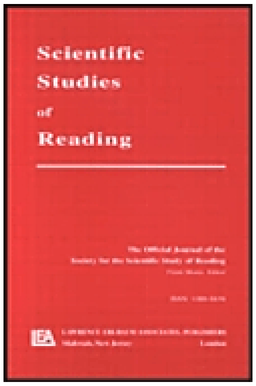


Letter-sound integration is modulated by automaticity demands and is related to reading performance in English-speaking children

Journal:	*Scientific Studies of Reading*
Manuscript ID	Draft
Manuscript Type:	Research Article
Keywords:	automaticity, letter knowledge, letter-sound integration, reading ability

Letter-sound integration is modulated by automaticity demands and is related to reading performance in English-speaking children

Abstract

Letter-sound integration, the ability to accurately and quickly associate letters and their corresponding speech sounds, has been suggested as an independent predictor of skilled reading, but few studies have examined how individual differences in letter-sound integration relate to reading. This study assessed letter-sound integration in school-age English-speaking children (*n* = 67, ages 8-11, 31% with dyslexia) via a behavioral letter-sound matching task. The task included conditions in which the presentation of visual letter and auditory sound stimuli pairs was synchronous or was offset by up to 400ms. Children's letter-sound integration across levels of temporal asynchrony was significantly correlated with measures of reading ability. Performance in the synchronous condition, which placed the largest demands on automaticity, accounted for variance in reading ability beyond measures of IQ, phonological awareness, and RAN. These findings suggest that letter-sound integration efficiency plays a unique role in children's reading performance.

When children learn to read, they must explicitly learn the correspondences between speech sounds (phonemes) and associated letters (graphemes). As reading development progresses, children typically achieve accuracy, then continue to build speed of letter-sound associations, reaching automaticity, when little or no explicit effort is required to make these associations (Ehri, 2005; Ehri & Wilce, 1983; Frost, 1998; LaBerge & Samuels, 1974; Snowling, 1980). This efficient and automatic association of letters and sounds is called letter-sound integration (Blomert, 2011). Some individuals, however, fail to achieve letter-sound integration, remaining inaccurate in matching letters and sounds or being accurate but slow and not automatic (e.g., Aravena et al., 2013; Blau et al., 2009; Blomert & Willems, 2010; Froyen et al., 2011). It has been established that letter name and letter sound knowledge accuracy account for unique variance in early reading ability (e.g., Caravolas et al., 2012; McBride-Chang, 1999), but it is not known whether letter-sound integration (which takes automaticity in to account) plays a unique role in reading ability. A deficit in letter-sound integration has been suggested as one potential mechanism underlying developmental dyslexia (Blomert, 2011; Blomert & Froyen, 2010). The variety of letter-sound integration tasks and different populations and orthographies studied, however, limits our understanding of how automaticity of letter-sound integration and its relationship with reading.

Letter-sound knowledge and reading ability

Letter-sound knowledge refers to the accuracy of associating printed letters with the speech sounds they represent (Snowling, 1980). In various orthographies such as English, Spanish, Slovak, and Czech, children's letter-sound knowledge in preschool predicts their word reading and spelling performance in first grade (Caravolas et al., 2012). This is also true for children with familial risk of dyslexia in English and in Dutch (Blomert & Willems, 2010; Hulme et al., 2015).

Though letter-sound knowledge is important for reading development, accuracy measures of this skill do not give substantial insight into individual differences in later, more skilled reading abilities. The association of letter-sound knowledge and reading ability is strongest at the age when children do not yet know all of their letters; the set of associations to be learned is constrained (limited to the number of letters in the language), and so children typically quickly reach ceiling in accuracy (Paris, 2005). Letter name and sound knowledge accuracy measures in kindergarten do not substantially predict 2nd grade reading beyond the contribution of phonological awareness and RAN (Schatschneider et al., 2004). However, these letter accuracy measures may be less related to later reading fluency measures, and 2nd graders may not have achieved automaticity. This is consistent with the notion that building automaticity in sublexical skills such as letter-sound associations supports more advanced fluent reading and comprehension (Norton & Wolf, 2012; Wolf & Katzir-Cohen, 2001).

Letter-sound integration and reading ability or dyslexia status

Letter-sound integration builds on accurate letter-sound knowledge. Several studies in the relatively transparent Dutch orthography (i.e. letters and speech sounds are reliably matched) support the notion that accuracy of letter-sound knowledge is intact in poor readers, but that letter-sound integration (measured in terms of speed or efficiency) is impaired, on average. When matching letters and sounds, Dutch-speaking adults and children with dyslexia had comparable accuracy to typical-reading peers, but were significantly slower (Blau et al., 2009; 2010). In another study, Dutch -speaking children 7-12 years old with and without dyslexia were trained to associate novel letters (Hebrew graphemes) with Dutch speech sounds and were trained under explicit (no time pressure), implicit (time pressured), or combined, conditions (Aravena et al., 2013). Both groups became accurate in associating the sounds and letters after training. Children with dyslexia, however, made more errors than typical readers during the time-pressured training conditions. Thus, letter-sound integration deficits in poor readers may manifest during tasks that require speed or have a time pressure element. This is perhaps not surprising, as it has been determined that performance is slower in the absence of automaticity or when automaticity is heavily taxed because of time pressure (Kail & Hall, 1994; Logan, 1992). Accordingly, we suggest that tasks that measure accuracy without a speed or time pressure element are likely not sensitive enough to differentiate individuals who have automaticity deficits from those whose automaticity is intact.

In contrast, studies in English, a comparatively opaque alphabetic orthography, have found that letter-sound integration is only weakly related to reading performance and dyslexia. However, studies in English have used quite different paradigms than those in Dutch, making it difficult to ascertain which differences are associated with orthography and which reflect varying task demands. Clayton and Hulme (2018) suggested that poor readers may have problems in accessing phonological representations or may be generally slow in reaction time tasks, rather than having deficits at the level of letter-sound integration. They assessed how English-speaking children's letter-sound integration ability related to their reading performance by using a priming paradigm. A prime (either an English letter or an unfamiliar Greek letter) was presented, followed by a sound; children were asked to decide whether the presented sound was a real speech sound or not. The researchers predicted that children who automatically associated letters and sounds would respond more quickly when the prime was a letter that matched the speech sound versus when the prime was an unrelated Greek letter. This 'facilitation effect,' (i.e. the difference in reaction time for the matched over the unrelated Greek letter prime), served as an index of letter-sound integration. The researchers compared the facilitation effect among three groups: children with dyslexia age 9-11 years, age-matched typical readers, and reading-matched typical readers age 6-8 years. All three groups showed similar facilitation effects, suggesting that letter-sound integration in children with dyslexia was not impaired. In another study sample, the researchers recruited children age 5-7 years with varied reading proficiency and examined the correlation between letter-sound integration performance and reading. Using the same priming paradigm, they measured children's facilitation effect residual scores (in which the reaction time of the matched condition was regressed on the baseline reaction time). These facilitation scores were only weakly correlated with reading performance (composite of timed and untimed word reading); this relation was non-significant after controlling for children's age. Interestingly, a reaction time composite score (derived from all priming conditions) correlated strongly with reading ability, even after controlling for age, phonological awareness (PA), and RAN. Similarly, in their three-group comparison study (Clayton & Hulme, 2018), the dyslexia group performed significantly slower in the priming task than the age-matched typical readers, even though they had comparable facilitation effects to typical readers. Although these results do not support letter-sound integration deficits in dyslexia, slow reaction time in this priming task aligns with other studies showing that individuals with dyslexia are generally slower in letter-sound matching tasks (Aravena et al., 2013; Blau et al.,2010; Blomert, 2011). Given that this priming task is not purely a simple speed task as it does involve activation of letter-sound processing, slow reaction time may reflect a lack of automaticity in linking letters and sounds.

