Published in final edited form as:

Mind Brain Educ. 2015 September ; 9(3): 145–153. doi:10.1111/mbe.12083.

N400 Event-Related Potential and Standardized Measures of Reading in Late Elementary School Children: Correlated or Independent?

Donna Coch and Clarisse Benoit¹

Dartmouth College

Abstract

We investigated whether and how standardized behavioral measures of reading and electrophysiological measures of reading were related in 72 typically developing, late elementary school children. Behavioral measures included standardized tests of spelling, phonological processing, vocabulary, comprehension, naming speed, and memory. Electrophysiological measures were composed of the amplitude of the N400 component of the event-related potential waveform elicited by real words, pseudowords, nonpronounceable letter strings, and strings of letter-like symbols (false fonts). The only significant brain-behavior correlations were between standard scores on the vocabulary test and N400 mean amplitude to real words ($r = -.272$) and pseudowords ($r = -.235$). We conclude that, while these specific sets of standardized behavioral and electrophysiological measures both provide an index of reading, for the most part, they are independent and draw upon different underlying processing resources.

[T]o completely analyze what we do when we read... would be to describe very many of the most intricate workings of the human mind, as well as to unravel the tangled story of the most remarkable specific performance that civilization has learned in all its history

(Huey, 1908/1968, p. 3).

Reading is a complex task that involves numerous kinds of skills and knowledge, as well as numerous neural networks (e.g., Adams, 1990; Dehaene, 2009). There is relatively strong behavioral evidence indicating that knowledge about print and orthographic awareness (e.g., Perfetti & Bolger, 2004; Treiman, 2000), knowledge of phonemes and phonological awareness (e.g., Bryant, MacLean, & Bradley, 1990; Shaywitz, 1996), knowledge concerning word meanings and vocabulary (e.g., Carlisle, 2000; Nagy & Townsend, 2012), and comprehension skills (e.g., Hulme & Snowling, 2011; Willingham, 2006–2007), amongst others, are integral to reading development and achievement. However, there is less evidence concerning how behavioral measures of these elements of reading may be correlated with electrophysiological measures of reading. From the perspective of Mind, Brain, and Education, with its emphasis on multiple levels of analysis (e.g., Ansari & Coch, 2006), exploring such brain-behavior relationships will be key in unraveling the story of

Corresponding author: Donna Coch, Dartmouth College, Department of Education, Reading Brains Lab, 3 Maynard Street, Raven House, HB 6103, Hanover, NH, 03755. Telephone: 603.646.3282. FAX: 603.646.3968. donna.coch@dartmouth.edu.

¹now a teacher at Samuel Morey Elementary School, Fairlee, Vermont

reading development. In this study, we investigated whether and how standardized measures of orthographic and phonological knowledge, vocabulary, and comprehension were related to the amplitude of the N400 component of the event-related potential (ERP) waveform in a group of typically developing, late elementary school children.

ERPs are a derivative of the electroencephalogram, and are time-locked to presentations of stimuli. The ERP waveform represents the averaged response to multiple presentations of a specific stimulus type and is characterized by a series of positive and negative deflections. These peaks and valleys in the waveform are typically labeled with a P (positive-going deflection) or an N (negative-going deflection) and a number indicating the time in milliseconds that the peak or valley maximally occurs after presentation of the stimulus. Each of the deflections in the ERP waveform is associated with different kinds of processing (e.g., Luck, 2005). The N400 component, a negative deflection in the waveform peaking at about 400 ms after stimulus presentation, was originally identified in sentence processing studies in conditions of semantic ambiguity or novelty. For example, the anomalous terminal word *socks* in the sentence *He spread the warm bread with socks* elicited a marked N400 in comparison to the canonical terminal word *butter* (Kutas & Hillyard, 1980). Since then, the N400 has consistently been associated with lexical and semantic processing in studies with adults (for reviews, see Kutas & Federmeier, 2011; Lau, Phillips, & Poeppel, 2008).

Comparatively few studies of reading with children have used the N400 to investigate the typical course of lexicosemantic development (e.g., Benau, Morris, & Couperus, 2011; Coch, Maron, Wolf, & Holcomb, 2002; Eddy, Grainger, Holcomb, Mitra, & Gabrieli, 2014; Holcomb, Coffey, & Neville, 1992). Those that have done so have suggested developmental changes in the component. For example, Holcomb and colleagues (1992) presented visual sentences with canonical or anomalous terminal words to 7- to 26-year-olds and found that anomalous words elicited larger N400s than canonical terminal words in all the age groups. However, canonical terminal words also elicited substantial N400s, only in the youngest children (age groups 7–8 and 9–10). Similarly, in a study with 10- and 11-year-olds presenting single stimuli in list form, Coch et al. (2002) found that real words, pseudowords, nonpronounceable letter strings, and strings of letter-like symbols (false fonts) all elicited marked N400s. This contrasts with data from adults in which the N400 is much larger to potentially meaningful strings, such as words and pseudowords, than to meaningless strings, such as random sequences of letters (e.g., Bentin, Mouchetant-Rostaing, Giard, Echallier, & Pernier, 1999; Nobre & McCarthy, 1994).

