

Predicting reading growth with event-related potentials: Thinking differently about indexing “Responsiveness”

Christopher J. Lemons^{a,*}, Alexandra P.F. Key^{b,c}, Douglas Fuchs^{a,c}, Paul J. Yoder^{a,c}, Lynn S. Fuchs^{a,c}, Donald L. Compton^{a,c}, Susan M. Williams^c, Bobette Bouton^a

^a Department of Special Education, Vanderbilt University, United States

^b Department of Hearing and Speech Sciences, Vanderbilt University, United States

^c Vanderbilt Kennedy Center for Research on Human Development, United States

ARTICLE INFO

Article history:

Received 19 January 2009

Received in revised form 13 November 2009

Accepted 21 November 2009

Keywords:

Event-related potential (ERP)

Reading

Reading growth

Response-to-intervention (RTI)

Prediction

ABSTRACT

The purpose of this study was to determine if event-related potential (ERP) data collected during three reading-related tasks (Letter Sound Matching, Nonword Rhyming, and Nonword Reading) could be used to predict short-term reading growth on a curriculum-based measure of word identification fluency over 19 weeks in a sample of 29 first-grade children. Results indicate that ERP responses to the Letter Sound Matching task were predictive of reading change and remained so after controlling for two previously validated behavioral predictors of reading, Rapid Letter Naming and Segmenting. ERP data for the other tasks were not correlated with reading change. The potential for cognitive neuroscience to enhance current methods of indexing responsiveness in a response-to-intervention (RTI) model is discussed.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

The 2004 reauthorization of the Individuals with Disabilities Education Improvement Act (Public Law 108–446; IDEA) describes and expresses a subtle preference for a new method of identifying students with learning disabilities (LD). More specifically, it encourages use of a child's response to research-based, generally effective instruction as a formal part of the disability identification process. This new method is called “Responsiveness-to-Intervention,” or RTI. Disappointingly, little guidance in the IDEA statute or in subsequently published regulations has been given to practitioners to help them proceduralize RTI, in effect permitting them to create their own ways of doing this. Nevertheless, a multi-tiered RTI model has been embraced by researchers and practitioners who seem to agree on its basic components (Fuchs, Mock, Morgan, & Young, 2003; National Joint Committee on Learning Disabilities, 2005).

These components include the provision of research-backed (“core”) instruction in the regular classroom and the use of a screening measure to identify students who will likely struggle to meet normative academic goals by year's end. Students scoring below a certain percentile or performance benchmark on the screening measure are identified as “at-risk.” Their academic progress is monitored in the regular classroom, also referred to as the first tier of instruction, usually for 6–8 weeks.

Those failing to demonstrate appropriate progress are referred to a second, more intensive instructional tier, where their progress is again monitored. Following 8 to 20 weeks (depending on the RTI model), children who continue to struggle are formally identified as non-responders. At this point in the process, a disability may be suspected because of an assumption that students should profit from interventions that have proven effective for a majority of their peers. Again, depending on the RTI model, nonresponders advance to a third and most intensive instructional tier in general education or participate in a comprehensive evaluation to determine if they are disabled and if so whether special education is appropriate.

1.1. A rush to orthodoxy

We and others believe this multi-tier operationalization of RTI holds promise for promoting earlier identification of at-risk children, eliminating poor instruction as the cause of low achievement, and generating data more relevant to instruction than that produced by conventional student evaluations (cf. Fuchs & Fuchs, 2006; Vaughn & Fuchs, 2003; Vellutino et al., 1996). Yet, we also believe that there has been a rush to make this multi-tier version of RTI *the* RTI approach. This premature consensus is unfortunate for several reasons.

A key problem with current thinking about RTI that justifies an exploration of modifications or alternatives is the rigidity in the assignment of children to the instructional tiers. Right now, all students typically move through them in the same way. This uniform movement is driven by a series of decision points: A student receives 6

* Corresponding author. Present Address: Department of Instruction & Learning, University of Pittsburgh, United States.

E-mail address: lemons@pitt.edu (C.J. Lemons).

to 8 weeks of general education to determine whether she is “at-risk.” If so, this is followed by as many as 20 weeks of more intensive instruction to determine whether she is a “nonresponder.” If so, there must then be a determination concerning the necessity of a most intensive level of instruction (Speece, 2005). In this progression from Tier 1 to Tier 2 and so forth, there seems little opportunity to move a child from Tier 1 directly to Tier 3. And yet, it is not hard to imagine a small group of children whose performance is so poor in Tier 1 that to place them in Tier 2 makes little sense.

Similarly, there will surely be a different group of students whose placement in Tier 2 proves unnecessary. That is, they are incorrectly identified in Tier 1 as requiring more intensive instruction and, once in Tier 2, quickly demonstrate a rate of improvement or level of performance that argues for their immediate return to the regular classroom. In short, the inherent inaccuracy in choosing children for the various tiers seems to argue for flexibility, not rigidity, of movement among them. Rigidity of movement contributes to unnecessary delay in placing children in the most appropriate instructional programs.

Although it appears that, according to consensual opinion, RTI should be defined by a multi-tiered approach, very different RTI models can be imagined. In this vein, we should be looking for different or additional methods that may be used to refine the capability of schools to make more accurate and timely distinctions between students who need intensive support and those who do not. Our hope is that the work described in this article—the use of event-related potentials (ERPs) to predict reading growth—will encourage researchers and practitioners to expand their thinking about what RTI may be. Any improvements in the accuracy and efficiency of RTI should decrease the time students wait to receive appropriate instruction or are identified for special education, and increase the likelihood that school resources are used appropriately.

1.2. Cognitive neuroscience and the search for solutions

We acknowledge that cognitive neuroscience may strike some in the field of education as an improbable field in which to search for ways to improve RTI. However, a number of researchers have demonstrated that ERPs can be an effective tool for exploring neural correlates of reading (Maurer & McCandliss, 2007; Molfese et al., 2006; Schlaggar & McCandliss, 2007) and that children with reading disabilities demonstrate different ERP responses from typically developing children (Harter, Anillo-Vento, & Wood, 1989; Molfese, 2000; Penolazzi, Spironelli, Vio, & Angrilli, 2006). In an attempt to expand current thinking about RTI, we decided to explore whether ERPs could be used in a practical sense to provide information that may be used to predict short-term reading change. We hope such thinking may be properly seen as heuristic. Before suggesting how cognitive neuroscience may be applied to RTI, we briefly discuss it in general terms.

Examining brain processes associated with specific behaviors can be very informative by illuminating differences that may be not easily seen using traditional behavioral assessments (Ward, 2006). ERPs offer one means of examining reading-related brain processes in children and adults. ERP is a measure of electrical brain activity associated with a stimulus event (Andreassi, 2000; Lyytinen et al., 2005; Molfese, Molfese, & Espy, 1999). ERPs are represented as complex waveforms of positive and negative deflections that vary in magnitude (amplitude) and temporal features (latencies), and are thought to reflect extent and speed of information processing (Kutas & Federmeier, 1998; Lyytinen et al., 2005; Molfese et al., 1999). One of the main benefits of ERP is their high temporal resolution that allows neural changes to be documented at the millisecond level, at the same rate that reading processes occur (Barber & Kutas, 2007; Molfese, Molfese, & Pratt, 2007). From a practical standpoint, ERPs may be the most feasible imaging technique for use in schools in the future because it is less invasive, less expensive, and requires less specialized laboratory settings than other imaging techniques (Andreassi, 2000).

