Are Reading Assessment Scores and School Intervention Predictive of Left Hemisphere N170 EEG Amplitude?

Introduction

The ability to read affectively is crucial for doing well in academics and it often translates to job success. However, there are plenty of elementary school students who struggle with reading, which builds up as they advance in their academics. This could be due to an inherent reading disability such as dyslexia, the lack of proper training in school, or the lack of resources at home. Language acquisition is found to be processed in the left hemisphere temporal lobe of the brain (Friederici, 2011). More specifically, word processing can be detected by an electroencephalography (EEG) component called the N170, which is a negative voltage deflection in the EEG waveform detected at 170 milliseconds (ms) after the onset of a word stimulus (Sanchez-Vincitore, Avery & Froud, 2017). Hence, the left hemisphere N170 negativity is known to be a marker for word processing automaticity (Maher, Brandeis & McCandliss, 2005). Moreover, it is observed that as reading skills advance, the more positive the N170 amplitude, possibly suggesting greater reading efficiency (Froud et al., 2019). However, reading skill is a general term and such language acquisition and decoding ability can be broken down into many components such as phonological awareness, reading fluency, sight word efficiency, and reading comprehension. Hence, the aim of this project is to find out which skill components, as represented by assessment scores of each component measure, may have a significant correlation with the N170 amplitude of the left hemisphere when participants are reading words. Also of interest is to see whether school intervention plays a role on top of assessment scores when correlating with amplitude measures. Of the four major assessments, scores on the Phonics Inventory, from here on referred to as “phonics scores”, is found to have a significant positive correlation with the N170 EEG amplitude. Moreover, there is a main effect of school intervention on top of this correlation, albeit insignificant. Those who received interventions had relatively more positive amplitudes than those who did not. These findings suggest that reading fluency, as measured by the Phonics Inventory may be an important component in the development of reading skills at the neurological level, and that school interventions in general may have a positive impact on reading ability.

Data Description

Data for this project are acquired from the Neurocognition of Language (NCL) Lab at Teachers College, Columbia University. The NCL lab recruited 21 students ranging from 10-12 years of age and with different levels of reading abilities. These participants were sampled by random and are thus independent of other observations. Each participant was given various assessments on reading ability, each reflecting a specific component of reading skill. The four assessment scores of interests are derived from the Reading Inventory, Phonics Inventory, Comprehensive Test of Phonological Processing (CTOPP), and the Test of Word Reading Efficiency (TOWRE) assessments. The Reading Inventory is a universal screener and growth monitor for grades K-12 that focuses solely on reading comprehension and readability as measured by the Lexile (Wagner, 2005). The Phonics Inventory is a foundational reading assessment for grades 3-12 and it measures the accuracy and fluency with which students identify individual letters and words (Wagner, 2009). CTOPP scores provide a normed measure of phonological awareness (Wagner, Torgesen, Rashotte & Pearson, 2013). Lastly, the TOWRE assessment looks at students' ability to efficiently sight read (Torgesen, Wagner & Rashotte, 2012). Descriptive statistics for all the assessment scores are listed in Table 1.

All participants also had their EEG amplitude measurements taken while reading a series of words that appeared on a screen. The words were either low frequency or high frequency words. Low frequency words are words that are rarely encountered in our daily lives, whereas high frequency words are those that appear often. To reduce the confound of word familiarity, this project will only use data on the high frequency words. In other words, participants who are advanced readers may encounter certain low frequency words more often than do those who struggle with reading disabilities. Whereas, the word familiarity of high frequency words is relatively similar across participants.