A number of studies have used the N400 to investigate lexicosemantic development, in combination with behavioral measures, but without using an ERP reading paradigm. For example, there have been interesting investigations of N400 effects in infants for spoken words paired with pictures (e.g., Friedrich & Friederici, 2005); changes in phonological representations (N400 measured to spoken words) over the course of early reading acquisition (e.g., Bonte & Blomert, 2004); relationships between the N400 elicited by picture and spoken word pairs and vocabulary knowledge in both typically developing and special populations (e.g., Byrne et al., 1999; Byrne, Dywan, & Connolly, 1995a; Byrne, Dywan, & Connolly, 1995b; McCleery et al., 2010; Molfese & Morris, 1990); and relationships between the size of the N400 elicited by pairs of pictures and spoken words

and scores on tests of listening comprehension, nonword decoding, and word reading in 8- to 10-year-olds (children with better reading and decoding skills showed smaller amplitude N400s, Henderson, Baseler, Clarke, Watson, & Snowling, 2011). However, these studies are limited in that they did not record ERPs to actual written word stimuli.

ERP studies using both standardized measures of reading and some sort of reading paradigm to elicit the N400 with typically developing children are scarce. In one, using a variant of the Reicher-Wheeler paradigm (e.g., Reicher, 1969; Wheeler, 1970) with 24 7- and 24 11-year-olds, the authors reported a moderate correlation ($r = -.418$) between the average peak amplitude of the N400 and standard scores on the Woodcock Reading Mastery Tests - Revised Word Identification subtest (Coch, Mitra, & George, 2012). In another, a composite score based on Woodcock Reading Mastery Tests - Revised Letter Identification, Word Identification, and Word Attack subtest scores was strongly correlated ($r = -.65$) with N400 amplitude averaged across known words, unknown words, and difficult words in a group of 14 first-grade girls, suggesting a “relatively robust relation between N400 amplitude and reading ability” in beginning readers (Coch & Holcomb, 2003, p. 160). However, these studies were small and spanned across early and late elementary school.

Studies of struggling readers that have included typically developing control groups, standardized behavioral measures, and an ERP reading task may also speak to the potential relationship between electrophysiological and standardized behavioral measures of reading. For example, Hasko, Groth, Bruder, Bartling, and Schulte-Körne (2013) recently reported a study in which 8-year-old children (29 controls and 52 with developmental dyslexia) were asked to judge whether singly presented written words, pseudohomophones, pseudowords, and false fonts sounded like a real German word or not. The mean peak amplitude of the N400, calculated across the three letter string conditions (i.e., excluding false fonts), was correlated with spelling scores ($r = -.25$), such that better spelling was related to larger N400 amplitude, regardless of diagnosis. Thus, there is some suggestion in the literature, across both typically and atypically developing populations, that N400 ERP amplitude to written strings and various standardized behavioral measures of reading and spelling are correlated.

Here, we built on these scattered reports and directly analyzed brain-behavior relationships in a group of 72 typically developing late elementary school students. We used the N400 elicited by single real words, pseudowords, nonpronounceable letter strings, and strings of false font characters as our electrophysiological measures (Coch, 2015) and scores on standardized tests of spelling, phonological processing, vocabulary, comprehension, naming speed, and memory as our behavioral measures. Based on the scant developmental literature, we hypothesized that the amplitude of the N400 to real words would be associated with behavioral semantic measures, such as measures of vocabulary and, perhaps, comprehension (e.g., Holcomb et al., 1992), and our measure of spelling (e.g., Hasko et al., 2013). Others have noted a “general need in this field to combine behavioural and neurological techniques to enable a better understanding of whether we are measuring the same underlying processes with each methodology” (Henderson et al., 2011, p. 90). This is the overarching question addressed in the current study.

Methods

Participants

Participants included 72 elementary school students in the third, fourth, and fifth grades in public and private schools in the northeastern United States (37 female, average age 10;0, *SD* 1;0).¹ Students were recruited through posters placed in public places, advertisements in the local newspaper, descriptions posted in school newsletters, flyers sent home through schools, postings on town listservs, and word of mouth. All participants were right-handed (Oldfield, 1971), monolingual English speakers (who learned English as a first language and were not fluent in any other language). By parent report, none had a history of neurological dysfunction or a diagnosed language or reading disorder, and none were taking any medications that would affect brain function. As tested with a standard Snellen chart, all had normal or corrected-to-normal binocular vision (20/40 or better). All participants were volunteers, and were given small prizes and paid \$20 for their time. The study was completed in one or two sessions at a college lab (standardized testing battery and ERP recording portions together or separately; for single sessions, behavioral testing was done first).