ERP studies of reading have demonstrated differential patterns of brain behavior for readers of differing reading abilities. For example, Molfese et al. (2006) reported that above-average readers evaluated phonological properties of printed words and nonwords more quickly than the average and below-average readers within 160–400 ms after stimulus onset. Above-average readers also demonstrated higher left-hemisphere amplitudes, whereas average and below-average readers demonstrated higher right-hemisphere amplitudes. Less hemispheric differentiation was associated with weaker reading skills. Hemispheric differences between dyslexic and non-dyslexic children were also reported by Penolazzi et al. (2006) who used a task requiring comparisons of words in terms of phonology, orthography, or semantics. Non-dyslexic readers demonstrated higher left-hemisphere amplitudes and greater hemispheric differentiation when compared to the dyslexic readers associated with phonological elaboration and delayed grapheme–phoneme conversion processes within 370–470 ms and 700–1500 ms after stimulus onset, respectively.

Additional ERP studies comparing reading-disabled children or reading-disabled adults to typical controls have also identified processing deficits that may be contributing to reading difficulties such as auditory/phonological processes (Molfese, Molfese, & Modgline, 2001; Lyytinen et al., 2005; Bonte & Blomert, 2004), visual processes (Barnea, Lamm, Epstein, & Pratt, 1994; Harter, Diering, & Wood, 1988; Lovrich, Cheng, & Velting, 2003), attentional processes (Bernal et al., 2000; Jonkman, Licht, Bakker, & Van Den Broek-Sandmann, 1992), and intermodal timing deficits (Breznitz & Meyler, 2003).

Further, researchers have demonstrated that ERPs can serve as neuropsychological predictors of later language and reading outcomes by examining brain activity associated with reading-related processes such as phonological discrimination. From longitudinal studies, there is evidence that ERP differences among newborns in response to phonological stimuli are predictive of the same children's performance on language tasks at ages 3 and 5, and reading measures at age 8 (Molfese et al., 1999; Molfese et al., 2007; Molfese et al., 2001). Lyytinen et al. (2005) identified differential hemispheric processing between children at-risk for dyslexia and non-risk peers. At-risk infants exhibited right-hemisphere dominance and non-risk children exhibited left-hemisphere dominance for processing auditory information. These differential patterns of ERP response were significantly correlated with poorer receptive language skills at age 2.5, poorer verbal memory skills at age 5, and lower scores on measures of word and nonword reading and reading fluency during the first year of school.

Overall, studies including both reading tasks and tasks requiring the use of reading-related language processes have shown that ERP can provide information about reading processes above and beyond information from behavioral measures (e.g., Molfese et al., 2006), and that readers with differing skill levels exhibit distinct patterns of brain responses (see Lyytinen et al., 2005 for a review). Researchers have also used ERP to make long-term predictions of language and reading skill (e.g., Molfese et al., 2001). However, the use of ERP responses, particularly those collected during reading tasks, to make short-term predictions of reading change and to index responsiveness to academic instruction has not been examined. Additionally, many of the existing ERP studies of reading used well-known ERP tasks that were not designed to target specific reading processes. Examples are the use of “oddball tasks” targeting attentional processes needed for detection of a rare stimulus among frequent distracters (Holcomb, Ackerman, & Dykman, 1986; Bernal et al., 2000), “priming” paradigms where paired stimuli are presented sequentially and properties of the earlier stimulus may be relevant to processing of the later one (McPherson, Ackerman, Oglesby, & Dykman, 1996), and “stop tasks” assessing the executive functioning ability to inhibit an otherwise appropriate response when such a request is unexpectedly presented shortly after the stimulus offset (van der Schoot, Licht, Horsley, & Sergeant, 2002). Attempting to adapt such tasks to assess reading processes could inadvertently lead us to examine latencies and scalp

regions that reflect more global processes such as attention, disinhibition or expectancy rather than more reading-specific processes (e.g., letter to sound conversion).

1.3. Purpose of study

The purpose of the current study was to explore the utility of ERPs to predict academic growth in reading during a single school year in a group of beginning readers. The young age of the children required the use of ERP tasks that would target beginning reading skills. While previous ERP studies have included various tasks exploring reading processes in adults and more experienced readers, limited work has been done with struggling, beginning readers. Therefore, we chose to use a set of ERP tasks that were variants of existing tasks which would tap the same early reading skills (i.e., letter–sound association, simple nonword decoding, rhyming) measured by previously validated behavioral measures, and that would presumably require similar processes to those exercised by children during literacy instruction in kindergarten and first grade.

Another consideration in our task design process was to keep the tasks relatively brief to ensure participants' optimal attention levels and minimal fatigue. Given the novelty of two of the tasks, and the young age and wide range of reading ability in the sample, our approach was conceptually exploratory. That is, we could not predict beforehand where on the scalp or when on the brain wave the ERP variable(s) would predict reading gain. However, we have used a disciplined statistical approach (i.e., omnibus tests to guide subsequent detailed analyses and an adjusted alpha for multiple significance testing) to reduce the probability of detecting sample-specific results.

We hypothesized that (1) these tasks would elicit ERP responses that vary by the children's reading skill and (2) the ERP responses would predict short-term reading growth after controlling for reading performance on relevant behavioral measures. Our hope was that this study would add to previous research in several related ways. First, to use ERPs to predict reading performance during a relatively brief interval, using ecologically valid reading tasks suitable for beginning readers; second, to demonstrate a practical application of ERP as a possible supplemental or alternative method of indexing student reading gain during instruction. Finally, this study was designed as heuristic and to encourage others to think creatively about alternatives when evaluating student responsiveness—alternatives that may eventually improve the efficiency and accuracy of current RTI approaches. At a more general level, we hoped to provide additional evidence that techniques borrowed from cognitive neuroscience can produce rich, unique, and important information as part of educational research.

2. Method

2.1. Participants

Study participants were 29 children (16 females), between the ages of 6 and 8 years ($M=6.92$, $SD=.43$). They were a volunteer sample recruited from 105 first-grade students participating in a larger study aimed at exploring the predictive utility of a dynamic assessment measure (Fuchs, et al., 2007). Students attended 4 elementary schools (two of which were high-poverty Title I) in the Metropolitan Nashville Public Schools. Seventeen students (58.6%) were African American, 10 (34.5%) were Caucasian, 1 (3.5%) was Hispanic, and 1 (3.5%) was Asian; 17 (58.6%) received free or reduced lunch; 3 (10.3%) had previously been retained; and 3 (10.3%) had Individualized Education Plans (IEPs). Of the three children with IEPs, one had a learning disability, one a speech impairment, and one both a learning disability and a speech impairment. All but one participant were right-handed ($M LQ=.73$, $SD=.34$) as determined by Edinburgh Handedness Inventory (Oldfield, 1971). No child was reported by his or her parent to have hearing loss.

All were native English speakers. Parents of all the children provided written informed consent and oral assent was obtained from each child.

2.2. Behavioral procedure

Students participating in the larger study completed a battery of reading assessments. This battery was administered by trained graduate students at the end of the fall semester and again 19 weeks later ($M=18.80$ wks, $SD=.34$).

2.2.1. Behavioral predictors

Two measures administered in the larger study were used as predictors of reading change in this study. Both index skills predictive of early reading growth in first grade (O'Connor & Jenkins, 1999). Students' scores by reading subgroup are displayed in Table 1 (see Section 2.5.3 for additional information on the formation of the subgroups). The *Rapid Letter Naming* (RLN) subtest adapted from the Comprehensive Test of Phonological Processing (CTOPP) was administered to measure the speed with which students could name letters (Wagner, Torgesen, & Rashotte, 1999). Each form consists of 7 rows of letters, with 7 or 8 letters in each row. The letters (a, c, k, n, s, and t) are arranged in random order on each form. The student is asked to name the letters on each page as quickly as possible. If a response is not offered in 3 s, the examiner names the letter and tells the student to move to the next letter. The examiner records the number of letters named correctly in 60 s. The test–retest reliability coefficient for the RLN subtest of the CTOPP is .97 for students 5 to 7 years.