EEG data were recorded at 129 electrodes, each placed at a different location on the scalp. There are seven electrodes of interest, which represent measurements of the left hemisphere. A total of 32 high frequency words were presented to each participant. Each electrode captured EEG data from 100ms before the stimulus (word) onset and 400ms after the stimulus onset, sampled at a rate of 2ms. However, the time frame of interest is mainly the window that surrounds the N170, which is a negativity dip at around 170ms into the stimulus onset. The problem with capturing the N170 time frame is that each participant may have a slightly different time frame for this negativity dip. For example, participant A can have a dip at 165ms after stimulus onset whereas participant B may have a dip at 175ms after stimulus onset. Therefore, data from a time window subset of 150ms to 220ms after stimulus onset, is first extracted. Then, all the EEG amplitudes across all seven electrodes of interest are averaged across each 2ms time sample within the subset time window. Then, the minimum amplitude within the subset time window is identified for each participant, as their negativity dip. Lastly, the adaptive mean of each participant’s negativity dip is calculated, yielding a single representative N170 amplitude for each participant. Descriptive statistics for amplitude are also listed in Table 1.

Each participants' parent was given a survey to fill out, of which one of the questions asks if the participant received any intervention(s) in school. The answer to the question is either a "yes" or a "no" choice. In this data set, "yes" is coded as one and "no" as zero. The number of participants in each group are shown in Table 2.

Methods

First, the normality of the amplitude is checked using qqPlots for normality with a 95% confidence envelope. If all the points lie within the confidence interval, then the variable is normally distributed. If, however there appears to be any points that lie outside the confidence interval, then a Shapiro-Wilks test is conducted for the variable to see if p is greater than the significance level ($\alpha = 0.05$), with which the variable would still be considered as normally distributed.

A linear regression model in the form of $Y_i = \beta_0 + \beta_1 X_1 + \epsilon_i$ is fit to model amplitude conditional on each assessment scores. These scores are also plotted against amplitude measurements, with the addition of their corresponding linear regression line. From looking at the multiple squares and p of each regression model as well as from the visual scanning of the correlation plots, each regression model either shows a significant correlation between assessment scores and amplitude or not. Any models with a p above the significance level ($\alpha = 0.05$) is considered as insignificant.

In the case that a significant correlation is found, the categorical variable of school intervention is added into the equation if the difference in score means for participants who received and did not receive intervention differ, as shown in Figure 3. The two independent variables are then plotted against amplitude. Next, a linear model in the form of $Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon_i$ is used to model amplitude conditional on assessment scores and intervention to see if the addition of school intervention can improve the previous model. Then, the presence of any an interaction between assessment scores and school intervention are checked by looking at the residuals of a linear model that includes the interaction. This linear equation, $Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \epsilon_i$, models amplitude conditional on the assessment score of interest, school intervention, and their interaction. The coefficients of this model are used to plot the interaction. Both the qqPlot and residualPlot are applied to the linear model of choice, to check for residual normality and linearity assumptions. If the model seems to fit well for the variables, then no further complication of the model is needed. If, however the fitted values seem to deviate from the Pearson residuals, then another model should be considered. Nevertheless, to triple check that a linear model is the best fit model, a quadratic regression model that considers the addition of a quadratic relationship between phonics scores and amplitude is assessed. The quadratic model in the form of $Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_1^2 + \epsilon_i$ is fit to model amplitude conditional on assessment score, school intervention, interaction between the two, and the quadratic component of assessment score. To finally determine which model fits best, ANOVA is applied to all four models, with the most complex model nested in the second most complex model and so on until the least complex model nests all the other models. Lastly, the significance of each model is reviewed.