In another study of English-speaking children, Nash and colleagues (2017) used a similar priming task with children age 7-13 years, employing both behavioral and brain event-related potential (ERP) methods. Similar to Clayton and Hulme's study, they compared the letter-sound integration performance among three groups: children with dyslexia age 9-13 years, age-matched typical readers, and reading-matched children (age 7-9 years). Similar to the previous work, the groups did not differ significantly in behavioral facilitation effects. In contrast, the ERP data revealed significant spatial differences in early (P1 component) brain responses between the typical-reading groups and the dyslexic group. Specifically, greater P1 amplitudes observed in fronto-central regions in children with dyslexia may reflect greater processing effort. This could indicate that the dyslexic group had deficits in automaticity in letter-sound integration, because automaticity is defined by minimal processing effort (Kail & Hall, 1994; Logan, 1997), especially because the younger age-matched group was similar to the older typical readers. The authors suggested that the discrepancy between behavioral and ERP findings might reflect a limitation of behavioral priming, in which the task may be "not as sensitive to the timing of the feedback" from visual letters to speech sounds (p. 13). Thus, it remains unclear whether letter-sound integration is associated with reading ability in English.

Taxing automaticity by varying SOA

Systematically varying the stimulus or task properties for letter-sound integration tasks may allow for more sensitive measurement of the relation between letter-sound integration and its relationship with reading. For example, typical-reading Dutch adults were slower to decide if letter-sound pairs were matched or not when the letter-sound stimuli were presented simultaneously (stimulus onset asynchrony [SOA] of 0 ms) compared to conditions when letters and sounds were presented further apart (van Atteveldt et al., 2007). Manipulating the SOA of letter-sound stimuli in each pair potentially adjusts the demands on automaticity for letter-sound integration: a larger SOA allows sequential processing of each stimulus, reducing the cognitive load and the tax on automaticity. For example, a presentation of the first stimulus will activate its phonological or orthographic representation, and such information is readily available for deciding if it matches with the second stimulus or not, thus facilitating reaction time. Van Atteveldt and colleagues (2007) showed that when letters and sounds were presented closer in time, overall reaction time increased and the magnitude of the congruency effect, an index of letter-sound integration, was reduced. The congruency effect is based on the premise that if letters and speech sounds are successfully integrated, a faster reaction time is expected for congruent (matched) letter-sound pairs than for incongruent (mismatched) pairs (Van Atteveldt et al., 2006; 2007). It is evident that different temporal asynchrony of letter-sound pairs has an effect on letter-sound integration performance and may present an opportunity to study how automaticity demands relate to letter-sound integration performance and children's reading performance.

The present study

This study addressed three research questions: (**1**) does altering demands on automaticity, by manipulating timing of letter-sound stimulus presentation, affect children's letter-sound integration performance?; (**2**) is letter-sound integration related to reading ability, and does this differ depending on level of automaticity demand?; (**3**) does letter-sound integration account for variance in reading ability beyond the contribution of reading-related skills such as PA and RAN?

To address these questions, we employed a letter-sound matching task that varied levels of SOA between the letter and sound stimuli in a pair (as in van Atteveldt et al., 2007). The congruency effect was the indicator of letter-sound integration. We reasoned that the simultaneous presentation of a letter and sound would impose greater cognitive demands because it requires simultaneous processing, whereas asynchronous presentation allows less demanding serial processing; thus, the process of mapping phonological and orthographic information would be more heavily taxed during shorter SOA conditions. We expected that a stronger congruency effect would be observed when the letter and sound were presented further apart in their start times than when presented simultaneously (SOA 0 ms).

To create individual letter-sound integration efficiency scores that could be analyzed for their correlation with reading ability (Question 2), we employed a binning procedure, which integrates reaction times and accuracy into a single measure (Draheim et al., 2016; Hughes et al., 2014). Recently, concerns have been raised regarding approaches that use an experimental measure that compares two highly correlated conditions (e.g., a difference score or cost score) for analyses of individual differences (e.g., correlation, regression). Measuring a difference between two tasks that are highly correlated may result in low measure reliability and impact individual differences analyses (see Hedge et al., 2018 for discussion). The binning procedure resolves the issue of low reliability and allows us to analyze reading abilities in a continuous manner. Further, this continuous analysis approach allows us to avoid arbitrary cut-offs because there is no agreement on the specific tests or criteria to define dyslexia.

Finally, if our data showed an association between letter-sound integration and reading skills, we wanted to investigate whether letter-sound integration contributed to reading proficiency beyond other key cognitive-linguistic skills such as phonological awareness (PA) and rapid automatized naming (RAN). PA and RAN have been well established not only as unique predictors of children’s reading abilities but also as causal deficits in dyslexia (Norton et al., 2014; Norton & Wolf, 2012; Wolf & Bowers, 1999). Therefore, considering the contribution of PA and RAN helps us determine whether letter-sound integration may relate to reading because of shared processes (e.g., phonological identification, speeded processing) or whether letter-sound integration plays a unique role in reading ability.

Method

Participants

Participants included 67 children of varied reading abilities (32 females, age *M* = 9.7, *SD* = 0.8, range 8.0-11.6 years) recruited from schools and the community in [US city]. Participants met the following inclusion criteria: (1) English as the primary language used at home, (2) no known hearing loss, (3) no history of neurodevelopmental or neurological disorders such as autism or epilepsy, (4) typical or above-average nonverbal cognitive ability (scaled score ≥7 on the matrix reasoning subtest of the Wechsler Intelligence Scale for Children, 5th edition; Wechsler, 2014). We oversampled for children with reading difficulties with the goal of including a sufficient sample to understand the role of letter-sound integration across the spectrum of reading ability. Thus, 21 children (31% of the sample) had a diagnosis of dyslexia or a specific learning disability in reading from an outside clinician or other professional. Because there is no consensus on the specific measures or scores to define dyslexia, reading scores were analyzed. Because of the common co-occurrence of ADHD and reading difficulties, children with ADHD (*n* = 6) were included in the sample; those taking stimulant medications continued to do so as usual. This study was approved by the [XXXXXXX University] Institutional Review Board, consistent with principles of the Declaration of Helsinki. Parents provided informed written consent and children provided assent to participate.

Four additional children were recruited and tested but excluded from analyses because data were not recorded due to a technical computer problem (*n* = 1) or because children had persistent attention issues throughout the tasks that led to compromised data (*n* = 3).