Standardized Behavioral Tests

Participants were given a battery of standardized tests measuring various aspects of reading-related skills. The Sight Spelling and Homophone Choice subtests comprising the Spelling Accuracy composite of the Test of Orthographic Competence (TOC, Mather, Roberts, Hammill, & Allen, 2008) provided a measure of orthographic knowledge. In Sight Spelling, the examiner says a word and the participant is shown part of the word, with one or more letters missing; the participant is asked to fill in the missing letter or letters to complete the spelling of the word. In Homophone Choice, the participant is shown a picture with two or three possible (homophonic) spelling choices, and is asked to circle the word that she thinks is the correct spelling for the picture.

The Elision and Blending Words subtests comprising the Phonological Awareness composite of the Comprehensive Test of Phonological Processing (CTOPP, Wagner, Torgesen, & Rashotte, 1999) and the Phonemic Decoding Efficiency subtest of the Test of Word Reading Efficiency (TOWRE, Torgesen, Wagner, & Rashotte, 1999) measured phonological knowledge. Elision involves removing phonological segments from spoken words to form other words, Blending Words measures the ability to synthesize sounds to form words, and Phonemic Decoding Efficiency measures accuracy and fluency in reading phonemically regular nonwords under timed conditions.

The Peabody Picture Vocabulary Test IIIA (PPVT, Dunn & Dunn, 1997), in which the participant is shown a field of four pictures and asked to select the one picture that has the same meaning as the word the examiner has said, provided a measure of receptive vocabulary. Comprehension was measured by the Passage Comprehension subtest of the Woodcock Reading Mastery Test - Revised (WRMT, Woodcock, 1987), which requires

¹A group of 24 college students also participated in the overall study, but data from this group are not included in this developmental report focused on elementary students.

silent reading of a passage and filling in the blank with a word that fits. The Letters and Numbers subtests from the Rapid Automatized Naming and Rapid Alternating Stimulus Tests (RAN/RAS, Wolf & Denckla, 2005) served as measures of naming speed; these tests involve naming lists of letters and numbers as quickly and accurately as possible under timed conditions. Finally, the Memory for Digits subtest of the CTOPP (Wagner et al., 1999) was used as a measure of short-term memory. Administration of these tests took about one hour. Due to experimenter error, standard scores were not available for one child for the TOC spelling composite (Mather et al., 2008), or for another child for the PPVT (Dunn & Dunn, 1997); in each case, the rest of the behavioral scores for that child were included in analyses, as were the child's electrophysiological data.

ERP Stimuli and Analysis

The ERP stimuli and analyses have been described in detail elsewhere (Coch, 2015). Briefly, the stimulus list consisted of 60 instances of each of five types of highly controlled stimuli: real words (e.g., *bed*, *bring*), pseudowords (e.g., *bem*, *fring*), nonpronounceable letter strings (e.g., *mbe*, *nrfgi*), false font strings (e.g., BEM,,FRING), and target animal names (e.g., *cat*, *duck*). Real words were single syllables, regular, and decodable. Pseudowords were formed by changing one letter of each real word. Letter strings were formed by randomly rearranging letters from each pseudoword into orthographically and phonologically illegal strings. False font strings were composed of characters based on each letter in a pseudoword. Stimuli were presented singly, in list form (i.e., one by one, one after another), at the center of an LCD monitor directly in front of each participant. They were presented in pseudorandom order, such that related stimuli were separated by at least four stimuli (e.g., *bed*, *bem*, *mbe*, and BEM were each separated by at least four intervening stimuli). Stimuli were presented for 500 ms each, in white Times New Roman font on a black background. Children were simply asked to press a button on a game controller whenever they saw an animal name (semantic categorization task). Recording time for all 300 stimuli was approximately 20 minutes. After recording, for each subject, separate ERPs to words, pseudowords, letter strings, and false fonts were averaged at each electrode site over a 1000 ms epoch, using a 200 ms pre-stimulus-onset baseline. Trials contaminated by artifacts such as eye movements, muscular activity, or electrical noise were rejected from these averages. The mean number of trials in each condition of interest was: 46.2 (*SD* 8.1) for real words, 47.0 (*SD* 8.1) for pseudowords, 47.1 (*SD* 8.4) for letter strings, and 47.5 (*SD* 8.0) for false fonts. Mean amplitude of the N400 component was subsequently measured within the 300–550 ms time window within each of the four conditions.

