = -0.047$, $p = 0.709$; tiredness, $r(64) = 0.094$, $p = 0.453$; stress, $r(64) = 0.088$, $p = 0.484$; ability to focus, $r(64) = -0.141$, $p = 0.259$ (Figure 7). The correlations between the number of artifacts and how participants reported they felt at the end of data collection are as follows: anxiety, $r(64) = -0.083$, $p = 0.508$; motivation, $r(64) = -0.051$, $p = 0.686$; tiredness, $r(64) = -0.064$, $p = 0.609$; stress, $r(64) = 0.033$, $p = 0.792$; ability to focus, $r(64) = -0.190$, $p = 0.126$. For both sets of affective reports, the correlations between the number of artifacts and participants' level of motivation were less than moderate (Figure 7). The correlations between the total number of trials containing artifacts per participant and participants' levels of tiredness and stress at the end of data collection were also less than moderate. The correlations with the other affective measures could not be conclusively judged as moderate or less than moderate but were not greater than moderate.

Number of Artifacts in Channel 2

As a validity check, the same statistical tests were run on data from channel 2 after additionally excluding participants if more than half of their trials in channel 2 across blocks or all trials in channel 2 in a single block contained artifacts. 56 participants remained, 29 who completed the "blink" condition first and 27 who completed the "no blink" condition first. Results were qualitatively similar in that the main finding was a temporal order effect, $F(1,54) = 19.348$, $p < 0.001$, with block 1 containing significantly more artifacts than the other three blocks, all p 's < 0.001 . No other tests or correlations were significant.

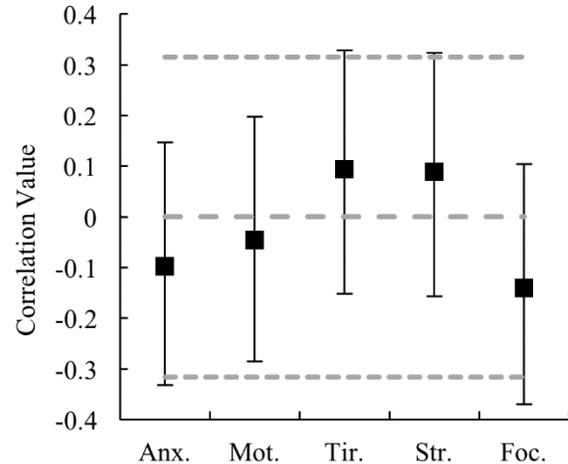


Figure 7. Correlations between the total number of trials containing artifacts per participant and how participants reported they felt at the beginning of data collection for each of the five affective measures: anxiety, motivation, tiredness, stress, and ability to focus. Error bars represent the upper and lower bounds of the 95% CIs. The grey dashed line with wider dashes marks $r = 0$, and the grey dashed lines with thinner dashes mark $r = \pm 0.316$, used here as the cutoffs for classifying a correlation as less than moderate.

Discussion

Minimizing the occurrence of artifacts in data is key for ERP researchers as ERP data has a low signal-to-noise ratio, requiring the conduction of large numbers of trials so that there will be enough data to isolate ERPs from EEG noise (Luck, 2005). This study attempted to uncover what factors are connected to EEG data quality by examining connections between the number of trials containing artifacts and a set of methodological variables, physical considerations, and individual differences with the goal of helping researchers develop research-supported ways to minimize the amount of data that must be discarded in ERP studies.

However, there was no significant difference between the number of artifacts observed in blocks during which participants were given specific opportunities to blink and blocks during which participants were asked not to blink. Additionally, there was not a significant effect of the order in which participants completed the two "blink" and the two "no blink" blocks. Therefore, quantitatively, giving participants the opportunity to blink during

blocks does not seem to make a difference in EEG studies in which they are being asked to view stimuli. However, participants tended to respond negatively when asked to blink as little as possible during the blocks, making unsolicited negative comments to investigators before and after blocks and giving negative feedback about that part of the procedure during debriefing. In order to ease the minds of participants and to help them focus on viewing stimuli, it may thus be best for researchers to build blink trials into their blocks. Having those trials forces participants to pay attention to the stimuli rather than merely focusing on avoiding blinking, particularly when images cueing blinking are randomly spaced in time so that participants cannot predict when they will appear.