Results

Amplitude and CTOPP scores are not significantly correlated, r = 0.130, p = 0.577. However, amplitude and phonics scores are significantly correlated, r = 0,476, p = 0.029. The combination of school intervention and phonics scores also significantly predict amplitude measurements, as shown in Table 3. The presence of school intervention has a main effect on amplitude. Those who received school intervention has a less negative amplitude measurement when compared to those who did not receive school intervention, as shown in Figure 1. However, such main effect is not significant, as shown in Table 3, where the nested model of the school intervention does not significantly improve the model with just the phonics scores. From looking at Figure 1 alone, the interaction between school intervention and phonics scores is minimal, which is further confirmed in Table 3. Nevertheless, the addition of interaction slightly improves the model, though very insignificantly. To confirm that the interaction model is a good fit for the data, a histogram of the residuals along with its kernel density curve and the normal residual distribution for the linear model are plotted, as shown in Figure 2. Lastly, the quadratic model is not found to significantly improve the interaction model, as shown in Table 3. Statistically speaking, the best fit model would be the simplest model that has a significant improvement from the model that has one less predictor involved. In this case, it is the simple linear model where amplitude is regressed on phonics scores alone. Nevertheless, for this project, many more predictors are tested and the interaction model is graphed for better visual understanding.

Conclusion

Phonics score is the only variable out of the four reading skills assessment scores that is shown to have a significant positive correlation with amplitude measurements. The gradual increase of the left hemisphere N170 amplitude along with the increase in phonics scores shows that students who are scoring higher on the Phonics Inventory assessment may be processing words more efficiently. This suggests that the accuracy and fluency with which students can identify individual letters and words may have a significant impact on word processing at the neurological level, more so than the other components. As a result, reading fluency may be of greater importance when considering for the development of reading interventions in schools. When phonics scores are combined with school intervention, there seems to be a main effect where participants who received school intervention have relatively more positive amplitudes than their counterparts. Though such main effect is not significant, the more positive amplitudes show that attending any type of school intervention may benefit reading fluency at the neural level. The lack of a main effect may be due to many factors, one of which could be the low sample size of 21 participants. Another reason may be the fact that the type of school intervention was not specified on the survey and thus some types of interventions may be driving a main effect while others hinder it. One other point to consider is that most participants who received intervention struggle with some type of language deficit. Thus, the fact that participants who received intervention still showed a neurological efficiency in processing words when compared to those who did not receive intervention is something worth noting. This may suggest that with the right type of intervention, significant improvements in reading skills may be established in students. Lastly, it is important to note that since this project is a correlational study, no conclusions of causation can be drawn, and hence future research would have to be conducted for detecting a clear directional effect.

References

Friederici, A.D (2011). The brain basis of language processing: form structure to function. *Physiological Reviews, 91*(4).

Froud, K., Levinson, L., Maddox, C., Avery, T., Lau, Jung, S….Wusinich, C. (2019). Poster from cognitive neuroscience society conference 2019. San Francisco, CA.

John Fox and Sanford Weisberg (2019). An {R} Companion to Applied Regression, Third Edition. Thousand Oaks CA: Sage.

URL: https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Maher, U., Brandeis, D. & McCandliss, B.D. (2005). Fast, visual specialization for reading in English revealed by the topography of the N170 erp response. *Behavioral Brain Function, 1*:1-13.

Sanchez-Vincitore L.V., Avery, T., & Froud, K. (2017). Word-related N170 responses to implicit and explicit reading task in neoliterate adults. *International Journal of Behavioral Development, 1833*:1-12.

Torgesen, J.K., Wagner, R.K. & Rashotte, C.A. (2012). *TOWRE-2 test of word reading efficiency- second edition.* Austin, TX: Pro Education.

Wagner, R.K., Torgesen, J.K., Rashotte, C.A., Pearson, N.A. (2013). *CTOPP-2: Comprehensive test of phonological processing-second edition.* Austin, TX: Pro Education.

Wagner, R.K. (2009). *Phonics inventory: Assessing phonics and sight words for older struggling readers.* New York, NY: Houghton Mifflin Harcourt Publishing Company.

Wagner, R.K. (2005). *Scholastic reading inventory target success with the lexile framework for reading.* New York, NY: Houghton Mifflin Harcourt Publishing Company.