Measures

Letter-sound matching task

A letter-sound matching task was used to assess letter-sound integration. This task presented a visual letter and an auditory speech sound in each trial that either matched (congruent trials) or mismatched (incongruent trials). To measure how different demands on automaticity affected letter-sound integration performance, letter and sound stimuli were presented at different stimulus onset asynchronies (SOAs): -400, -200, 0, 200, and 400 ms. These values were selected to cover a range of timing differences similar to previous research (van Atteveldt et al., 2007). For data presentation and analysis, negative SOAs indicate that a visual letter was presented before an auditory letter sound, whereas positive SOAs indicate that an auditory letter sound was presented before a visual letter (the choice for visual first to be negative versus positive was arbitrary). At the SOA of 0 ms, a visual letter and an auditory letter sound were presented simultaneously.

Task procedure. Children were given instructions on the task and completed 10 practice trials to ensure that they understood the procedure before the experiment began. Figure 1 shows a schematic of the task and stimulus presentation. In each trial, a fixation cross was first presented in the center of the screen for 800 ms. Then, the letter and sound stimuli were presented. Participants were instructed to respond whether the pair was matched (congruent) or mismatched (incongruent) as quickly and accurately as possible by pressing their left or right index finger on keys labeled 'Yes' or 'No' on a keyboard. The Yes and No responses (i.e., left and right hand) were counterbalanced across subjects.

Conditions and Trials. In total, there were 10 conditions, including 2 levels of congruency (matched vs. mismatched letter-sound pairs) and 5 levels of SOA (-400, -200, 0, 200, and 400 ms). In each condition, 30 trials were presented, for a total of 300 trials. The order of the trials was pseudo-randomized with regard to condition, and a given visual or auditory stimulus was not repeated in successive trials. Participants completed the trials in the same order and took a short break after completing a block of 75 trials. The entire task took about 15 minutes to complete.

Stimuli. The stimuli used in this experiment were consonants, because they have more reliable letter-sound mapping than vowels in English. Six letter stimuli (f, g, k, m, p, t) were chosen. These letters were unlikely to be visually confused with other letters and were relatively easy to discriminate from each other phonologically. Corresponding auditory letter sounds, /fə/, /gə/, /kə/, /mə/, /pə/, and /tə/, were recorded from a female native English speaker. We included the schwa sound /ə/ after each of the consonants, as it made the sounds more intelligible. The duration of the sounds was approximately 350 ms (±10 ms). The sounds were presented to participants through earphones at a sound intensity level of approximately 65 dB SPL. Letters were presented in white lowercase Arial font on a black background for 350 ms.

Task presentation and response measurement. The task was presented using Presentation software (Version 20.2; Neurobehavioral Systems, Berkeley, CA) on a laptop running Microsoft Windows 10. Screen size was approximately 27.5 cm x 17.3 cm. The accuracy and precision of the stimulus presentation timing were verified using an oscilloscope. Both reaction time (RT) and accuracy were recorded for each trial. RT was recorded from the onset time of the second stimulus (or at onset of both for the 0 ms SOA condition). For example, for a trial with SOA of 200 ms, RT was measured relative to the onset of the visual letter stimulus that was presented 200 ms after the onset of the speech sound.

Letter-sound matching task reaction time data cleaning

For simple reaction time analyses (question 1), only correct responses were included, as this is a common practice. Prior to calculating the mean reaction time for each participant within each of the conditions, reaction time data were cleaned to address responses that were extreme outliers. The interquartile range (IQR) rule, also known as Tukey fences, was used to identify outliers (Tukey, 1977). This method is less influenced by extreme values than approaches like the standard deviation method (Seo, 2006). The IQR method sets a minimum and maximum for the range based on the 1st (Q1) and 3rd (Q3) quartile, which are values below Q1 - (1.5 * (Q3 - Q1)) or above Q3 + (1.5 * (Q3 – Q1)), respectively. Minimum and maximum values were calculated for each participant for each of the congruency and SOA conditions. If the participant had RTs below the minimum IQR cutoff, the outlier RTs were replaced by that participant's minimum IQR cutoff value (the lowest acceptable RT) for that condition. Correspondingly, if the participant had RTs above the maximum IQR value, those outlier RTs were replaced by that participant's maximum IQR cutoff value for that condition. After the outliers were replaced, the mean reaction times for each participant within each of the conditions were calculated. The reaction time means for each participant and condition were log-transformed to reduce the positive skew of the distributions prior to analysis.

Binning procedure for deriving letter-sound integration efficiency scores

For analyses related to individual differences (questions 2 and 3), we included a measure of task performance that incorporated accuracy and RT, yet avoided concerns about low reliability. Using difference scores derived from highly similar or correlated conditions has been raised as problematic for correlational/individual differences studies (Hughes et al., 2014; Draheim et al., 2019) because this approach eliminates a considerable amount of systematic variance, resulting in low reliability of the difference measure (Edwards, 2001). In our case, the correlation of RT within participants for the conditions of interest (congruent vs. incongruent trials) ranged from $r = 0.88$ to 0.93 across SOA levels. We employed a binning procedure that integrated accuracy as well as reaction time (Hughes et al., 2014; Draheim et al., 2016) Briefly, the binning procedure compared each individual's RTs for congruent vs. incongruent correct trials, and each resulting value was assigned to an ordinal bin based on the histogram of all values across participants; incorrect trials are assigned a high bin score. The total bin score reflects both accuracy and RT (see Supplemental Materials for a detailed description of the binning procedure). This procedure also provides the advantage that it considers the speed-accuracy tradeoff that occurs in most reaction time studies.

Standardized reading and language assessments

Children completed a battery of standardized reading and reading-related measures. All analyses used raw scores (indicating children's absolute performance) rather than age-based standard scores (indicating performance relative to peers) because we were interested in the relationship of overall reading ability to letter-sound integration, as well as for comparability with the experimental letter-sound integration task. Age-based standard scores are presented for descriptive purposes and to allow comparison with other samples.

To create composite scores for constructs of phonological awareness, RAN, and reading ability, we converted the raw scores for each relevant subtest into z-scores within the study sample, then took the mean of the z-scores for each construct for each participant.

Nonverbal IQ. Nonverbal cognitive ability was assessed with the Matrix Reasoning subtest of the Wechsler Intelligence Scale for Children, 5th Edition (WISC-V) (Wechsler, 2014).

Phonological awareness. Phonological awareness was assessed with the Elision, Blending Words, and Phoneme Isolation subtests of the Comprehensive Test of Phonological Processing, 2nd Edition (CTOPP-2; Wagner et al., 2013).

Rapid automatized naming (RAN). RAN was assessed via the Color and Letter subtests of the Rapid Automatized Naming/Rapid Alternating Stimulus Tests (RAN/RAS, Wolf & Denckla, 2005).

Reading ability. Untimed single real word (Letter-Word Identification) and nonword (Word Attack) reading accuracy were assessed using the Woodcock Johnson Tests of Achievement, 4th Edition (WJ IV, Schrank et al., 2014). Timed real word (Sight Word Efficiency) and nonword reading (Phonemic Decoding Efficiency) efficiency were assessed using the Test of Word Reading Efficiency, 2nd Edition (TOWRE-2, Torgesen et al., 2012). Timed sentence reading with a comprehension component was assessed using the WJ IV Sentence Reading Fluency subtest.