A test of students' word *Segmentation* skills was adapted from the word segmentation subtest of the CTOPP (Wagner, et al., 1999). The test begins with 3 practice items. If the student responds incorrectly to all items, the test is discontinued. If the student responds correctly to at least 1 item, the test is administered until the student misses 4 in succession. The test consists of 22 test items. The score is the total number of sounds segmented correctly in 1 minute. This subtest of the CTOPP has a test–retest reliability of .79 for children between the ages of 8 and 17 in the normative population.

2.2.2. Behavioral outcome

Because the purpose of this study was to determine if ERP could be used to predict reading change over a relatively brief time, *curriculum-based measurement (CBM) word identification fluency* lists were used to measure reading skill. Previous work indicates this measure is a reliable and robust indicator of first-grade reading level and growth (Deno, Mirkin, & Chiang, 1982; Fuchs, Fuchs, & Compton, 2004). Reading change was operationalized as the difference between scores at Time 1 and Time 2. At each time point, the examiner presents two lists of 50 words randomly sampled from high-frequency word lists. The examiner begins with a practice list of 6 words. The student's score is the average number of words read correctly in 1 min on the two lists. Scores are prorated if a student completes reading the list in less than 1 min. Students' scores by reading subgroup are displayed in Table 1.

Table 1
Behavioral measures by reading subgroup.

Measure	Low ($n=10$)		Average ($n=10$)		High ($n=9$)	
	<i>M</i>	(<i>SD</i>)	<i>M</i>	(<i>SD</i>)	<i>M</i>	(<i>SD</i>)
RLN	36.30	(15.35)	41.50	(10.47)	60.78	(15.38)
Segmentation	22.60	(7.78)	24.50	(8.66)	26.00	(3.64)
Time 1 CBM	6.35	(4.67)	14.95	(6.13)	36.39	(12.35)
Time 2 CBM	11.70	(7.00)	29.80	(6.08)	65.00	(12.77)
CBM Change	5.35	(3.32)	14.85	(2.59)	28.61	(12.59)

Note. RLN = Rapid Letter Naming. CBM = Curriculum-Based Measurement.

2.3. ERP procedure

A high-density array of 128 Ag/AgCl electrodes embedded in soft sponges (EGI, Inc., Eugene, OR) was used to record the ERPs. During data collection, all electrodes were referenced to Cz (vertex) and then were re-referenced offline to an average reference. Electrode impedance levels were at or below 40 k Ω (checked before and after testing). The data were sampled at 250 Hz with filters set to .1–30 Hz.

ERPs were recorded at the university ERP lab within 7 weeks ($M = 6.56$ wks, $SD = 2.31$) of each behavioral assessment battery. ERPs were collected twice for each child with Time 1 and Time 2 sessions separated by 14 weeks ($M = 14.69$, $SD = .91$). ERPs were obtained using three measures meant to assess different reading processes: (a) letter sound knowledge, (b) the rhyming of nonsense words, and (c) the decoding of simple nonsense words. The order of presentation of the tasks was fixed across students, starting with the one presumed most easy (Letter Sound Matching) and progressing to the one expected to be most difficult (Nonword Reading). Recording of brainwaves was controlled by Net Station software (v. 4.1; EGI, Inc.). Stimulus presentation was controlled by E-Prime (PST, Inc., Pittsburgh, PA). During the entire test session, the participant's EEG and behavior were continuously monitored and stimulus presentation was suspended during periods of inattention or motor activity.

2.4. ERP tasks

To ensure that children were paying attention and engaging in targeted reading processes, all tasks required participants to evaluate pairs of stimuli (e.g., a printed letter and a spoken sound) and to indicate whether the two were the same or different by pressing buttons on a hand-held response pad. Button assignment to response type was counterbalanced across children. Children were allowed to respond after the second stimulus in a pair as soon as they knew the answer.

Furthermore, because many young children have limited attention spans, we chose to use an equiprobable stimulus presentation where same and different stimulus pairs were presented an equal number of times. The rationale for this choice is that oddball designs necessitate a large number of trials to achieve the desired balance between rare target and frequent distracter trials and thus result in a lengthy recording session. Increasing the duration of the recording increases the likelihood of participant attrition due to fatigue, increased noise in ERPs (e.g., movement and eye artifacts), and loss of focus on the task. Also, proper analysis of an oddball design would require discarding a large number of standard trials to equate the number of trials going into averages for the standard and target conditions to prevent averages from being based on largely uneven trial numbers which could bias the findings toward larger amplitudes being observed for the smaller trial count average (Thomas, Grice, Najm-Briscoe, & Miller, 2004). Equiprobable designs require shorter recording sessions and all recorded data are used for data analysis. Furthermore, in a pilot study in our lab comparing the utility equiprobable and oddball designs for detecting speech sound discrimination, we found the equiprobable design to be no less efficient than an oddball design.

Brief practice sessions were provided for each task. A researcher monitored children's compliance with instructions during the recording session. For the three tasks, auditory stimuli were presented through a speaker positioned above the child's head, at the level of 75 dB SPL(A) as measured at the ear level. Visual stimuli were presented on a 17" monitor positioned 3' in front of the child.

2.4.1. Letter Sound Matching

Stimuli included 7 printed lower-case letters (t, k, n, d, p, g, j) and their corresponding sounds recorded by a male native English speaker. The letters were presented in Century Gothic font (size 96) and their on-screen size was 1" wide by 1 to 1.5" tall. Each trial began

with a 500 ms fixation point (a plus sign) in the center of the screen, followed by a 2000 ms presentation of the printed letter and then a spoken letter sound. On same trials, a printed letter was followed by the correct letter sound. On different trials, a wrong letter sound (with three alternatives used for each letter) was presented following the letter. Participants were asked to press one button if the spoken letter sound matched the printed letter and another if the two did not match. The intertrial interval varied randomly between 1500 and 2500 ms to prevent habituation. There were 84 trials.

2.4.2. Nonword rhyming

Stimuli included 60 pairs of spoken nonwords (selected from the list previously used by Coch, Grossi, Skendzel, & Neville, 2005), where 30 pairs rhymed and 30 did not. Each nonword was included in one rhyming and one non-rhyming pair, and each pair was presented only once during the task resulting in 60 trials. All stimuli were recorded using a male native English speaker. Each trial began with a 500 ms fixation point (a centered plus sign), followed by auditory presentation of the first and second words of a pair. To assist participants' tracking of which words were to be checked for rhyming, numbers "1" and "2" were visually presented 250 ms prior to the onset of each spoken nonword. Participants were instructed to press one button if the two words rhymed and another if the words did not rhyme after they heard the second word. The intertrial interval varied randomly between 1500 and 2500 ms.

2.4.3. Nonword reading

Stimuli included 10 CVC nonwords printed in lower-case (Century Gothic, size 96). Ten spoken nonwords were the same as the CVC nonwords; 10 spoken nonwords differed from the CVC nonwords in the last sound, and an additional 10 spoken nonwords differed from the CVC nonwords in the first sound. Spoken nonwords were recorded by a male native English speaker. Each trial began with a 500 ms fixation point (a centered plus sign). The fixation was replaced by the printed CVC nonword for 2250 ms. The on-screen size of nonwords was 2" to 3" wide x 1" to 1.5" tall. The angle for the visual stimuli was 3.3–4.9° wide by 1.63–2.45° high. On same trials, the printed nonword was followed by the correct spoken nonword (e.g., 'mip' and 'mip'). On different trials, the spoken nonword differed in the first or last sound (e.g., 'mip' paired with 'fip' or 'min'). Participants were asked to press one button if the spoken word matched the printed word and another if the two did not match. Intertrial interval varied randomly between 1500 and 2500 ms to prevent habituation. The task included 80 trials.