Procedure

Participants were given a brief overview of the study and tour of the lab, and questions were addressed before parents signed a permission form and children signed an assent form. Following behavioral testing as described above, participants were fitted with an electrode cap for electroencephalogram recording. Electrophysiological recording details are reported elsewhere (Coch, 2015). Participants were then seated in a comfortable chair in a sound attenuating and electrically shielded booth for the ERP semantic categorization task. Following ERP recording, participants were given a brief paper-and-pencil post-test in which they were asked to circle all of the real words on a sheet of paper with half the stimuli

used in the ERP task printed in five columns; the post-test data, which confirmed that the children could read the real words as intended [average accuracy: 97.3% (*SD* 3.3%)], are not considered further here.

Data Analysis

A Pearson correlation analysis was run using the standard scores from the behavioral tests and the mean amplitude measures of the N400 component at lateral and medial recording sites to word, pseudoword, letter string, and false font stimuli (see Figure 1). The two-tailed tests were considered significant at $p = .05$. Preliminary analyses showed no significant correlations between age and any of the ERP measures (all $ps > .28$), and standard scores controlled for age for the behavioral test data, so age was not partialled. Follow-up regression analyses were used to determine relative predictiveness for significant correlation results.

Results

Behavioral Measures

Descriptive statistics (mean, standard deviation) and the results of the correlation analysis are summarized in Table 1. Standard scores on the behavioral measures of spelling, phonological awareness, vocabulary, and comprehension were all significantly correlated with one another (rs from .256 to .714). Standard scores on the three timed behavioral tasks (TOWRE and RAN/RAS) were also significantly correlated (rs from .370 to .735). Further, scores on the digit span memory measure were significantly correlated with all scores on the behavioral reading measures except the RAN/RAS scores (rs from .312 to .491).

Electrophysiological Measures

Similarly, the electrophysiological measures were all significantly correlated with one another (rs from .654 to .845).

ERP-Behavioral Measure Correlations

Addressing the primary research question, the only significant correlations between the behavioral and ERP measures were small, negative correlations between standard scores on the PPVT and the mean amplitude of the N400 to words, $r = -.272$, $p = .022$, 95% CI $[-.474, -.043]$ and pseudowords, $r = -.235$, $p = .048$, 95% CI $[-.443, -.004]$.

Regression of ERP Measures on Behavioral Measures

Regression analyses confirmed the correlation findings, indicating that the mean amplitude of the N400 to words explained a small but significant amount of the variance in PPVT standard scores, adjusted $R^2 = .061$, $F(1, 69) = 5.51$, $p = .022$, $\beta = -.27$, as did the mean amplitude of the N400 to pseudowords, adjusted $R^2 = .042$, $F(1, 69) = 4.04$, $p = .048$, $\beta = -.24$. Stepwise regression indicated that the mean amplitude of the N400 to pseudowords did not add any significant predictive value, $p = .94$, to the model that already included the mean amplitude of the N400 to words.

Discussion

In an exploration of whether and how standardized behavioral measures of spelling, phonological processing, vocabulary, comprehension, naming speed, and memory were correlated with electrophysiological measures of the N400 component elicited by single word, pseudoword, letter string, and false font stimuli in a group of 72 typically developing late elementary school students, we found few statistically significant brain-behavior relationships. Indeed, the only such relationships were between standard scores on the vocabulary test (PPVT, Dunn & Dunn, 1997) and N400 mean amplitude to words and pseudowords; thus, N400 amplitude was not related to most elements of reading assessed, except semantics. Overall, this pattern of results suggests that, in typically developing late elementary school students, the specific behavioral and electrophysiological measures of reading used here, for the most part, do not access the same underlying processes.

The correlations between PPVT scores and N400 amplitude to words ($r = -.272$) and pseudowords ($r = -.235$), but not letter strings or false font strings, are consistent with ERP research indicating that the N400 indexes lexicosemantic processing (e.g., Kutas & Federmeier, 2011; Lau et al., 2008); these measures appear to be proxies for partly overlapping, but substantially different, aspects of word knowledge. Although this may not seem to be the case for pseudowords, these stimuli are thought to elicit an N400, despite lacking meaning or lexicosemantic representations, due to partial activation of similar real word representations (e.g., Holcomb, Grainger, & O'Rourke, 2002). That N400 mean amplitude to real words was a better predictor of PPVT scores than N400 mean amplitude to pseudowords supports this interpretation. That higher PPVT scores were associated with more negative N400 amplitudes (i.e., larger N400s) is consistent with a previous report showing that better reading scores (across a composite of subtests) were associated with larger N400s in a small group of first-grade girls ($r = -.65$, Coch & Holcomb, 2003); however, the strength of this relationship appears to decrease over time from the early to late elementary grades.

These small, negative correlational findings are reminiscent of a recent report of a weak, negative relationship ($r = -.25$) between spelling scores and average N400 amplitude to words, pseudohomophones, and pseudowords (combined) in 8-year-olds (Hasko et al., 2013). Here, we did not find significant correlations between our spelling scores and N400 measures, but our sample included only typically developing children, while Hasko et al. included children with dyslexia. Further, our stimuli were in English, a language with a deep orthography, while Hasko et al.'s were in German; reading and spelling development differ with depth of orthography (e.g., Ziegler et al., 2010). Unfortunately, it does not appear that Hasko et al. used a standardized test of vocabulary that would allow for direct comparison with the present findings.