Appendix A

Table 1

Descriptive Statistics of Continuous Variables

Descriptives				
Variables	Minimum	Maximum	Mean	Standard deviation
Amplitude	-6.37	3.20	-2.32	2.89
Lexile	137.00	1060.00	682.70	10.27
Phonics Inventory Score	2.00	36.00	18.19	10.27
CTOPP Score	11.00	32.00	21.52	6.10
TOWRE Score	12.00	232.00	167.60	48.34

Table 1. Shows the minimum, maximum, mean and standard deviation values of amplitude measures and assessment scores for Reading and Phonics Inventory, CTOPP, and TOWRE.

Table 2

Frequency Table of School Intervention

School Intervention	
Yes	No
9	12

Table 2. Shows the number of participants who did and did not receive school intervention.

Table 3

ANOVA Table for Nested Linear and Quadratic Models

Models for Amplitude Regressed on	*df*	*F*	Adjusted R^2	*p*
Intercept (Phonics Mean)	0, 20			
Phonics Score	1,19	5.124	0.227	0.038*
School Intervention	1,18	1.443	0.212	0.247
Interaction Between Score and Intervention	1,17	0.024	0.167	0.879
Phonics Quadratic	1,16	0.004	0.115	0.953

Note. Each subsequent variable listed is added onto the previous variable and the resulting linear model is nested in the previous one. A significant model improvement with a *p* indicated by * is found in the correlation between phonics scores amplitude when tested against the mean-only model; however, none of the subsequent models seem to further improve this model. The adjusted R^2 shows that as more predictors are added, the less improvement each additional predictor provides the model.

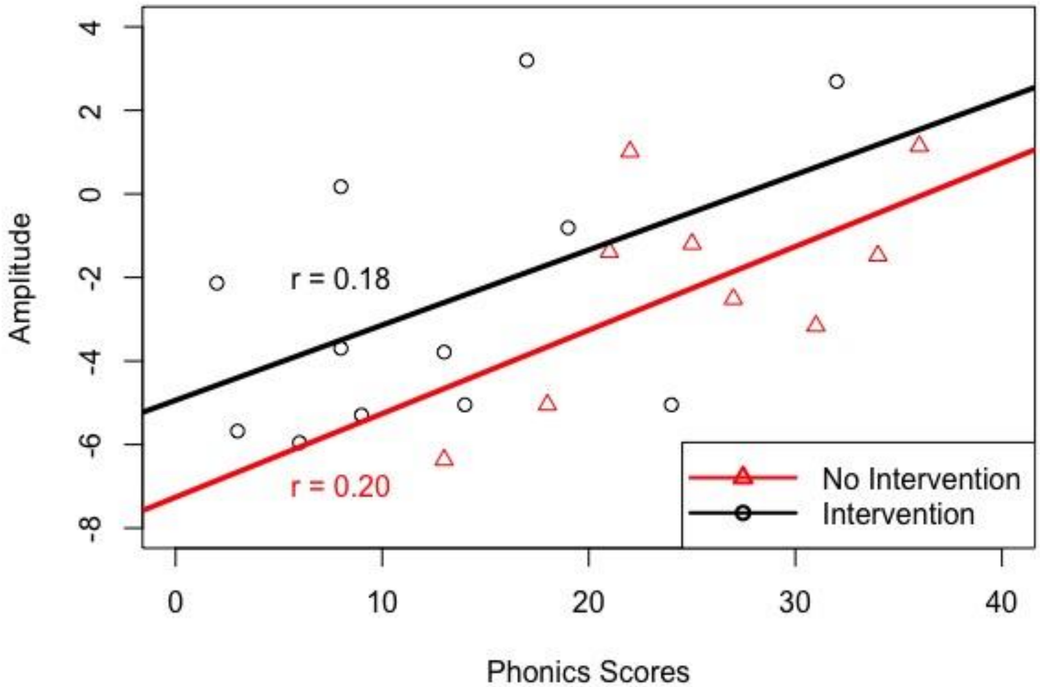

Figure 1. Amplitudes seem to be higher for participant who received school intervention as opposed to those who did not. No interaction seems to be present as both lines have similar slopes