2.5. ERP data analysis

2.5.1. Artifact removal

For each of the three tasks, individual ERPs were derived by segmenting the ongoing EEG on auditory stimulus onset to include a 100 ms pre-stimulus baseline and a 600 ms post-stimulus interval. For the Letter Sound Matching task, the trials were segmented on auditory stimulus onset; segments for the other two tasks (i.e., Nonword Rhyming and Nonword Reading) were time-locked to the onset of the second stimulus in a pair. The 600 ms window was expected to capture all of stimulus-related brain activity in children; the window was 100-ms longer than the 500 ms window used in previous work due to the diverse reading ability of our sample (e.g. Coch et al., 2005). Trials with eye and movement artifacts were rejected. Data from electrodes with poor signal quality were replaced using spherical spline interpolation procedures (Srinivasan, Nunez, Tucker, Silberstein, & Cadusch, 1996). Averaged data for each condition were re-referenced to an average reference and baseline-corrected by subtracting the average microvolt value across the 100 ms pre-stimulus interval from the post-stimulus segment. For a data set to be included in the remaining analyses, each stimulus condition had to include a minimum of 10 good

trials. Although this criterion may appear low to some, previous ERP studies with young children demonstrated that 10–15 trials are sufficient to obtain reliable data (e.g., Taylor & Keenan, 1990). Young children can tire and lose interest in a task very quickly, leading to increased movement artifacts and decreased performance due to inattention rather than poor skills. Balancing task length with an optimal number of trials is necessary for any ERP study with children.

Because we used novel tasks with children younger than those typically included in ERP studies of reading problems, statistical analyses, by necessity, had to be exploratory as no a priori predictions regarding a temporal window or scalp location of group differences could reasonably be made. Therefore, we chose to implement a disciplined analytical approach to reduce experiment-wise error.

2.5.2. Data reduction

ERP variables are specified in terms of scalp location and time samples. As a result, ERP tasks generate very large data sets. Without data reduction, there would be 2 conditions \times 124 electrodes (4 lower eye channels were not included in the analysis) \times 150 time samples (600 ms interval sampled every 4 ms), or 14,880 variables per person. Because we did not have peak-specific predictions, using a traditional ERP analysis approach (e.g., peak amplitude/latency measures) was not feasible. It would result in an extremely large number of statistical tests. Therefore, to derive the maximum benefit of this rich data set without arbitrarily selecting just a few data points and discarding the rest, we implemented a spatio-temporal analysis initially described by Spencer, Dien, and Donchin (1999). Specifically, we first reduced the number of electrodes by using spatial principal components analysis (sPCA) with Varimax rotation (using SPSS v.13 package) to identify “virtual electrodes,” or clusters of spatially contiguous electrodes that detect highly inter-correlated data. Amplitude values from electrodes within each cluster were averaged and served as input for the temporal principal components analysis (tPCA) with Varimax rotation aimed at reducing 150 individual time samples to a smaller number of time intervals. The refined factor scores from the tPCA served as ERP variables for further analyses. The PCA analysis approach, including the spatio-temporal method described here, has been widely used in the ERP community for the purposes of data reduction, exploration, and description (Dien & Frishkoff, 2005; Donchin, 1966; van Boxtel, 1998; Kayser & Tenke, 2005).

Both sPCA and tPCA procedures included data from the same and different conditions to provide important contrasting information in order to make the same condition ERP variables more interpretable. Additionally, although our research questions focused on the predictive utility of Time 1 ERPs, data from the Time 2 ERPs were included in the PCA analyses to provide a single data reduction method for the entire ERP data set. As structural changes would not be expected between the time points, the topographical distributions were anticipated to remain stable over time. Thus, including Time 2 ERPs permitted a more reliable estimate of spatial clusters and allowed for an assessment of the stability of the ERP variables correlated with change in reading.

2.5.3. Statistical analysis

Following spatial and temporal data reduction, a series of repeated measures analyses of variance (ANOVA; SPSS v.13) were used to identify which ERP variables (defined by their temporal range and electrode cluster) at Time 1 were most likely to index psycho-physiological processes related to reading. This analytic method required the reading measure to be a categorical variable, and to be entered in the same analysis with the electrode cluster and condition variables. To achieve this, all participants were divided into three reading subgroups (low, average, and high readers; $n = 10, 10$, and 9 , respectively) as defined by the mean, $-.5$ SD and $+.5$ SD on the Time 2 CBM reading test. Children with IEPs were included in the low ($n = 2$) and average ($n = 1$) subgroups. Statistically significant differences were found between the

groups for RLN ($p < .01$) and all CBM measures ($p < .001$), but not for Segmentation ($p = .59$) or age ($p = .41$). This admittedly arbitrary division resulted in loss of information about individual differences in reading, but provided a conservative test of whether ERP variables were related to reading while providing protection against experiment-wise error.

Next, the repeated measures ANOVAs, which included data from the same and different conditions, permitted exploration of the possible presence of any reliable reading group \times electrode cluster \times condition effects (using an adjusted alpha of .01). These analyses helped identify time periods relevant to indexing a process that varied by reading skill. Main effects were not evaluated following nonsignificant interactions as our purpose was to identify ERP responses predictive of reading change. In the case of a significant interaction, a one-way ANOVA with reading group as the between-subject factor and condition difference (different–same) as the dependent variable helped identify the cluster at which ERP responses to condition contrast differed as a function of reading skill. The correlations between these ERP responses and reading change (i.e., the CBM difference score between Times 1 and 2) were next examined to test our hypothesis that the ERP responses would predict short-term reading change. In cases where the selected ERP variable(s) were statistically significantly correlated with reading change, we examined the predictive utility of ERP by using regression with reading change as the criterion (dependent) variable and the Time 1 ERP as the predictor variable while controlling for Time 1 accuracy of behavioral responding during the selected ERP task.

3. Results

3.1. Behavioral performance

To interpret ERP response as an index of the cognitive processes involved in a specified task, and to legitimately use brainwaves as predictors of future outcomes, we must have confidence that participants were indeed performing the designated task (e.g., comparing visually presented and spoken stimuli). One obvious way to verify that participants were discriminating same from different stimuli is to look at the accuracy of their responses. In other words, because ERPs associated with *incorrect* responses may reflect processes in addition to or instead of those targeted by the task (e.g., inattention), one may choose to focus ERP analyses on trials with only correct responses. This solution, however, is problematic with young children as participants because it is often difficult to obtain from them an adequate number of correct, artifact-free responses when using a task targeting a developing skill and keeping the testing session relatively brief. Hence, we analyzed all ERP data regardless of the accuracy of children's responses, thereby providing a conservative test of our hypotheses.