Block 1 contained significantly more artifacts than any other block, suggesting that the first block of data collection is likely to have the lowest amount of quality data. This may be because participants were distracted from the task at hand by the novelty of the setup. They also could have been uncomfortable and unable to relax until they had experienced a full data collection block. As a result, researchers may benefit from having a practice block for participants to complete in order to minimize the number of artifacts in actual data collection blocks.

The number of artifacts in block 3, when participants transitioned to a new stimulus presentation condition, was significantly lower than in all other blocks. This could be due to the switching of condition or due to improvement over time in maintaining stillness and reducing blinking. The relative increase in artifacts in block 4 could suggest that the lower number of artifacts in block 3 resulted from the switch in condition rather than improvement over time, but it could also have taken place due to fatiguing of participants' eyes or even due to their knowledge that data collection was almost over. Without additional blocks and condition switches, these options cannot be differentiated.

There were no significant findings in relation to the physical considerations examined, and the only significant individual difference was that participants who knew one or both investigators, ranging from close friends of

investigators to mere acquaintances, had significantly fewer trials with artifacts. This suggests that, if possible, researchers may benefit from personally recruiting people that they know to participate in their studies.

This difference in data quality could have resulted from differences in motivation as participants who knew one or both investigators may have felt more motivated to do what they could to ensure the quality of their data due to their personal connection to the research team. No differences in motivation levels were found between these groups, and the correlation between number of artifacts and motivation was less than moderate, seemingly refuting the possibility that there was a difference in motivation. However, participants were not specifically asked about how motivated they felt to do their best in participating in the study and were likely answering based more on what they felt their general level of motivation was at the time as the measure was grouped with other general affective measures. Thus, this lack of significant difference and lack of correlation cannot be taken as evidence against an influence of motivation to do well on the particular task.

Furthermore, participant motivation was measured through self-report in this study. Self-report measures are inherently less reliable than objective measures, as there is no guarantee that participants are responding honestly or are interpreting the question in the intended way. Additionally, for this study, participants were asked to retroactively rate how they were feeling at the beginning of the session on the survey given after EEG data collection. They were only surveyed at the end of data collection to avoid the possibility that completing a survey and having to think about their state of being would somehow affect their performance during the EEG, but that inherently makes those reports even more unreliable. Future research could examine whether attempts to increase participant motivation, perhaps by giving participants motivating statements before the beginning of data collection, increase data quality.

The difference between the number of trials containing artifacts in the data of participants who knew investigators and those who did not could also have resulted from a difference in comfort levels. Participants could

have felt more relaxed and thus had an easier time focusing on the task while attempting to remain still and physically relaxed. This interpretation aligns with the possibility that the number of trials with artifacts decreases after block 1 because participants get more comfortable. More research would be needed to determine whether comfort levels really are the source of the difference. Studies could additionally investigate whether there is a systematic difference in the behavior of researchers and participants when they know each other as compared to when they do not.

The lack of significant findings with regard to the physical considerations and individual differences examined suggests that there is a network of individual factors at play affecting the quality of participants' EEG data, rather than one factor or a few central factors. That seems to be the case with spontaneous blink rates, which would be expected to correlate with EEG data quality (Bentivoglio et al., 1997; Doughty, 2002; Kojima et al., 2002; Crevits et al., 2003; Himebaugh et al., 2009; Cruz et al., 2011). As such, there may be no overarching way for researchers to screen participants to predict what the quality of their data will be and subsequently take measures to minimize the occurrence of artifacts. Research comparing the quality of data across groups who experience an intervention and those who do not would be needed to confirm whether any methodological manipulation could successfully decrease the number of artifacts in EEG data.

Furthermore, as the electrode locations used for this study are particularly susceptible to muscular interference, this dataset likely contains more artifacts than it would have if other electrode locations had been chosen. However, these locations were specifically selected for Shields et al. (2017), the procedures of which are mimicked in this study, to record neural activity from brain areas relevant for the project's research question. EEG studies of neural activity cannot always avoid the possibility of additional muscular interference by changing electrode placement locations, and thus, different locations were not selected for this study either. It would be interesting though to investigate whether a similar study to this one, with electrodes placed in locations less

susceptible to muscular interference, would find additional significant differences.