Results

Descriptive statistics and correlations

Table 1 shows the means, standard deviations, and ranges for all assessment measures. The mean overall response accuracy on the letter-sound matching task, as expected, was high, with a mean of 92%, and range of 72 to 99%. Importantly, response time (RT) varied substantially across participants. Table 2 shows the descriptive information for accuracy, RT, and letter-sound integration efficiency (bin scores) by condition.

Both zero-order correlations among the variables of interest and partial correlations controlling for nonverbal IQ are shown in Table 3. It is noteworthy that age was not significantly correlated with most reading-related measures, potentially due to the heterogeneity of reading ability in the sample and poor readers tending to be older. Because age did not play a significant role in explaining our sample's reading ability, age was not included in further analysis.

Question 1: The effect of SOA and congruency on letter-sound integration reaction time

To assess how different automaticity demands (determined by SOA) impacted children's letter-sound integration, we examined the interaction between congruency condition and SOA on reaction time. These data are plotted in Figure 2. A repeated measures ANOVA (2 congruency conditions x 5 SOA levels) was conducted (for violations of sphericity assumptions, *p*-values are reported with Greenhouse-Geisser corrections and the original degrees of freedom). The analysis revealed significant main effects of congruency ($F(1, 66) = 84.55$, $p < .001$, $\eta_p^2 = .56$), with faster responses in the congruent than in the incongruent letter-sound condition and a main effect of SOA ($F(4, 264] = 127.09$, $p < .001$, $\eta_p^2 = .66$), as reaction times tended to be slower in higher timing-demand conditions, e.g., 0 ms SOA. There was also a significant interaction between congruency and SOA ($F(4, 264) = 4.17$, $p = .004$, $\eta_p^2 = .06$). To better understand this interaction, we ran follow-up planned contrasts to determine if the congruency effect in the 0 ms SOA condition was significantly smaller than other SOA conditions. *P*-values reported below were adjusted for 5 multiple comparisons using the Holm-Bonferroni (controlling family-wise error rate) method (Holm, 1979), with an alpha of .05. Congruency effects in the -400 ms ($t(66) = 3.64$, $p = .002$) as well as 400 ms condition ($t(66) = 2.69$, $p = .027$) were significantly larger than the SOA 0 ms condition. The -200 and 200 ms conditions were not significantly different than 0 ms (-200 ms: $t(66) = 1.17$, $p = .493$; 200 ms: $t(66) = 0.92$, $p = .493$).

Question 2: The relation between letter-sound integration efficiency and reading ability

We next examined how letter-sound task efficiency scores related to composite measures of reading, PA, and RAN, after controlling for nonverbal IQ (see Table 3). Spearman correlations were used as they are robust to the skewed distribution of letter-sound integration efficiency scores. There were broad significant associations between letter-sound integration efficiency scores in partial correlations controlling for nonverbal IQ. Better performance in three of the SOA conditions (-200, 200, and 400 ms) was significantly correlated with PA (r_s = -.26 to -.29, $p < .05$). These conditions and also the 0 ms SOA condition were all significantly associated with better reading (r_s = -.29 to -.37, $p < .05$). There were no significant correlations between letter-sound integration and RAN. Further, efficiency scores for the -400 ms condition (visual stimulus appeared 400 ms before the onset of the speech sound) were not significantly associated with PA, RAN, or reading abilities.

Question 3: The unique contribution of letter-sound integration to children's reading ability

Given the significant partial correlation results, we proceeded to determine if letter-sound integration accounted for unique variance in reading beyond the contribution of IQ, PA, and RAN. We ran regression analyses separately for the 400 ms and 0 ms conditions; these conditions were selected a priori for testing because they provide a contrast between low levels of cognitive demand (400 ms, sound then visual letter) and high cognitive demand (simultaneous presentation). Regression analyses are shown in Table 4. Nonverbal IQ was entered in step 1, then PA and RAN in step 2. As expected, these measures together accounted for substantial variance in the composite measure of reading ability (57%, F [2,63] = 50.18, $p < .001$). In the next step of the regression, we entered either 0 or 400 ms efficiency scores. Letter-sound integration performance at 0 ms SOA (step 3a) explained an additional 3% of variance, which was statistically significant ($\Delta R^2 = 0.03$, $p = .02$). However, performance at 400 ms SOA (step 3b) did not account for significant additional variance in reading ability ($\Delta R^2 = .01$, $p = .21$).

Discussion

This study both assessed the extent to which automaticity demands affect letter-sound integration performance and the relation between letter-sound integration efficiency and reading in English-speaking children. Our findings showed that children's letter-sound integration performance, as indexed by the congruency effect (faster reaction time for congruent vs. incongruent trials), changed as a function of temporal asynchrony of the stimulus presentation. Specifically, letter-sound integration performance was better in conditions with less automaticity demand. Children's letter-sound integration efficiency scores (bin scores combining accuracy and reaction time) in four of the five SOA conditions were significantly correlated with their reading ability composite score. Regression analyses revealed that the relation between letter-sound integration efficiency in the highest automaticity demand condition (0 ms SOA) accounted for unique variance in reading ability after accounting for nonverbal IQ, phonological awareness (PA), and rapid automatized naming (RAN).

Letter-sound integration and automaticity demands

Consistent with previous literature in Dutch adults (van Atteveldt et al., 2007), English-speaking children responded faster in congruent than incongruent letter-sound trials (see Figure 2), demonstrating a pronounced congruency effect. In addition, children performed faster in both congruent and incongruent conditions when a letter and sound were presented sequentially rather than simultaneously, suggesting that different levels of automaticity demand (achieved by manipulating temporal asynchrony of letter-sound stimuli) influenced letter-sound integration performance. Notably, there was a significant interaction effect between congruency and SOA, suggesting that the efficiency of letter-sound integration can be modulated by temporal asynchrony of stimuli, with more efficient letter-sound integration when the stimuli are presented further apart in time. We propose that this effect reflects how the serial presentation of stimuli lessens the load on cognitive resources, allowing children to process the stimuli over time. For example, the first stimulus activates a phonological or orthographic representation, and such information is readily available for comparison with the successive stimulus, facilitating reaction time. Interestingly, RAN was only weakly and non-significantly correlated with the letter-sound integration efficiency scores. This indicates that the letter-sound integration efficiency measure taps a different subset of skills despite shared demands of automaticity. Together, the present study and existing literature highlight the critical role of timing in letter-sound integration and suggest that a congruency effect for letter-sound integration is apparently present despite differences in participant age and orthographic transparency of the language.