Variation in accuracy of performance (i.e., the proportion of correct button presses) at Time 1 and Time 2 was examined with hierarchical linear regression (see Table 2). Results indicate that accuracy did not vary by type (e.g., same vs. different, $t(304) = -0.84, p = .40$); however, statistically significant differences were found across tasks ($t(304) = -3.36, p = .001$) and time ($t(304) = 3.08, p = .003$). Follow up analyses indicated higher levels of accuracy on Letter Sound Matching compared to Nonword Reading ($p = .001$) and Nonword Rhyming ($p = .005$) with no differences between Nonword Reading and Nonword Rhyming ($p = .70$). Overall, students were more accurate on all tasks at Time 2 ($p = .003$). Due to a statistically significant interaction of group with task ($p = .000$), follow up one-way analyses of variance were conducted to examine differences between the groups for each task at each time. Significant group differences were found only for Nonword Rhyming at Time 1 ($F(2,19) = 9.80, p = .001$) and Time 2 ($F(2,19) = 3.87, p = .04$). Post hoc analyses indicated that at Time 1, Low readers were less accurate than Average ($p = .012$) and High ($p = .001$) readers with no differences found between the latter two groups ($p = .86$). At Time 2, Low readers were less accurate than Average

Table 2
Accuracy rates for the ERP tasks by task, time, and reading subgroup.

ERP task	Time 1		Time 2	
	M	(SD)	M	(SD)
Letter Sound Matching				
Low (<i>n</i> = 10)	0.82	(0.14)	0.91	(0.10)
Average (<i>n</i> = 8)	0.80	(0.10)	0.94	(0.04)
High (<i>n</i> = 8)	0.85	(0.10)	0.80	(0.35)
Nonword Rhyming				
Low (<i>n</i> = 7)	0.60	(0.13)	0.79	(0.12)
Average (<i>n</i> = 8)	0.79	(0.09)	0.93	(0.03)
High (<i>n</i> = 7)	0.85	(0.12)	0.90	(0.13)
Nonword Reading				
Low (<i>n</i> = 10)	0.66	(0.17)	0.82	(0.22)
Average (<i>n</i> = 5)	0.75	(0.17)	0.99	(0.01)
High (<i>n</i> = 7)	0.80	(0.16)	0.91	(0.15)

($p = .044$) readers; no statistically significant differences were found when comparing the remaining groups. Additionally, CBM scores at Time 2 were correlated with reading change (i.e., CBM difference): .80 (Nonword Reading and Rhyming) and .90 (Letter Sound Matching).

3.2. ERP results

3.2.1. Letter Sound Matching

Data from 3 participants (1 female, 2 males) were excluded due to an excessive number of artifacts in the ERP data (at least one condition had less than 10 clean trials). For the remaining 26 participants, numbers of artifact-free trials were comparable across stimulus conditions (M same = 26.54 \pm 6.97; M different = 26.46 \pm 7.42 trials). This number of trials is consistent with the typical trial retention in ERP studies with children where averages are based on 15–20 trials (e.g., Bernal et al., 2000; Khan, Frisk, & Taylor, 1999; Silva-Pereyra et al., 2003). There were no reading group differences in the number of trials retained for analysis ($p > .05$). Furthermore, the

number of retained trials did not correlate with the behavioral accuracy or CBM scores (p 's $> .10$).

The sPCA resulted in 5 factors accounting for 80.35% of the total variance. These factors corresponded to 7 electrode clusters (see Fig. 1) that included 98 of 124 electrodes (79% of the net). The tPCA resulted in 5 factors that accounted for 80.87% of the total variance. These factors divided the waveform into intervals corresponding to 0–88 ms (Factor 4), 104–232 ms (Factor 2), 240–344 ms (Factor 3), 368–416 ms (Factor 5), and 408–600 ms (Factor 1). At Time 1, a statistically significant reading group \times electrode cluster \times condition effect was found only for the interval latest in temporal sequence (408–600 ms; $F(12,138) = 3.285$, $p = .005$, partial eta squared = .222). A one-way ANOVA with reading group as the between-subject factor and the difference in the refined factor scores for the same-different conditions as the dependent variable was used to identify the electrode cluster at which ERP responses differed as a function of reading. Statistically significant group effects were found at the frontal cluster ($F(2,23) = 8.628$, $p = .002$, partial eta squared = .429) and posterior cluster ($F(2,23) = 4.575$, $p = .021$, partial eta squared = .285).

Therefore, Time 1 ERP responses to the same conditions for these two effects (i.e., those occurring latest in time over frontal and posterior scalp locations) were selected as putative predictors of reading change. Intercorrelations between behavioral measures (segmentation, rapid letter naming, CBM at Times 1 and 2, reading change) and ERP predictors (posterior and anterior ERPs at Time 1) are presented in Table 3. Time 1 late ERPs for the same condition over the posterior scalp electrode cluster (See Fig. 2) were correlated with reading change ($r = .481$, $p = .013$). When these Time 1 ERPs were regressed onto reading change while controlling for Time 1 accuracy, ERPs were a significant contributor to predicting reading change (R^2 change = .222; $F(1,23) = 6.683$, $p = .017$).

Contributions of the posterior ERPs at Time 1 to the prediction of reading change remained significant after controlling for each of two commonly used behavioral predictors of reading change: (a) Time 1 Rapid Letter Naming scores (R^2 change = .17; $t(23) = 2.53$, $p = .019$) and (b) Segmentation scores (R^2 change = .25; $t(23) = 2.866$, $p = .009$).

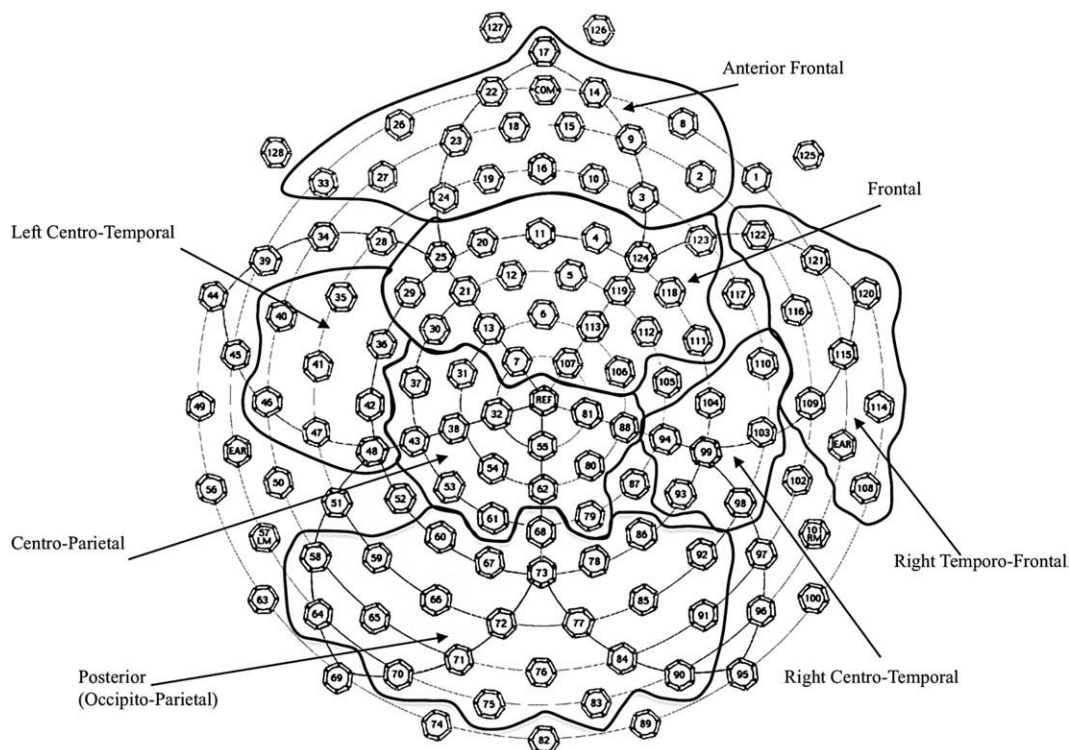


Fig. 1. Electrode clusters identified by sPCA for the Letter Sound Matching task.