Finally, with regard to these findings, it is interesting to note that the N400 elicited by picture-spoken word pairings in a previous study was reportedly *smaller* in 8- to 10-year-olds with better reading and decoding skills, with positive correlations ranging between .57 and .63 (Henderson et al., 2011). Henderson and colleagues (2011) used a baseline-to-peak amplitude measure (as compared to our mean amplitude measure) and averaged across their

congruent and incongruent conditions (as compared to our four isolated conditions), so direct comparison with the present study is difficult. However, the pattern across studies does suggest potential brain-behavior differences between N400 measures involving reading and those involving images and listening.

Not surprisingly, standard scores on the behavioral measures of spelling, phonological awareness, vocabulary, and comprehension were all significantly correlated with one another. Previous studies have shown that these elements of reading are related in children in late elementary school (e.g., fourth graders, Ouellette, 2006). Scores on the RAN/RAS subtests (Wolf & Denckla, 2005) were correlated with scores on our third timed measure, the Phonemic Decoding Efficiency subtest of the TOWRE (Torgesen et al., 1999), presumably as indices of processing speed. However, RAN/RAS scores were not correlated with scores on the Phonological Awareness composite of the CTOPP (Wagner et al., 1999), consistent with the view that RAN/RAS tasks are not simply a measure of phonological processing (e.g., Wolf & Bowers, 1999). Further supporting this distinction, standard scores on the digit span test (Wagner et al., 1999) were moderately, positively correlated with standard scores on all of the other behavioral tests except the RAN/RAS. If digit span performance can be considered a measure of short-term phonological memory, these moderate, positive correlations are consistent with the importance of phonological memory for reading.

Also not surprisingly, each of the N400 ERP measures was strongly, positively correlated with the other N400 ERP measures. Thus, the overall pattern of findings indicates consistency among the electrophysiological measures, and relative consistency across the standardized behavioral measures, but few correlations between the electrophysiological and behavioral measures. Such differentiation has been reported before in both the adult (e.g., Coch & Mitra, 2010; Landi & Perfetti, 2007) and developmental (e.g., Coch et al., 2012; Henderson et al., 2011) ERP literatures. It has been noted that behavioral measures may not be as sensitive to subtle processing differences as electrophysiological measures, and that behavioral measures “tap the end point of processing,” while ERP measures provide a direct, on-line index of processing in real time (e.g., Henderson et al., 2011, p. 97); therefore, they may provide “complementary information” (e.g., Coch et al., 2012, p. 76).

Given this complementarity, the relevance of these findings to the classroom might not be immediately clear. Indeed, the specificity of the significant correlations between the neural and behavioral measures and the lack of widespread correlations could lead educators to believe that the worlds of neuroscience research and classroom practice are essentially parallel lines of inquiry, moving in the same direction in attempting to understand how people learn, but not intersecting or interacting with one another. Practicing teachers generally do not bring students into the neuroscience lab for individual testing, and the results of a neuroscience study – the average across a select sample of children doing a controlled task in a lab setting – cannot be taken to dictate what teachers should do with individual children in a real-life classroom. However, although these findings do not directly inform teaching practice, this does not mean that they have no bearing on the classroom.

Taken together, the behavioral and ERP data create a more complete picture of “reading,” revealing aspects of both the processes and the products of reading. Similar to many classroom assessments, the behavioral measures used here are outcome measures of the products of processing. In contrast, electrophysiological measures reflect the processing itself, the subcomponents and steps along the way that, combined, lead to the products. Thus, studies like this may help teachers to re-conceptualize “reading” as a complex skill comprised of many different elements that work together and develop over time. In our experience, many master teachers have developed such insight into the processes and products of reading through years of experience working with students in the classroom. Teachers-in-training, without the benefit of such experience, could potentially accelerate the development of this insight through exposure to neuroscience research. This sort of connection between neuroscience and education is subtle, but it has potentially important implications for both classroom practice and teacher training. Exposing teachers-in-training to relevant neuroscience research could deepen their understanding of the complexities of reading. In turn, this could influence their decision making in the classroom, facilitate principled deconstruction of the processes and products of reading, and allow teachers to devise teaching strategies and activities that address the components of reading and enhance the process of combining the many parts of reading into a coherent whole skill that is truly more than the sum of its parts.

Beyond direct connection to the classroom, this study has a number of limitations. Among these are the nature of the volunteer sample, which, despite variability, tended to include relatively strong readers (refer to Table 1). Although few electrophysiological studies have considered such a large sample of typically developing late elementary readers, the findings may not generalize to a wider population, poorer readers of the same age, or younger or older readers. Whereas the study was designed as an investigation of the N400 component, another limitation is that we have not considered possible correlations between our standardized behavioral measures and other ERP components, due to space constraints.