The relation between letter-sound integration and children's reading proficiency

A key finding from this study is the unique contribution of letter-sound integration to English-speaking children's reading ability. Our results demonstrate that greater letter-sound integration efficiency (as indicated by smaller efficiency scores) at SOAs of -200, 0, 200, and 400 ms were significantly associated with better performance (*r*s = -.29 to -.37) on a composite reading measure that involved accuracy, speed, and fluent comprehension. We predicted that closer temporal synchrony of letter-sound pairs would be particularly associated with children's reading performance due to greater automaticity demands, as integration across multiple modalities such as vision and audition can be influenced by varying task demands (Miller & D'Esposito, 2005; van Atteveldt et al., 2007; Van Wassenhove et al., 2007). Our results generally aligned with this view; the patterns for the 0, -200 and 200 ms SOA conditions were all similarly and significantly related to reading ability. The -200 ms and 200 ms conditions may impose a cognitive load essentially similar to the 0 ms condition. An interesting observation is that the 400 ms condition (speech sound presented before letter), similar to other SOA conditions, correlated with reading performance. The only SOA condition that did not show a significant association with reading ability was the -400 ms condition (visual letter presented before speech sound), which was likely the easiest because visual letters can be processed quickly (recognized within 100-200 ms; Liu et al., 2009) and the whole visual stimulus is available for processing immediately, whereas sounds unfold over time. Therefore, in the -400 ms condition, participants could recognize and process the letter stimulus before the auditory stimulus was presented, allowing serial rather than more demanding parallel processing. The SOA values examined here were not chosen because of particular hypothesized differences in timing but were selected to cover a range of timing offsets. Thus, we did not plan *a priori* to test the differences between the strength of each of these conditions' correlations with reading.

The unique contribution of letter-sound integration to reading ability

Hierarchical regression analyses showed that letter-sound integration efficiency in the most demanding condition (0 ms SOA) uniquely accounted for variance in reading performance beyond the contributions of PA, RAN, and nonverbal IQ, whereas performance in the 400 ms SOA condition did not. We posit that because the 0 ms SOA condition required greater automaticity due to parallel processing of letter-sound stimuli, it may be more sensitive in revealing the relation between letter-sound integration and reading even after controlling for other skills. These findings also indicate that letter-sound integration is a skill that is unique and not subsumed by the phonological processes required by our PA measures or the speed and automaticity processes also involved in RAN.

These results contrast with previous studies in English (Clayton & Hulme, 2018; Nash et al., 2017) that showed no significant relation between letter-sound integration and dyslexia status at the group level. Several methodological differences between their study and the current study may account for these contrasting results. First, the studies used different experimental design (priming approach vs. explicit letter-sound matching task). Second, we used a binning procedure that combined both reaction time and accuracy of letter-sound integration. It could be the case that the combination of accuracy and reaction time is more sensitive in revealing the relation between letter-sound integration and reading performance, as reading relies on both these factors. Third, the binning approach allowed us to use continuous rather than categorical analyses, which may be more sensitive than a heterogeneous diagnostic category like dyslexia.

It could also be the case that accuracy versus automaticity of letter-sound association have different predictive influence on reading performance depending on age or reading stage. In a longitudinal study, Clayton and colleagues (2019) demonstrated that 4-5-year-olds' letter-sound knowledge accuracy, but not the letter-sound integration facilitation effect, predicted reading performance a year later. In addition, German children's accuracy in learning novel symbol-sound relations predicted their later reading fluency, even after controlling for PA, RAN, IQ, and home literacy (Horbach et al., 2015).

Overall, our findings suggest that letter-sound integration plays a role in reading ability and thus a deficit in this area may contribute to reading breakdown for some children with dyslexia. Our data show that PA and RAN together account for about 60% of variance in children's overall reading ability, consistent with the notion that these two factors are predominant causal deficits in children with dyslexia (e.g., Norton & Wolf, 2012; Wolf & Bowers, 1999). To our knowledge, our finding is the first showing that letter-sound integration accounts for reading performance in school-aged children even after considering for PA and RAN. Therefore, letter-sound integration could potentially be an important predictor for children's reading, especially because letter-sound knowledge accuracy measures typically reach a ceiling after first grade and lose their predictive power. Further, our finding suggests that it is plausible that some children may have letter-sound integration deficits. If this can be identified early, this is a strong target for intervention.

Limitations and future directions

Some limitations of this study should be noted. Because stimuli in the letter-sound matching task included only six single letters, the variety and complexities of letter-sound correspondences in English orthography were not fully captured. Many single phonemes are represented by two or more letters, e.g. /f/ as in "phone", and graphemes can represent various phonemes, e.g. letter c pronounced as /k/ as in "cat" and /s/ as in "city". These complexities likely influence letter-sound integration performance. Future studies should expand the scope of assessing letter-sound integration by more comprehensively attending to the intricacies of English letter-sound mapping. Further, the binning procedure for deriving efficiency scores offers an important advantage but is just one of many potential ways to assess letter-sound integration.

Our sample included children with a broad range of reading abilities, relatively over-sampled for poor readers, and some children with ADHD. Although our findings suggest a specific role for letter-sound integration in reading and thus are consistent with a multifactorial view of reading impairment (O'Brien & Yeatman, 2019; Peterson & Pennington, 2015) in which this could be a potential deficit, further investigation is needed. It might be assumed that a subgroup of children with dyslexia would have poor letter-sound integration efficiency, and that this deficit may co-occur with other deficits such as in PA or RAN. Unfortunately, our study lacked statistical power to examine such subgroups. This is an important research question in light of identifying different subtypes of dyslexia and targeting effective intervention.

Finally, these data add to previous findings of compromised letter-sound integration in poor readers in more transparent orthographies such as Dutch and Finnish. Of particular note, even though children's letter-sound knowledge accuracy is commonly assessed and some standardized measures are available in English (e.g., York Assessment of Reading for Comprehension, Hulme et al., 2009), there currently is no consensus on how letter-sound integration (efficiency) should be measured. We propose that development of a standardized measure of letter-sound integration would allow researchers and clinicians to quantify this ability more readily and gain a more comprehensive understanding of reading ability and dyslexia.

References

Aravena, S., Snellings, P., Tijms, J., & van der Molen, M. W. (2013). A lab-controlled simulation of a letter-speech sound binding deficit in dyslexia. *Journal of Experimental Child Psychology*, *115*(4), 691-707.

Blau, V., Reithler, J., van Atteveldt, N., Seitz, J., Gerretsen, P., Goebel, R., & Blomert, L. (2010). Deviant processing of letters and speech sounds as proximate cause of reading failure: a functional magnetic resonance imaging study of dyslexic children. *Brain, 133*(3), 868-879.

Blau, V., van Atteveldt, N., Ekkebus, M., Goebel, R., & Blomert, L. (2009). Reduced neural integration of letters and speech sounds links phonological and reading deficits in adult dyslexia. *Current Biology, 19*(6), 503-508.

Blomert, L. (2011). The neural signature of orthographic-phonological binding in successful and failing reading development. *Neuroimage, 57*(3), 695-703.

Blomert, L., & Froyen, D. (2010). Multi-sensory learning and learning to read. *International Journal of Psychophysiology*, *77*(3), 195-204.

Blomert, L., & Willems, G. (2010). Is there a causal link from a phonological awareness deficit to reading failure in children at familial risk for dyslexia? *Dyslexia*, *16*(4), 300-317.

Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litavský, M., Onochie-Quintanilla, E., ... & Seidlová-Málková, G. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychological Science, 23*(6), 678-686.

Clayton, F. J., & Hulme, C. (2018). Automatic activation of sounds by letters occurs early in development but is not impaired in children with dyslexia. *Scientific Studies of Reading, 22*(2), 137-151.