The correlation between the late ERP variable at the frontal cluster and reading change was not significant ($r = -.325, p = .105$).

Additionally, we examined the stability of the late occurring ERP to letter comparison at the posterior cluster over time (14 weeks), by calculating the intra-class correlation (ICC) for this variable between Time 1 and Time 2 ($ICC = .612, p < .001$; confidence interval $\text{min} = .301, \text{max} = .805$). This degree of stability compares favorably to the stability of ERP variables reported in the extant literature.

3.2.2. Nonword rhyming

Data from 7 participants (4 females, 3 males) on this measure were excluded due to excessive noise in the ERP data. For the remaining 22 children, numbers of artifact-free trials were comparable across stimulus conditions ($M \text{ same} = 19.09 \pm 1.19$; $M \text{ different} = 20.50 \pm 2.97$ trials). The sPCA resulted in 7 factors accounting for 81.98% of the total variance. These factors corresponded to 8 electrode clusters that included 89 of 124 electrodes (72% of the net). The tPCA focused on the 100–600 ms post-stimulus interval resulted in 5 factors that accounted for 90.857% of the total variance. These factors divided the waveform into intervals corresponding to 100–144 ms (Factor 5), 152–256 ms (Factor 2), 280–328 ms (Factor 4), 336–472 ms (Factor 3), and 448–600 ms (Factor 1). No statistically significant effects ($p < .05$) were obtained involving reading group \times condition variables. Therefore, putative predictors of reading change were not evaluated.

3.2.3. Nonword reading

Data from 7 participants (4 females, 3 males) were excluded due to excessive noise. For the remaining 22 children, numbers of artifact-free trials were comparable across stimulus conditions ($M \text{ same} = 19.45 \pm 2.18$; $M \text{ different} = 19.73 \pm 1.58$ trials). The sPCA resulted in 7 factors accounting for 79.50% of the total variance. These factors corresponded to 7 electrode clusters that included 77 of 124 electrodes (62% of the net). The tPCA resulted in 5 factors that accounted for 83.35% of the total variance. These factors divided the waveform into intervals corresponding to 0–64 ms (Factor 5), 80–176 ms (Factor 4), 184–312 ms (Factor 3), 312–464 ms (Factor 1), and 464–600 ms (Factor 2). A statistically significant reading group \times condition interaction ($F(2,19) = 7.436, p = .004$, partial eta squared = .439) occurred for the late component (464–670 ms). A one-way ANOVA with reading group as the between-subject factor and condition difference (different–same) as the dependent variable resulted in a marginally statistically significant effect ($F(2,21) = 3.419, p = .054$, partial eta squared = .265). Time 1 ERPs for the same condition at this time period did not correlate with the change in reading ($r = .278, p = .211$). Due to this lack of correlation, further regression analyses were not conducted.

4. Discussion

The purpose of this study was to determine if the ERP data from three reading-related tasks could be used to predict short-term reading change. We hypothesized that Time 1 ERP responses to all

Table 3
Intercorrelations among behavioral measures and ERP predictors for Letter Sound Matching task.

Variable	1	2	3	4	5	6	7
1. RLN	–						
2. Segmentation	–.04	–					
3. Time 1 CBM	.54	.18	–				
4. Time 2 CBM	.56	.22	.92	–			
5. Reading Change	.47	.21	.67	.91	–		
6. Posterior ERP Time 1 Match	.16	–.09	.27	.41	.48	–	
7. Anterior ERP Time 1 Match	–.35	.02	–.30	–.34	–.33	–.77	–

Note. Correlations above .39 were significant at $p < .05$, above .50 were significant at $p < .01$, and above .61 were significant at $p < .001$ (two-tailed). RLN=Rapid Letter Naming. CBM=Curriculum-Based Measurement.

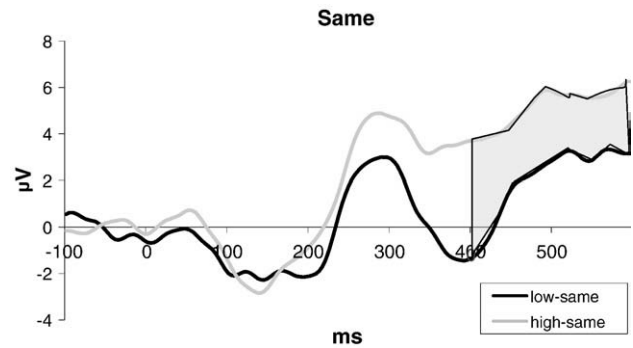


Fig. 2. Time 1 late ERPs for the same condition of the Letter Sound Matching task over the posterior scalp electrode cluster for children with low- and high-reading changes.

three tasks would be correlated with reading change over 19 weeks. Results indicated that late ERP responses to the Letter Sound Matching task over posterior sites were reliably predictive.

The late temporal window where we found the predictive effects in the letter-sound matching task may be viewed by some as delayed compared to findings from other ERP studies of reading. Previous research has typically identified differences earlier in the wave and attributed them to alterations in sensory processing (e.g., N1–P2 effects of Molfese et al. (2007)) or in integration of immediate context (e.g., N2/N400 effects in priming/sentence reading tasks—McPherson et al., 1996; Brandeis, Vitacco & Steinhausen, 1994; Neville, Coffey, Holcomb, & Tallal, 1993). In the current study, the predictive effect occurs in the 400–600 ms window where higher positive amplitudes were observed for children with better reading scores at Time 2. This effect overlaps in time and space with the memory-related effects observed in previous studies of recognition and recall in adults (e.g., Wilding & Rugg, 1997; Donaldson & Rugg, 1998) and suggests that early reading acquisition may depend heavily on the ability to recall specific information, such as letter–sound associations, rather than merely recognizing familiar pairings among distracters.

Furthermore, ERP responses on Letter Sound Matching remained a statistically significant predictor after controlling for two well-regarded reading measures, Rapid Letter Naming and Segmenting. In other words, ERP had a value added in predicting which students would and would not demonstrate reading growth. The Nonword Rhyming and Nonword Reading tasks were not predictive of short-term reading change in our sample.

4.1. Importance of the study

We believe these findings extend prior related work in several ways. First, we demonstrated that an ERP reading task, Letter Sound Matching, can help predict short-term reading change, and that the ERP responses provided predictive information beyond that of two well-regarded, commonly used behavioral measures. In prior studies, by contrast, researchers have used ERPs to make successful long-term (i.e., 5 to 8 year) predictions of reading performance (Lyytinen, et al., 2005; Molfese, et al., 2007). Whereas long-term predictions can have theoretical and practical importance, their accuracy may decrease as more intensive interventions are provided to at-risk students. Additionally, teachers and administrators typically must determine students' instructional needs in the here and now (i.e., during the current academic year), and therefore would benefit from measures sensitive to ongoing changes in student performance. Thus, accurate short-term predictions are necessary.

Second, results demonstrate the importance of task design. We developed our three tasks to be closely related to critical first-grade reading skills and to be aligned with validated behavioral measures. However, ERP responses to the Nonword Rhyming task—the task that seems to reflect the early reading skill of phonological awareness—were

not predictive of short-term reading change. Similarly, we did not obtain statistically significant predictions of change by using the Nonword Reading task, which seems most closely related to our word identification fluency outcome. Thus, our findings suggest that merely “translating” behavioral measures into psycho-physiological tasks may not always be fruitful. The most likely explanation for why these two tasks did not predict reading change is that they were too difficult at Time 1 (behavioral data indicate substantial improvement in performance at Time 2) and thus did not provide the needed measure of brain processes associated with these skills.