Despite these limitations, we conclude that, while the standardized behavioral and electrophysiological measures used here both provide an index of reading, for the most part, the two sets of measures are independent and draw upon different underlying reading resources; it was only in the specific case of semantic processing that we observed overlap. Thus, an investigative approach that includes both behavioral and electrophysiological measures and an educational approach that includes understanding of both processes and products – that is, multiple levels of analysis (e.g., Ansari & Coch, 2006) – will likely be more powerful in unraveling the full story of reading development than an approach that considers either alone.

Acknowledgments

Many thanks are due to the parents and children who elected to participate in this study, and to the local schools, libraries, businesses, and programs that allowed us to share information about the study with families and children. The project could not have been completed without help from undergraduate research assistants Jennifer Bares, Natalie Berger, Sarah Brim, Ayesha Dholakia, Elyse George, Emily Jasinski, Gabriela Meade, Priya Mitra, and Anna Roth, and help with programming from Ray Vukceвич and Mark Dow. This research was supported by grant R03HD058613 from the National Institutes of Health, Eunice Kennedy Shriver National Institute of Child Health and Human Development.

References

- Adams, MJ. Beginning to read: thinking and learning about print. Cambridge: MIT Press; 1990.
- Ansari D, Coch D. Bridges over troubled waters: education and cognitive neuroscience. *Trends in Cognitive Sciences*. 2006; 10(4):146–151.10.1016/j.tics.2006.02.007 [PubMed: 16530462]
- Benau EM, Morris J, Couperus JW. Semantic processing in children and adults: incongruity and the N400. *Journal of Psycholinguistic Research*. 2011; 40:225–239.10.1007/s10936-011-9167-1 [PubMed: 21479571]
- Bentin S, Mouchetant-Rostaing Y, Giard MH, Echallier JF, Pernier J. ERP manifestations of processing printed words at different psycholinguistic levels: time course and scalp distribution. *Journal of Cognitive Neuroscience*. 1999; 11(3):235–260.10.1162/089892999563373 [PubMed: 10402254]
- Bonte ML, Blomert L. Developmental changes in ERP correlates of spoken word recognition during early school years: a phonological priming study. *Clinical Neurophysiology*. 2004; 115:409–423.10.1016/S1388-2457(03)00361-4 [PubMed: 14744584]
- Bryant P, MacLean M, Bradley L. Rhyme, language, and children's reading. *Applied Psycholinguistics*. 1990; 11:237–252.10.1017/S0142716400008870
- Byrne JM, Connolly JF, MacLean SE, Dooley JM, Gordon KE, Beattie TL. Brain activity and language assessment using event-related potentials: development of a clinical protocol. *Developmental Medicine and Child Neurology*. 1999; 41(11):740–747.10.1017/S0012162299001504 [PubMed: 10576638]
- Byrne JM, Dywan CA, Connolly JF. Assessment of children's receptive vocabulary using event-related brain potentials: development of a clinically valid test. *Child Neuropsychology*. 1995a; 1(3): 211–223.10.1080/09297049508400226
- Byrne JM, Dywan CA, Connolly JF. An innovative method to assess the receptive vocabulary of children with cerebral palsy using event-related brain potentials. *Journal of Clinical and Experimental Neuropsychology*. 1995b; 17(1):9–19.10.1080/13803399508406576 [PubMed: 7608306]
- Carlisle JF. Awareness of the structure and meaning of morphologically complex words: impact on reading. *Reading and Writing: An Interdisciplinary Journal*. 2000; 12(3/4):169–190.10.1023/A: 1008131926604
- Coch D. The N400 and the fourth grade shift. *Developmental Science*. 2015; 18(2):254–269.10.1111/desc.12212 [PubMed: 25041502]
- Coch D, Holcomb PJ. The N400 in beginning readers. *Developmental Psychobiology*. 2003; 43(2): 146–166.10.1002/dev.10129 [PubMed: 12918093]
- Coch D, Maron L, Wolf M, Holcomb PJ. Word- and picture-processing in children: An event-related potential study. *Developmental Neuropsychology*. 2002; 22(1):373–406.10.1207/S15326942dn2201_3 [PubMed: 12405510]
- Coch D, Mitra P. Word and pseudoword superiority effects reflected in the ERP waveform. *Brain Research*. 2010; 1329:159–174.10.1016/j.brainres.2010.02.084 [PubMed: 20211607]
- Coch D, Mitra P, George E. Behavioral and ERP evidence of word and pseudoword superiority effects in 7- and 11-year-olds. *Brain Research*. 2012; 1486:68–81.10.1016/j.brainres.2012.09.041 [PubMed: 23036274]
- Dehaene, S. Reading in the brain: the science and evolution of a human invention. New York: Viking; 2009.
- Dunn, LM.; Dunn, LM. Peabody picture vocabulary test. 3. Circle Pines, MN: American Guidance Service; 1997.
- Eddy MD, Grainger J, Holcomb PJ, Mitra P, Gabrieli JDE. Masked priming and ERPs dissociate maturation of orthographic and semantic components of visual word recognition in children. *Psychophysiology*. 2014; 51(2):136–141.10.1111/psyp.12164 [PubMed: 24313638]
- Friedrich M, Friederici AD. Phonotactic knowledge and lexical-semantic processing in one-year-olds: brain responses to words and nonsense words in picture contexts. *Journal of Cognitive Neuroscience*. 2005; 17(11):1785–1802.10.1162/089892905774589172 [PubMed: 16269114]