Clayton, F. J., West, G., Sears, C., Hulme, C., & Lervåg, A. (2019). A longitudinal study of early reading development: letter-sound knowledge, phoneme awareness and RAN, but not letter-sound integration, predict variations in reading development. *Scientific Studies of Reading*, *24(*2), 91-107.

Draheim, C., Hicks, K. L., & Engle, R. W. (2016). Combining reaction time and accuracy: The relationship between working memory capacity and task switching as a case example. *Perspectives on Psychological Science*, *11*(1), 133-155.

Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. *Psychological Bulletin*, *145*(5), 508-535.

Edwards, J. (2001). Ten difference score myths. *Organizational Research Methods*, *4*, 265-287.

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading, 9*(2), 167-188.

Ehri, L. C., & Wilce, L. S. (1983). Development of word identification speed in skilled and less skilled beginning readers. *Journal of Educational Psychology, 75*(1), 3-18.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: true issues and false trails. *Psychological Bulletin, 123*(1), 71-99.

Froyen, D., Willems, G., & Blomert, L. (2011). Evidence for a specific cross-modal association deficit in dyslexia: an electrophysiological study of letter-speech sound processing. *Developmental Science*, *14*(4), 635-648.

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods, 50*, 1166-1186.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal* *of Statistics*, *6*(2), 65–70.

Horbach, J., Scharke, W., Cröll, J., Heim, S., & Günther, T. (2015). Kindergarteners' performance in a sound–symbol paradigm predicts early reading. *Journal of Experimental Child Psychology, 139*, 256-264.

Hughes, M. M., Linck, J. A., Bowles, A. R., Koeth, J. T., & Bunting, M. F. (2014). Alternatives to switch-cost scoring in the task-switching paradigm: Their reliability and increased validity. *Behavior Research Methods, 46*(3), 702–721.

Hulme, C., Nash, H. M., Gooch, D., Lervåg, A., & Snowling, M. J. (2015). The foundations of literacy development in children at familial risk of dyslexia. *Psychological Science, 26*(12), 1877-1886.

Hulme, C., Stothard, S. E., Clarke, P., Bowyer-Crane, C., Harrington, A., Truelove, E., & Snowling, M. J. (2009). *YARC: York Assessment of Reading for Comprehension: Early Reading.* GL Publishers.

Kail, R., & Hall, L. K. (1994). Processing speed, naming speed, and reading. *Developmental Psychology*, *30*(6), 949-954.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*(2), 293-323.

Liu, H., Agam, Y., Madsen, J. R., & Kreiman, G. (2009). Timing, timing, timing: fast decoding of object information from intracranial field potentials in human visual cortex. *Neuron*, *62*(2), 281-290.

Logan, G. D. (1997). Automaticity and reading: Perspectives from the instance theory of automatization. *Reading & Writing Quarterly*, *13*(2), 123-146.

McBride-Chang, C. (1999). The ABCs of the ABCs: The development of letter-name and letter-sound knowledge. *Merrill-Palmer Quarterly, 45*, 285-308.

Miller, L. M., & D'Esposito, M. (2005). Perceptual fusion and stimulus coincidence in the cross-modal integration of speech. *Journal of Neuroscience*, *25*(25), 5884-5893.

Nash, H., Gooch, D., Hulme, C., Mahajan, Y., McArthur, G., Steinmetzger, K., & Snowling, M. (2017). Are the literacy difficulties that characterize developmental dyslexia associated with a failure to integrate letters and speech sounds? *Developmental Science, 20*, e12423.

Norton, E. S., Black, J. M., Stanley, L. M., Tanaka, H., Gabrieli, J. D., Sawyer, C., & Hoeft, F. (2014). Functional neuroanatomical evidence for the double-deficit hypothesis of developmental dyslexia. *Neuropsychologia*, *61*, 235-246.

Norton, E. S., & Wolf, M. (2012). Rapid automatized naming (RAN) and reading fluency: Implications for understanding and treatment of reading disabilities. *Annual Review of Psychology, 63*, 427-452.

O'Brien, G., & Yeatman, J. (2019). Bridging sensory and language theories of dyslexia: towards a multifactorial model. *bioRxiv*. doi: 10.1101/773853

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly, 40*(2), 184-202.

Peterson, R. L., & Pennington, B. F. (2015). Developmental dyslexia. *Annual Review of Clinical Psychology, 11*, 283-307.

Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., & Foorman, B. R. (2004). Kindergarten prediction of reading skills: A longitudinal comparative analysis. *Journal of Educational Psychology*, *96*(2), 265-282.

Schrank, F. A., McGrew, K. S., & Mather, N. (2014). *Woodcock-Johnson IV Tests of Cognitive Abilities.* Rolling Meadows, IL: Riverside.

Snowling, M. J. (1980). The development of grapheme-phoneme correspondence in normal and dyslexic readers. *Journal of Experimental Child Psychology*, *29*(2), 294-305.

Torgesen, J. K., Wagner, R., & Rashotte, C. (2012). *Test of Word Reading Efficiency–Second Edition (TOWRE-2)*. Austin, TX: Pro-Ed.

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

Van Atteveldt, N. M., Formisano, E., Blomert, L., & Goebel, R. (2006). The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. *Cerebral Cortex*, *17*(4), 962-974.

Van Atteveldt, N. M., Formisano, E., Goebel, R., & Blomert, L. (2007). Top–down task effects overrule automatic multisensory responses to letter-sound pairs in auditory association cortex. *Neuroimage, 36*(4), 1345-1360.

Van Wassenhove, V., Grant, K. W., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia*, *45*(3), 598-607.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2013). *Comprehensive Test of Phonological Processing–Second Edition (CTOPP-2).* Austin, TX: Pro-Ed.

Wechsler, D. (2014). *Wechsler Intelligence Scale for Children* (5th ed.). Bloomington, IN: PsychCorp.