Moreover, studies using alternate EEG equipment or setups may find different results in regard to what factors may impact data quality. For instance, electrode caps may more securely hold electrodes in place, reducing the number of non-participant-related artifacts in data and potentially elucidating significant differences in data quality resulting from methodological variables, physical considerations, and individual differences. In a study comparing individual electrodes and electrode caps, participants did not rate caps as being more comfortable during data collection (Shields et al., 2016), so comfort would not likely be a confound that could unintentionally increase data quality. Different results may also be obtained if electrical impedance between electrodes and the skin is checked and minimized before data collection, which we were not able to do for this study. Perhaps taking that step to try to reduce the noise in the electrical signal would also elucidate effects not detected in this research. By studying variability in data quality across tasks, physical factors, individual difference variables, and equipment setups, researchers will be able to gain a fuller picture of what could be impacting the quality of their EEG data and what they could do to better that quality before and during data collection.

Acknowledgements

Thank you to the Roanoke College Psychology Department for providing the equipment, materials, and facilities needed for this study.

Corresponding Author

Stephanie Shields
Roanoke College
221 College Lane
Salem, VA 24153
smshields@mail.roanoke.edu

References

- Bentivoglio AR, Bressman SB, Cassetta E, Carretta D, Tonali P, Albanese A (1997) Analysis of blink rate patterns in normal subjects. *Mov Disord* 12:1028-1034.
- Cohen J (1988) *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cohen MX (2014) *Analyzing neural time series data: Theory and practice*. Cambridge, MA: MIT Press.
- Corby JC, Kopell BS (1972) Differential contributions of blinks and vertical eye movements as artifacts in EEG recording. *Psychophysiology* 9(6):640-644.
- Crevits L, Simons B, Wildenbeest J (2003) Effect of sleep deprivation on saccades and eyelid blinking. *Eur Neurol* 50(3):176-180.
- Cruz A, Garcia DM, Pinto CT, Cechetti SP (2011) Spontaneous eyeblink activity. *Ocul Surf* 9(1):29-41.
- Doughty MJ (2002) Further assessment of gender- and blink pattern-related differences in the spontaneous eyeblink activity in primary gaze in young adult humans. *Optometry Vision Sci* 79(7):439-447.
- Gevins AS, Yeager CL, Zeitlin GM, Ancoli S, Dedon MF (1977) On-line computer rejection of EEG artifact. *Electroen Clin Neuro* 42:267-274.
- Goncharova II, McFarland DJ, Vaughan TM, Wolpaw JR (2003) EMG contamination of EEG: Spectral and topographical characteristics. *Clin Neurophysiol* 114(9):1580-1593.
- Himebaugh NL, Begley CG, Bradley A, Wilkinson JA (2009) Blinking and tear break-up during four visual tasks. *Optometry Vision Sci* 86(2):E106-E114.
- Hurlless N, Metic A, Pena S, Humphries E, Gentry H, Nichols DF (2013) Music genre preference and tempo alter alpha and beta waves in human non-musicians. *Impulse* 1-11.
- Kilicarslan A, Grossman RG, Contreras-Vidal JL (2016) A robust adaptive denoising framework for real-time artifact removal in scalp EEG measurements. *J Neural Eng* 13(2):026013.
- Kojima M, Shioiri T, Hosoki T, Sakai M, Bando T, Someya T (2002) Blink rate variability in patients with panic disorder: New trial using audiovisual stimulation. *Psychiatry Clin Neurosci* 56(5):545-549.
- Luck SJ (2005) *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.
- Ma J, Tao P, Bayram S, Svetnik V (2012) Muscle artifacts in multichannel EEG: Characteristics and reduction. *Clin Neurophysiol* 123(8):1676-1686.
- Mognon A, Jovicich J, Bruzzone L, Buiatti M (2011) ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology* 48(2):229-240.
- Phillips PJ, Raus PJ, Der SZ (1996) FERET (face recognition technology) recognition algorithm development and test results. Army Res Lab.
- Plöchl M, Ossandón JP, König P (2012) Combining EEG and eye tracking: Identification, characterization, and correction of eye movement artifacts in electroencephalographic data. *Front Hum Neurosci* 6(278):1-23.
- Radüntz T, Scouten J, Hochmuth O, Meffert B (2015) EEG artifact elimination by extraction of ICA-component features using image processing algorithms. *J Neurosci Methods* 243:84-93.
- Radüntz T, Scouten J, Hochmuth O, Meffert B (2017) Automated EEG artifact elimination by applying machine learning algorithms to ICA-based features. *J Neural Eng* 14(4):046004.
- Semlitsch HV, Anderer P, Schuster P, Presslich O (1986) A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. *Psychophysiology* 23(6):695-703.
- Shields SM, Morse CE, Applebaugh ED, Muntz TL, Nichols DF (2016) Are electrode caps worth the investment? An evaluation of EEG methods in undergraduate neuroscience laboratory courses and research. *J Undergrad Neurosci Educ* 15(1):A29-A37.
- Shields SM, Morse CE, Arrington P, Nichols DF (2017) Exploration of the influence of smiling on initial reactions across levels of