- Hasko S, Groth K, Bruder J, Bartling J, Schulte-Körne G. The time course of reading processes in children with and without dyslexia: an ERP study. *Frontiers in Human Neuroscience*. 2013; 7:Article 570.10.3389/fnhum.2013.00570 [PubMed: 24109444]
- Henderson LM, Baseler HA, Clarke PJ, Watson S, Snowling MJ. The N400 effect in children: relationships with comprehension, vocabulary and decoding. *Brain and Language*. 2011; 117:88–99.10.1016/j.bandl.2010.12.003 [PubMed: 21272930]
- Holcomb PJ, Coffey SA, Neville HJ. Visual and auditory sentence processing: A developmental analysis using event-related brain potentials. *Developmental Neuropsychology*. 1992; 8:203–241.10.1080/87565649209540525
- Holcomb PJ, Grainger J, O'Rourke T. An electrophysiological study of the effects of orthographic neighborhood size on printed word perception. *Journal of Cognitive Neuroscience*. 2002; 14(6): 938–950.10.1162/089892902760191153 [PubMed: 12191460]
- Huey, EB. *The psychology and pedagogy of reading*. Cambridge, MA: MIT Press; 1908/1968.
- Hulme C, Snowling MJ. Children's reading comprehension difficulties: nature, causes, and treatments. *Current Directions in Psychological Science*. 2011; 20:139–142.10.1177/0963721411408673
- Kutas M, Federmeier K. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*. 2011; 62:621–647.10.1146/annurev.psych.093008.131123
- Kutas M, Hillyard SA. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*. 1980; 207:203–204.10.1126/science.7350657 [PubMed: 7350657]
- Landi N, Perfetti CA. An electrophysiological investigation of semantic and phonological processing in skilled and less-skilled comprehenders. *Brain and Language*. 2007; 102:30–45.10.1016/j.bandl.2006.11.001 [PubMed: 17188346]
- Lau EF, Phillips C, Poeppel D. A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience*. 2008; 9:920–933.10.1038/nrn2532
- Luck, SJ. *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press; 2005.
- Mather, N.; Roberts, R.; Hammill, DD.; Allen, EA. *Test of orthographic competence*. Austin, TX: Pro-Ed; 2008.
- McCleery JP, Ceponiene R, Burner KM, Townsend J, Kinnear M, Schreibman L. Neural correlates of verbal and nonverbal semantic integration in children with autism spectrum disorders. *Journal of Child Psychology and Psychiatry*. 2010; 51(3):277–286.10.1111/j.1469-7610.2009.02157.x [PubMed: 20025622]
- Molfese DL, Morris RD. Semantic discrimination in nonspeaking youngsters with moderate or severe retardation: electrophysiological correlates. *Brain and Language*. 1990; 38(1):61–74.10.1016/0093-934X(90)90102-M [PubMed: 2302546]
- Nagy W, Townsend D. Words as tools: learning academic vocabulary as language acquisition. *Reading Research Quarterly*. 2012; 47(1):91–108.10.1002/RRQ.011
- Nobre AC, McCarthy G. Language-related ERPs: Scalp distributions and modulation by word-type and semantic priming. *Journal of Cognitive Neuroscience*. 1994; 6:233–255.10.1162/jocn.1994.6.3.233 [PubMed: 23964974]
- Oldfield RC. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*. 1971; 9:97–113.10.1016/0028-3932(71)90067-4 [PubMed: 5146491]
- Ouellette GP. What's meaning got to do with it: the role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*. 2006; 98(3):554–566.10.1037/0022-0663.98.3.554
- Perfetti CA, Bolger DJ. The brain might read that way. *Scientific Studies of Reading*. 2004; 8(3):293–304.10.1207/s1532799xssr0803_7
- Reicher GM. Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*. 1969; 81(2):275–280.10.1037/h0027768 [PubMed: 5811803]
- Shaywitz SE. Dyslexia. *Scientific American*. 1996; 275(5):98–104.10.1038/scientificamerican1196-98 [PubMed: 8875810]
- Torgesen, JK.; Wagner, R.; Rashotte, C. *Test of word reading efficiency*. Austin, TX: Pro-Ed; 1999.