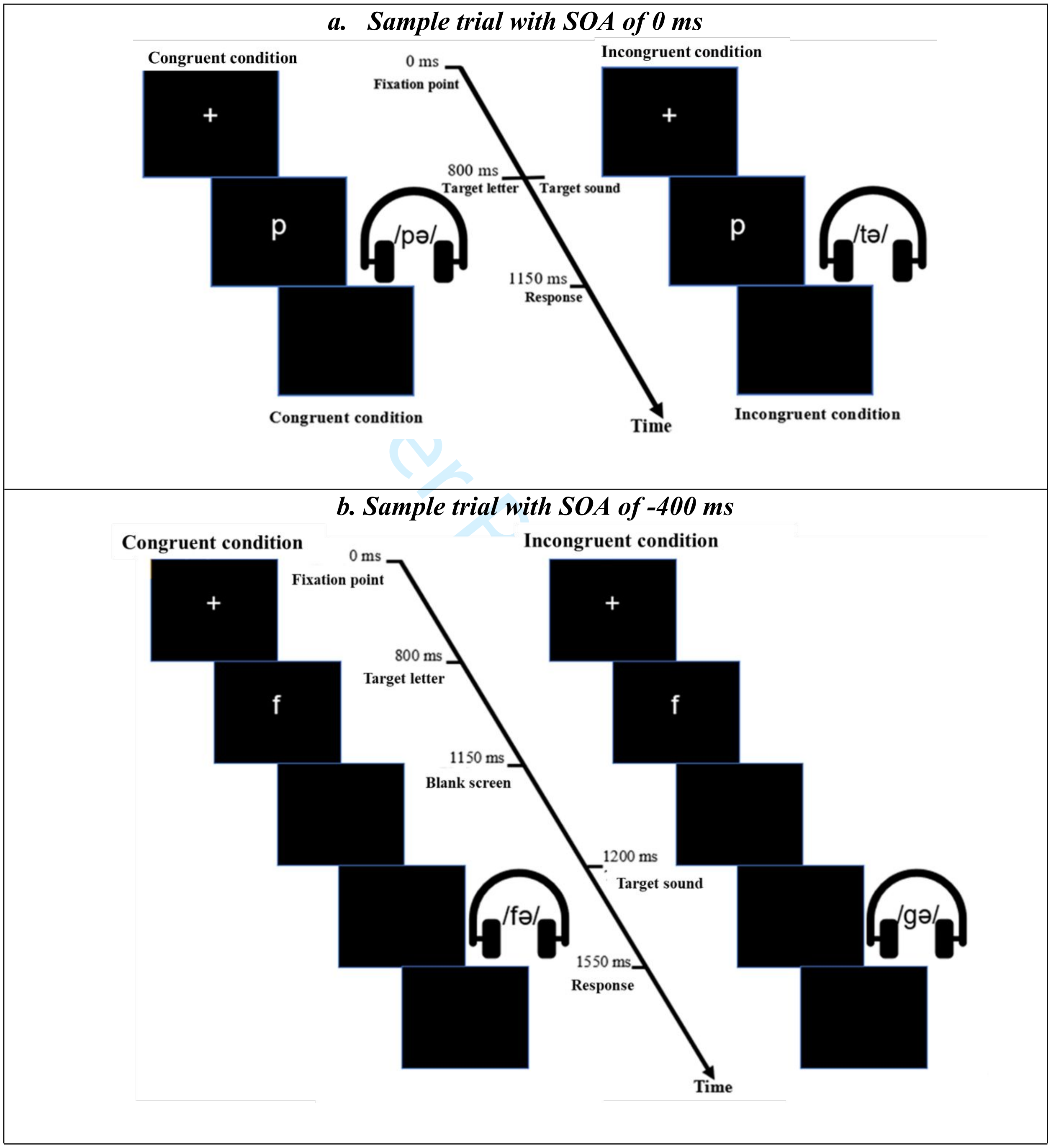
Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology*, *91*(3), 415.

Wolf. M, & Denckla, M. B., (2005). *Rapid Automatized Naming and Rapid Automatized Stimulus Tests.* Austin, TX: Pro-Ed.

Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Scientific Studies of Reading, 5*(3), 211-239.

Figure 1

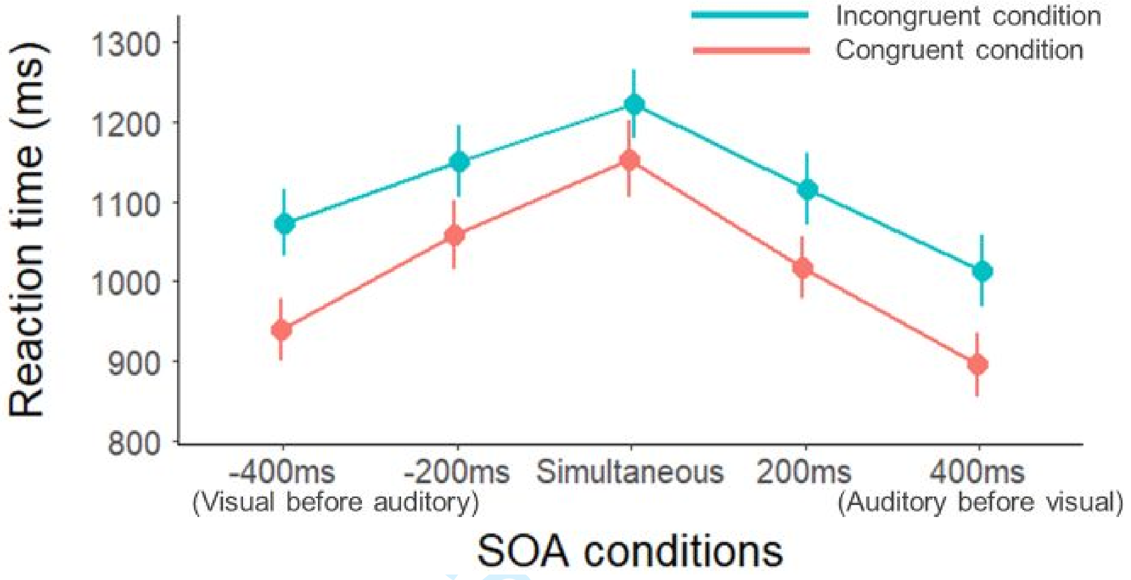
Schematic of Letter-Sound Matching Task Stimulus Presentation at SOA of (a) 0 ms and (b) -400 ms.



Note: Times listed are the onset of each event relative to the onset of the trial.

Figure 2

Mean Reaction Times by SOA and Congruency Condition



Note: On the X-axis: Negative ms conditions represent a visual stimulus presented before an auditory stimulus; positive ms conditions represent an auditory stimulus presented before a visual stimulus. Error bars represent 1 SE.

Table 1

Descriptive Statistics for Standardized Assessments.

Construct	Subtest	Raw scores		Standard scores	
		M (SD)	Range	M (SD)	Range
Nonverbal IQ	Matrix reasoning	19.28 (2.65)	14 – 16	11.12 (2.25)	7 – 17
PA	Elision	26.22 (5.11)	12 – 33	9.87 (2.68)	3 – 15
	Blending words	24.16 (2.94)	17 – 29	10.06 (2.50)	5 – 15
	Phoneme isolation	25.63 (3.38)	11 – 32	9.40 (2.22)	4 – 15
RAN	Colors	46.12 (11.38)	25 – 85	98.04 (17.40)	59 – 145
	Letters	28.87 (7.52)	14 – 60	98.18 (13.19)	70 – 145
Reading	Letter-word identification	56.12 (9.78)	23 – 73	102.42 (17.29)	41 – 142
	Word attack	22.72 (5.49)	9 – 32	107.37 (18.54)	63 – 154
	Sight word efficiency	64.34 (15.68)	18 – 106	97.19 (17.86)	55 – 145
	Phonemic decoding efficiency	34.54 (14.53)	6 – 64	100.16 (18.89)	62 – 145
	Sentence reading fluency	47.36 (19.22)	2 – 109	101.48 (20.57)	44 – 160

Notes. Nonverbal IQ and PA scaled scores have M = 10, SD = 3. RAN and reading standard scores M = 100, SD = 15.

Table 2

Descriptive Information for the Letter-Sound Integration (LSI) Task: Accuracy, Response Time (RT), and Letter-Sound Efficiency Bin Scores by SOA Condition.