- facial attractiveness. *Am J Undergrad Res* 14(2):65-79.
- Teplan, M (2002) Fundamentals of EEG measurement. *Meas Sci Rev* 2(2):1-11.
- Urigüen JA, Garcia-Zapirain B (2015) EEG artifact removal — State-of-the-art and guidelines. *J Neural Eng* 12(3):031001.
- Viggiano MP, Mecacci L (2000) Identification of Poppelreuter-like pictures as indexed by blinking. *Neuropsychologia* 38(7):1098-1100.
- Viola FC, Thorne J, Edmonds B, Schneider T, Eichele T, Debener S (2009) Semi-automatic identification of independent components representing EEG artifact. *Clin Neurophysiol* 120(5):868-877.

Appendix

Following are the questions from the survey completed by participants.

- 1) How old are you?
 - 2) Year in college
a. 1 b. 2 c. 3 d. 4 e. 5 f. ≥ 6
 - 3) What is your major/minor/concentration?
 - 4) Are you interested in or have you declared the neuroscience concentration?
Yes / No
 - 5) Why are you participating in this study?
Circle all that apply.
 - a. To receive required class credit (If so, which course?)
 - b. To receive extra credit (If so, which course?)
 - c. My professor suggested I participate
 - d. I enjoy participating in studies
 - e. I am interested in research
 - f. I know one of the investigators
 - g. Other
 - 6) Do you use corrective lenses? Yes / No
If so, circle which kind you are currently wearing: Glasses Contacts None
If you are wearing contacts, how many hours have they been in?
 - 7) Have you had Lasik eye surgery? Yes / No
 - 8) Have you ever had any eye surgery, excluding Lasik procedures? Yes / No
If so, what for?
 - 9) Do you have any preexisting eye conditions? Yes / No (e.g. eye twitching, strabismus, glaucoma)
If so, what are they?
 - 10) Are your eyes currently dry? Yes / No
 - 11) Do you have any allergies that are currently affecting your eyes? Yes / No
 - 12) Do you have any prior experience with EEGs? (Circle all that apply)
 - a. Yes, I have had one conducted on me _____ time(s).
 - b. Yes, I have conducted one before.
 - c. Yes, I have seen one be conducted before.
 - d. No, I do not.
 - 13) How many hours of sleep did you get last night?
 - 14) Have you had caffeine today? Yes / No
 - 15) Approximately how many hours of screen time have you had today (i.e., time looking at a phone, computer, tablet, or TV)?
 - 16) Have you had a migraine within the last year? Yes / No
Are you on medication for migraines?
Yes / No
 - 17) Have you had a seizure within the last year? Yes / No
Are you on medication for seizures?
Yes / No
 - 18) Do you have metal in your head? Yes / No
- For the following questions, please rate on a scale from 1 (Not at all) to 7 (Extremely) how you felt when we were beginning the EEG recordings:
- 19) Anxious
 - 20) Motivated
 - 21) Tired
 - 22) Stressed

23) Able to focus

For the following questions, please rate on a scale from 1 (Not at all) to 7 (Extremely) how you felt during the last EEG recording session:

24) Anxious

25) Motivated

26) Tired

27) Stressed

28) Able to focus