- Treiman R. The foundations of literacy. *Current Directions in Psychological Science*. 2000; 9(3):89–92.10.1111/1467-8721.00067
- Wagner, RK.; Torgesen, JK.; Rashotte, CA. *The comprehensive test of phonological processing*. Austin, TX: Pro-Ed; 1999.
- Wheeler DD. Processes in word recognition. *Cognitive Psychology*. 1970; 1:59–85.10.1016/0010-0285(70)90005-8
- Willingham DT. The usefulness of brief instruction in reading comprehension strategies. *American Educator*. 2006–2007:39–45. 50.
- Wolf M, Bowers PG. The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology*. 1999; 91(3):415–438.10.1037/0022-0663.91.3.415
- Wolf, M.; Denckla, MB. *Rapid automatized naming and rapid alternating stimulus tests (RAN/RAS)*. Austin, TX: Pro-Ed; 2005.
- Woodcock, RW. *The Woodcock reading mastery tests - revised*. Circle Pines, MN: American Guidance Service; 1987.
- Ziegler JC, Bertrand D, Tóth D, Csépe V, Reis A, Faísca L, Blomert L. Orthographic depth and its impact on universal predictors of reading: a cross-language investigation. *Psychological Science*. 2010; 21(4):551–559.10.1177/0956797610363406 [PubMed: 20424101]

Figure 1.

Grand average ERP waveforms elicited by words (solid black line), pseudowords (dotted line), letter strings (dashed line), and false fonts (solid gray line) for the group of upper elementary students ($N = 72$) across lateral and medial recording sites. Each vertical tick marks 100 ms, the calibration bar marks $4.0 \mu\text{V}$, and negative is plotted up. Mean amplitude of the N400 component was measured within the 300–550 ms time window, shaded light gray, and used in correlation analyses.

Table 1

Summary of correlations, means, and standard deviations for standard scores on the behavioral tests (items 1 through 8) and mean amplitude of the N400 to word and word-like stimuli (items 9 through 12)

	1	2	3	4	5	6	7	8	9	10	11	12
1. TOC SPACC	----											
2. CTOPP PA	.265*	----										
3. CTOPP MD	.312**	.376**	----									
4. PPVT	.372**	.256*	.342**	----								
5. WRMT PC	.505***	.387**	.491***	.670***	----							
6. TOWRE PD	.596***	.337**	.428***	.378**	.714***	----						
7. RAN NUM	.004	.055	.136	.051	.133	.370**	----					
8. RAN LETT	.115	.023	.173	.098	.221	.443***	.735***	----				
9. N400 WORD	-.135	-.019	-.097	-.272*	-.139	-.106	.028	.011	----			
10. N400 PW	-.047	.016	-.007	-.235*	-.078	-.056	-.093	-.049	.845***	----		
11. N400 LS	-.010	.071	-.006	-.146	-.057	-.086	-.013	.043	.747***	.759***	----	
12. N400 FF	-.086	.117	-.041	-.152	-.111	-.144	.011	-.035	.654***	.731***	.789***	----
<i>M</i>	114.5	112.5	11.1	124.2	118.0	113.6	106.2	104.3	-4.9	-5.7	-3.1	-3.6
<i>SD</i>	20.7	9.8	2.8	13.0	10.8	13.5	10.3	9.6	3.9	3.8	3.7	3.5

Note. TOC SPACC = Test of Orthographic Competence Spelling Accuracy Composite (Mather et al., 2008); CTOPP PA = Comprehensive Test of Phonological Processing Phonological Awareness Composite (Wagner et al., 1999); CTOPP MD = Comprehensive Test of Phonological Processing Memory for Digits (Wagner et al., 1999); PPVT = Peabody Picture Vocabulary Test IIIA (Dunn & Dunn, 1997); WRMT PC = Woodcock Reading Mastery Test – Revised Passage Comprehension (Woodcock, 1987); TOWRE PD = Test of Word Reading Efficiency Phonemic Decoding Efficiency (Torgesen et al., 1999); RAN NUM = Rapid Automatized Naming and Rapid Alternating Stimulus Tests Numbers (Wolf & Denckla, 2005); RAN LETT = Rapid Automatized Naming and Rapid Alternating Stimulus Tests Letters (Wolf & Denckla, 2005); N400 WORD = mean amplitude of the N400 to words (in μV); N400 PW = mean amplitude of the N400 to pseudowords (in μV); N400 LS = mean amplitude of the N400 to letter strings (in μV); N400 FF = mean amplitude of the N400 to false fonts (in μV).

* $p < .05$,

** $p < .01$,

*** $p < .001$.