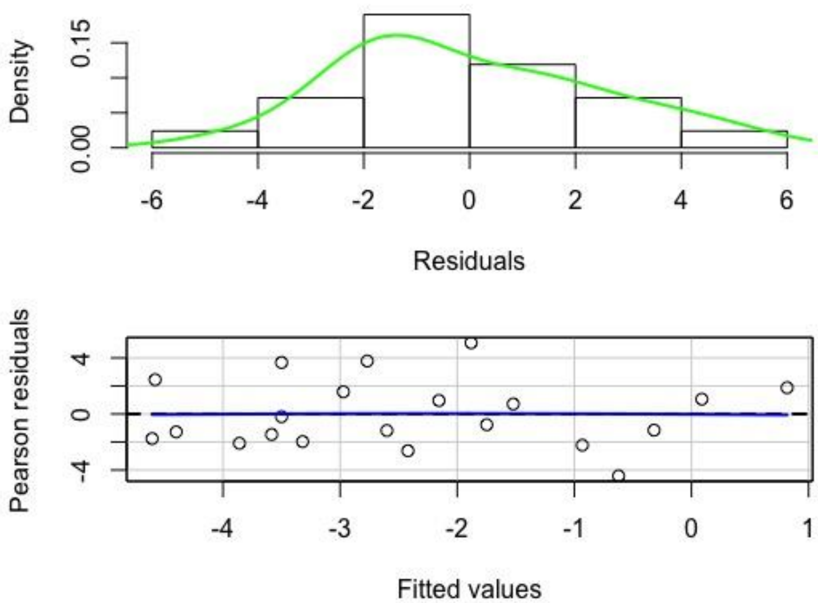

Figure 2. The residual density distribution for the interaction model is close to being normally distributed. As a confirmation, the Pearson residual plot for the same model seem to fit the values well and thus indicate that the interaction model appropriate portrays the prediction of amplitudes.

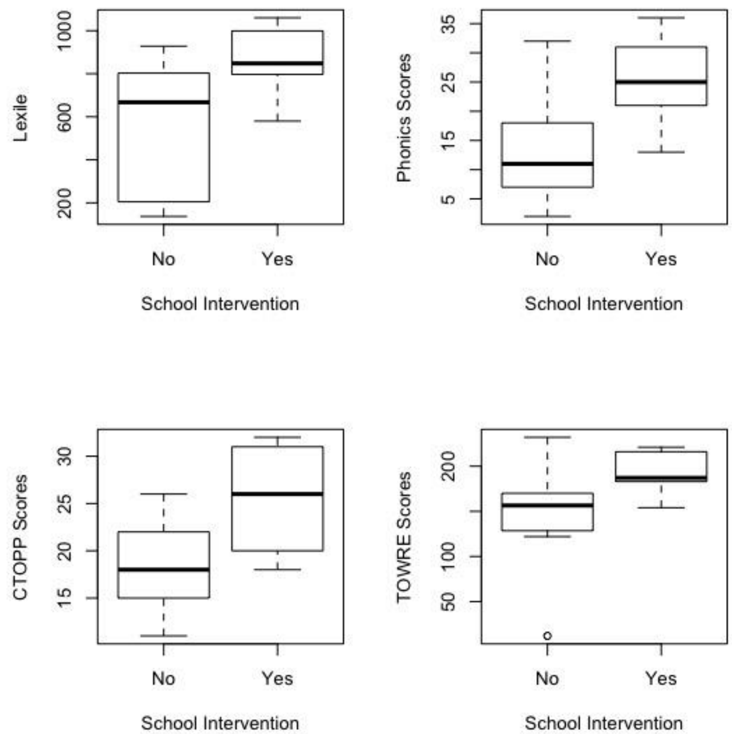

Figure 3. These boxplots show the differences in assessment scores for participants who received intervention and for those who did not receive intervention. Differences in score between intervention groups seem to be the greatest for phonics and CTOPP scores.