By contrast, the task with greatest predictive utility (e.g., Letter Sound Matching) was the one on which most students did well at Time 1. Similarly, researchers in prior studies in which ERP tasks have successfully differentiated group performance, or predicted behavioral outcomes, have reported high levels of behavioral accuracy, ranging from 87% to 97% (c.f., Penolazzi, et al., 2006; Yang, Perfetti, & Schmalhofer, 2007). Thus, an easier ERP task may be more informative about future outcomes than a more difficult one.

In terms of RTI, we have provided an “outside-the-box” example of how to think about predicting short-term reading change because we believe that current approaches are neither sufficiently effective nor efficient. Which is to say, we believe we and others can do better. Hence, we are obligated to continue exploring supplemental or alternative approaches. Our findings, however preliminary, suggest that cognitive neuroscience techniques like ERP may be one such approach.

4.2. Study limitations

Before discussing future research directions, we would like to note several limitations of the current study. First, we wish to highlight the difference between predicting change and predicting response to intervention (see Yoder & Compton, 2004). This study was not an evaluation of an RTI model, and the “response” used as an outcome variable was change in reading score that may or may not have been due to the instruction each child received. Indeed, we have little understanding of the nature of instruction directed at our 29 children in their regular classrooms. Had we greater understanding of this instruction, we would still have cause to be skeptical because interpreting predictors of change within the treatment group as reflecting response to treatment is tantamount to claiming that *all* change is due to treatment. This is probably not so.

Furthermore, it is likely that the predictive utility of ERP tasks (and behavioral assessments) will change depending upon the amount and quality of instruction provided. Similarly, the selection of outcome variable may change the predictive utility of ERP tasks. In other words, our Nonword Reading and Nonword Rhyming tasks may have resulted in statistically significant predictions if a measure other than CBM was selected as the outcome variable.

Some readers may wonder about the quality of data averaged across 26 trials per condition. First, this number of trials is more typical than not for ERP studies involving young children (cf. Taylor & Keenan, 1990). Second, an insufficient number of trials would produce random, not correlated, measurement error, making it more difficult to detect significant differences (i.e., increasing type II, not type I, error). Put differently, for this reason, we consider our study to be a conservative test of ERP utility. Furthermore, the Time 2 ERP recoding provided for a test–retest opportunity. The intra-class correlation between Time 1 and Time 2 for the ERP variable reflected high stability, something unlikely to occur if the data contained large amounts of noise. Nevertheless, a replication study is needed to validate our findings.

Finally, our analysis represented just one of a vast array of possible methods to identify relevant ERP variables. Had we identified these by another method, results may have been different. We emphasize that this was a pilot study with relatively few participants and that replication is necessary before conclusions may be made regarding

the predictive value of these tasks or the related ERP components for other students and outcome measures.

4.3. Future directions

This study has provided preliminary evidence that cognitive neuroscience techniques like ERP may have practical application in school settings. Although it is unlikely that we will soon leave “no child unscanned,” it is likely that technological advances will continue, increasing the likelihood that practitioners will have access to techniques like ERP in the future. One can imagine that, in 10 to 20 years, schools may be using imaging technology to help efficiently and effectively determine a most appropriate level of intervention for children. It is possible that ERP tasks like the ones used in this study could be used as alternate or supplemental methods to determine which students need to quickly advance to the most intensive level of instruction and which students do not need this level of intensity to make academic progress. It is also likely that cognitive neuroscience will continue to provide additional insights into the best ways to remediate nonresponsive children by expanding our understanding of the processing that underlies nonresponsiveness and the changes in this processing that occur when effective interventions are provided.

In the field of learning disabilities, more researchers should focus on the use of neuroscience approaches to understand processing differences between students who learn to read easily and those who do not. This information likely will lead to the development of improved behavioral assessments and academic interventions. Furthermore, as the application of neuroscience techniques to educational research is new, much work is needed in the development and replication of effective, reliable ERP tasks, and a better understanding of what cognitive processes are measured by which tasks (Barber & Kutas, 2007). Additionally, future studies should examine the predictive utility of ERP tasks across a range of behavioral outcome measures to determine which ERP tasks are most relevant to meaningful academic growth.

Finally, researchers should continue examining the practical application of neuroscience techniques into systems of identification and intervention evaluation. This study was an initial exploration of this type of work, but more is needed. Future studies could be conducted in which validated neuroscience screening methods could be used to index student responsiveness in an RTI framework and to evaluate whether efficiency and accuracy of the model is enhanced with these methods. Similarly, this type of technology could be used to evaluate brain-based changes that occur due to specific interventions. Cognitive neuroscience techniques, like ERP, are likely to expand our current understanding of learning disabilities and best ways to address the needs of academically vulnerable students. We encourage greater collaboration between educational researchers and cognitive neuroscientists to further this exploration.

Acknowledgements

This research was supported in part by the National Research Center on Learning Disabilities, Grant H324U010004, from the Office of Special Education Programs in the U.S. Department of Education; the National Institute of Child Health and Human Development Core Grant HD15052; and the Experimental Education Research Training (ExpERT) predoctoral training program, Grant R305B04110 from the Institute of Educational Sciences to Vanderbilt University. Statements do not necessarily reflect the positions or policies of these agencies, and no official endorsement by them should be inferred.

References

- Andreassi, J. L. (2000). *Psychophysiology: Human behavior and physiological response*, 4 ed. Mahwah, NJ: Lawrence Erlbaum.

- Barber, H. A., & Kutas, M. (2007). Interplay between computational models and cognitive electrophysiology in visual word recognition. *Brain Research Reviews*, 53, 98–123.
- Barnea, A., Lamm, O., Epstein, R., & Pratt, H. (1994). Brain potentials from dyslexic children recorded during short-term memory tasks. *International Journal of Neuroscience*, 74(1–4), 227–237.
- Bernal, J., Harmony, T., Rodríguez, M., Reyes, A., Yáñez, G., Fernández, T., et al. (2000). Auditory event-related potentials in poor readers. *International Journal of Psychophysiology*, 36, 11–23.
- Bonte, M. L., & Blomert, L. (2004). Developmental dyslexia: ERP correlates of anomalous phonological processing during spoken word recognition. *Cognitive Brain Research*, 21, 360–376.
- Brandeis, D., Vitacco, D., & Steinhausen, H. C. (1994). Mapping brain electric micro-states in dyslexic children during reading. *Acta Paedopsychiatrica*, 56, 239–247.
- Breznitz, Z., & Meyler, A. (2003). Speed of lower-level auditory and visual processing as a basic factor in dyslexia: Electrophysiological evidence. *Brain & Language*, 85, 166–184.
- Coch, D., Grossi, G., Skendzel, W., & Neville, H. (2005). ERP nonword rhyming effects in children and adults. *Journal of Cognitive Neuroscience*, 17, 168–182.
- Deno, S. L., Mirkin, P. K., & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children*, 49, 36–45.
- Dien, J., & Frishkoff, G. (2005). Principal components analysis of the ERP data. In T. C. Handy (Ed.), *Event-related potentials: A methods handbook* (pp. 189–207). Cambridge, MA: MIT Press.
- Donaldson, D., & Rugg, M. (1998). Recognition memory for new associations: Electrophysiological evidence for the role of recollection. *Neuropsychologia*, 36, 377395.
- Donchin, E. (1966). A multivariate approach to the analysis of average evoked potentials. *IEEE Transactions on Biomedical Engineering*, 13, 131–139.
- Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly*, 41, 93–99.
- Fuchs, D., Fuchs, L. S., Compton, D. L., Bouton, B., Caffrey, E., & Hill, L. (2007). Dynamic assessment as responsiveness to intervention: A scripted protocol to identify young at-risk readers. *Teaching Exceptional Children*, 39, 58–63.
- Fuchs, L. S., Fuchs, D., & Compton, D. (2004). Monitoring early reading development in first grade: Word identification fluency versus nonsense word fluency. *Exceptional Children*, 71, 7–21.
- Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research & Practice*, 18, 157–171.
- Harter, M. R., Anllo-Vento, L., & Wood, F. B. (1989). Event-related potentials, spatial orienting, and reading disabilities. *Psychophysiology*, 26, 404–421.
- Harter, M. R., Diering, S., & Wood, F. B. (1988). Separate brain potential characteristics in children with reading disability and attention deficit disorder: Relevance-independent effects. *Brain and Cognition*, 7, 54–86.
- Holcomb, P. J., Ackerman, P. T., & Dykman, R. A. (1986). Auditory event-related potentials in attention and reading disabled boys. *International Journal of Psychophysiology*, 3, 263–273.
- Jonkman, I., Licht, R., Bakker, D. J., & Van Den Broek-Sandmann, T. M. (1992). Shifting of attention in subtyped dyslexic children: An event-related potential study. *Developmental Neuropsychology*, 8, 243–259.
- Kayser, J., & Tenke, C. (2005). Trusting in or breaking with convention: Towards a renaissance of principal components analysis in electrophysiology. *Clinical Neurophysiology*, 116, 1747–1753.
- Khan, S. C., Frisk, V., & Taylor, M. J. (1999). Neurophysiological measures of reading difficulty in very-low-birthweight children. *Psychophysiology*, 36, 76–85.
- Kutas, M., & Federmeier, K. D. (1998). Minding the body. *Psychophysiology*, 35, 135–150.
- Lovrich, D., Cheng, J., & Velting, D. (2003). ERP correlates of form and rhyme letter tasks in impaired reading children: A critical evaluation. *Child Neuropsychology*, 9, 159–174.
- Lyytinen, H., Guttorm, T. K., Huttunen, T., Hamalainen, J., Leppanen, P. H. T., & Vesterinen, M. (2005). Psychophysiology of developmental dyslexia: A review of findings including studies of children at risk for dyslexia. *Journal of Neurolinguistics*, 18, 167–195.
- Maurer, U., & McCandliss, B. D. (2007). The development of visual expertise for words: The contribution of electrophysiology. In E. L. Grigorenko, & A. Naples (Eds.), *Single-word reading: Cognitive, behavioral and biological perspectives* (pp. 43–64). Mahwah, NJ: Lawrence Erlbaum.
- McPherson, W. B., Ackerman, P. T., Oglesby, D. M., & Dykman, R. A. (1996). Event-related brain potentials elicited by rhyming and non-rhyming pictures differentiate subgroups of reading disabled adolescents. *Integrative Psychological and Behavioral Science*, 31, 3–17.
- Molfese, D. L. (2000). Predicting dyslexia at 8 years of age using neonatal brain responses. *Brain and Language*, 72, 238–245.
- Molfese, D. L., Key, A. F., Kelly, S., Cunningham, N., Terrell, S., Ferguson, M., et al. (2006). Below-average, average, and above-average readers engage different and similar brain regions while reading. *Journal of Learning Disabilities*, 39, 352–363.
- Molfese, D. L., Molfese, V. J., & Espy, K. A. (1999). The predictive use of event-related potentials in language development and the treatment of language disorders. *Developmental Neuropsychology*, 16, 373–377.
- Molfese, D. L., Molfese, V. J., & Pratt, N. L. (2007). The use of event-related evoked potentials to predict developmental outcomes. In M. de Haan (Ed.), *Infant EEG and event-related potentials* New York: Psychology Press.
- Molfese, V. J., Molfese, D. L., & Modgline, A. A. (2001). Newborn and preschool predictors of second-grade reading scores: An evaluation of categorical and continuous scores. *Journal of Learning Disabilities*, 34, 545–554.
- National Joint Committee on Learning Disabilities. (2005). Responsiveness to intervention and learning disabilities. Bethesda, MD: Author. Retrieved from: http://www.lidonline.org/pdf/rti_final_august_2005.pdf
- Neville, H., Coffey, S., Holcomb, P., & Tallal, P. (1993). The neurobiology of sensory and language processing in language-impaired children. *Journal of Cognitive Neuroscience*, 5, 235–253.
- O'Connor, R. E., & Jenkins, J. R. (1999). Prediction of reading disabilities in kindergarten and first grade. *Scientific Studies of Reading*, 3, 159–197.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97–113.
- Penolazzi, B., Spironelli, C., Vio, C., & Angrilli, A. (2006). Altered hemispheric asymmetry during word processing in dyslexic children: An event-related potential study. *Cognitive Neuroscience and Neuropsychology*, 17, 429–433.
- Schlaggar, B. L., & McCandliss, B. D. (2007). Development of neural systems for reading. *Annual Review of Neuroscience*, 30, 475–503.
- Silva-Pereyra, J., Rivera-Gaxiola, M., Fernández, T., Díaz-Comas, L., Harmony, T., Fernández-Bouzas, A., et al. (2003). Are poor readers semantically challenged? An event-related brain potential assessment. *International Journal of Psychophysiology*, 49, 187–199.
- Speece, D. L. (2005). Hitting the moving target known as reading development: Some thoughts on screening children for secondary interventions. *Journal of Learning Disabilities*, 38, 487–493.
- Spencer, K., Dien, J., & Donchin, E. (1999). A componential analysis of the ERP elicited by novel events using a dense electrode array. *Psychophysiology*, 36, 409–414.
- Srinivasan, R., Nunez, P. L., Tucker, D. M., Silberstein, R. B., & Cadusch, P. J. (1996). Spatial sampling and filtering of EEG with spline-laplacians to estimate cortical potentials. *Brain Topography*, 8, 355–366.
- Taylor, M. J., & Keenan, N. K. (1990). Event-related potentials to visual and language stimuli in normal and dyslexic children. *Psychophysiology*, 27, 318–327.
- Thomas, D. G., Grice, J. W., Najm-Briscoe, R. G., & Miller, J. W. (2004). The influence of unequal numbers of trials on comparisons of average event-related potentials. *Developmental Neuropsychology*, 26, 753–774.
- van Boxtel, G. J. M. (1998). Computational and statistical methods for analyzing event-related potential data. *Behavioral Research Methods, Instruments, and Computers*, 30, 87–102.
- van der Schoot, M., Licht, R., Horsley, T. M., & Sergeant, J. A. (2002). Fronto-central dysfunctions in reading disability depend on subtype: Guessers but not spellers. *Developmental Neuropsychology*, 22, 533–564.
- Vaughn, S., & Fuchs, L. S. (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research & Practice*, 18, 137–146.
- Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., et al. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology*, 88, 601–638.
- Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive test of phonological processing*. Austin, TX: PRO-ED, Inc.
- Ward, J. (2006). *The student's guide to cognitive neuroscience*. New York: Psychology Press.
- Wilding, E. I., & Rugg, M. D. (1997). Event-related potential and the recognition memory exclusion task. *Neuropsychologia*, 35, 119–128.
- Yang, C. L., Perfetti, C. A., & Schmalhofer, R. (2007). Event-related potential indicators of text integration across sentence boundaries. *Journal of Experimental Psychology*, 33, 55–89.
- Yoder, P., & Compton, D. (2004). Identifying predictors of treatment response. *Mental Retardation and Developmental Disabilities*, 10, 162–168.