SOA condition	**Accuracy (trials correct out of 30)**		**RT (ms)**		**LSI efficiency**
	Congruent **M (SD)** **Range**	**Incongruent** **M (SD)** **Range**	**Congruent** **M (SD)** **Range**	**Incongruent** **M (SD)** **Range**	**Bin scores** **M (SD)** **Range**
-400 ms	27.13 (2.91)	28.24 (1.64)	939 (317)	1074 (337)	190.60 (34.40)
	16 – 30	22 – 30	509 – 2223	576 – 2235	131 – 275
-200 ms	27.49 (2.33)	28.25 (1.91)	1059 (358)	1150 (371)	190.38 (37.75)
	19 – 30	21 – 30	576 – 2185	642 – 2297	125 – 299
0 ms	26.90(2.97)	27.99(1.81)	1154 (399)	1223 (359)	194.25 (39.80)
	16 – 30	21 – 30	647 – 3015	713 – 2288	126 – 356
200 ms	26.58 (2.74)	27.97 (1.90)	1017 (324)	1116 (376)	194.49 (40.74)
	18 – 30	22 – 30	556 – 2096	545 – 2515	112 – 295
400 ms	26.52 (2.77)	27.85 (2.34)	895 (333)	1013 (368)	196.22 (41.11)
	19 – 30	21 – 30	434 – 1989	480 – 2360	119 – 303

Note. Negative SOAs indicate that a visual letter is presented before auditory letter sound. RT is reported relative to the onset of the second stimulus.

Table 3

Spearman Correlations and Partial Correlations Controlling for Nonverbal IQ Among Standardized Assessments and LSI Task Efficiency Scores.

Measures	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. Age	–	–	–	–	–	–	–	–	–	–
2. Nonverbal IQ	.11	–	–	–	–	–	–	–	–	–
3. Reading z-score composite	.12	.21	–	.49***	-.71***	-.10	-.29*	-.35**	-.37**	-.32**
4. PA z-score composite	.10	.24*	.51***	–	-.34**	-.21	-.29*	-.22	-.27*	-.26*
5. RAN z-score composite	-.18	-.19	-.72***	-.37**	–	.01	.11	.17	.12	.18
6. LSI score -400 ms	-.09	-.17	-.13	-.24	.04	–	.28*	.38**	.31*	.26*
7. LSI score -200 ms	-.18	-.19	-.32**	-.32**	.14	.31*	–	.50***	.45***	.41***
8. LSI score 0 ms	.00	-.19	-.37**	-.25*	.20	.40**	.51***	–	.54***	.45***
9. LSI score 200 ms	.09	-.13	-.39**	-.29*	.14	.33**	.46***	.55***	–	.57***
10. LSI score 400 ms	-.07	-.12	-.33**	-.28*	.20	.28*	.42***	.46***	.57***	–

Notes. *$p < .05$, **$p < .01$, ***$p < .001$. Zero-order correlations are below the diagonal. Partial correlations controlling for nonverbal IQ are displayed above the diagonal.
PA = phonological awareness, RAN = rapid automatized naming, LSI = letter-sound integration; LSI scores are efficiency bin scores. Negative SOAs indicate that a visual letter is presented before auditory letter sound.

Table 4

Hierarchical Regression Analyses: Letter-sound Integration Efficiency (SOA 0 ms and 400 ms) Explaining Children's Reading Ability

Step and variables	R^2	ΔR^2	ΔF	Final B	SE B	Final β
Step 1	.08		5.31*			
Non-verbal IQ				.02	.03	.06
Step 2	.64	.57	50.18***			
PA				.38	.09	.32***
RAN				-.58	.08	-.57***
Step 3a	.68	.03	6.07*			
Letter-sound integration efficiency **(0 ms)**				-.004	.002	-.19*
Step 3b	.65	.01	1.57			
Letter-sound integration efficiency **(400 ms)**				-.002	.002	-.10

Notes. N= 67, *$p < 0.05$, ***$p < 0.001$. This table presents results for two separate regression analyses. The variables entered in Step 1 and 2 are the same for both analyses. Steps 3a and 3b contrast the addition of the two different SOA conditions. The final values for Blocks 1 and 2 are similar for both analyses, so the values for SOA 0 ms are presented.

Supplementary Materials

Supplementary methods

The binning procedure of combining RT and accuracy for letter-sound integration

An alternative to using simple difference scores between conditions is to use a binning procedure (Hughes et al., 2014; Draheim et al., 2016). This procedure is appropriate for experiments that intend to use difference scores between two experimental conditions' reaction times and accuracy in a meaningful context. It also provides the advantage that it considers the speed-accuracy tradeoff that occurs in most reaction time studies. The difficulty levels in the congruent and incongruent conditions were relatively comparable in our experiment, thus we did not expect that accuracy would be significantly different between the conditions. However, we reasoned that accuracy should be taken into account for children's performance in this letter-sound matching task because errors in this task may reflect that a child does not automatically integrate letters and sounds.

The following steps were completed to calculate a letter-sound integration efficiency score for each participant for each SOA level (based on Hughes et al., 2014; Draheim et al., 2016). In **step 1**, we calculated each child's mean RT for their correct trials in the congruent condition. In **step 2**, we subtracted the reaction times of each correct trial in the congruent condition from the mean RT of the congruent trials from step 1. For example, if a child had 28 correct trials in the incongruent condition, that child would have 28 scores calculated; each score represents an incongruent condition trial's reaction time relative to the mean reaction time of the congruent condition. Thus, positive scores represent longer reaction time in the incongruent condition than the congruent condition. Negative scores represent faster reaction times in the incongruent than the congruent condition. In **step 3**, we assigned a bin value to each score from step 2. To do this, we grouped all scores from all subjects together and sorted them by rank order, forming a distribution of RT scores (similar to a histogram) each representing one trial's difference between congruent and incongruent conditions. We then grouped the rank-ordered scores into 10 deciles. Figure S1 shows a schematic of this distribution of this binning procedure (adapted from Hughes et al., 2014). The highest 10% of the scores were assigned a bin value of 1, then the next 10% of the scores had a bin value of 2, and the process continued until the lowest 10% of the scores were given a bin value of 10. A bin value of 1 is associated with the largest congruency effect, in which a child's response to the congruent condition is faster compared to that particular trial from the incongruent condition. A bin value of 10 is associated with the smallest or no congruency effect magnitude (RTs in two conditions are equal or the incongruent condition is faster). In **step 4**, after assigning bin values to the correct items in the incongruent condition, a bin value of 20 was assigned to each incorrect trial in the incongruent condition. Ultimately, each of the 30 trials in the incongruent condition (per each SOA level) was assigned a bin value (possible values of 1 to 10, or 20). Finally, in **step 5**, the sum of the bin values for each SOA for each participant was calculated; this is the metric for letter-sound integration efficiency used in analyses. A smaller total bin score indicates better letter-sound integration (performance is faster in the congruent condition than in the congruent condition, and more accurate).

References

Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. *Psychological Bulletin*, *145*(5), 508-535.

Hughes, M. M., Linck, J. A., Bowles, A. R., Koeth, J. T., & Bunting, M. F. (2014). Alternatives to switch-cost scoring in the task-switching paradigm: Their reliability and increased validity. *Behavior Research Methods, 46*(3), 702–721.

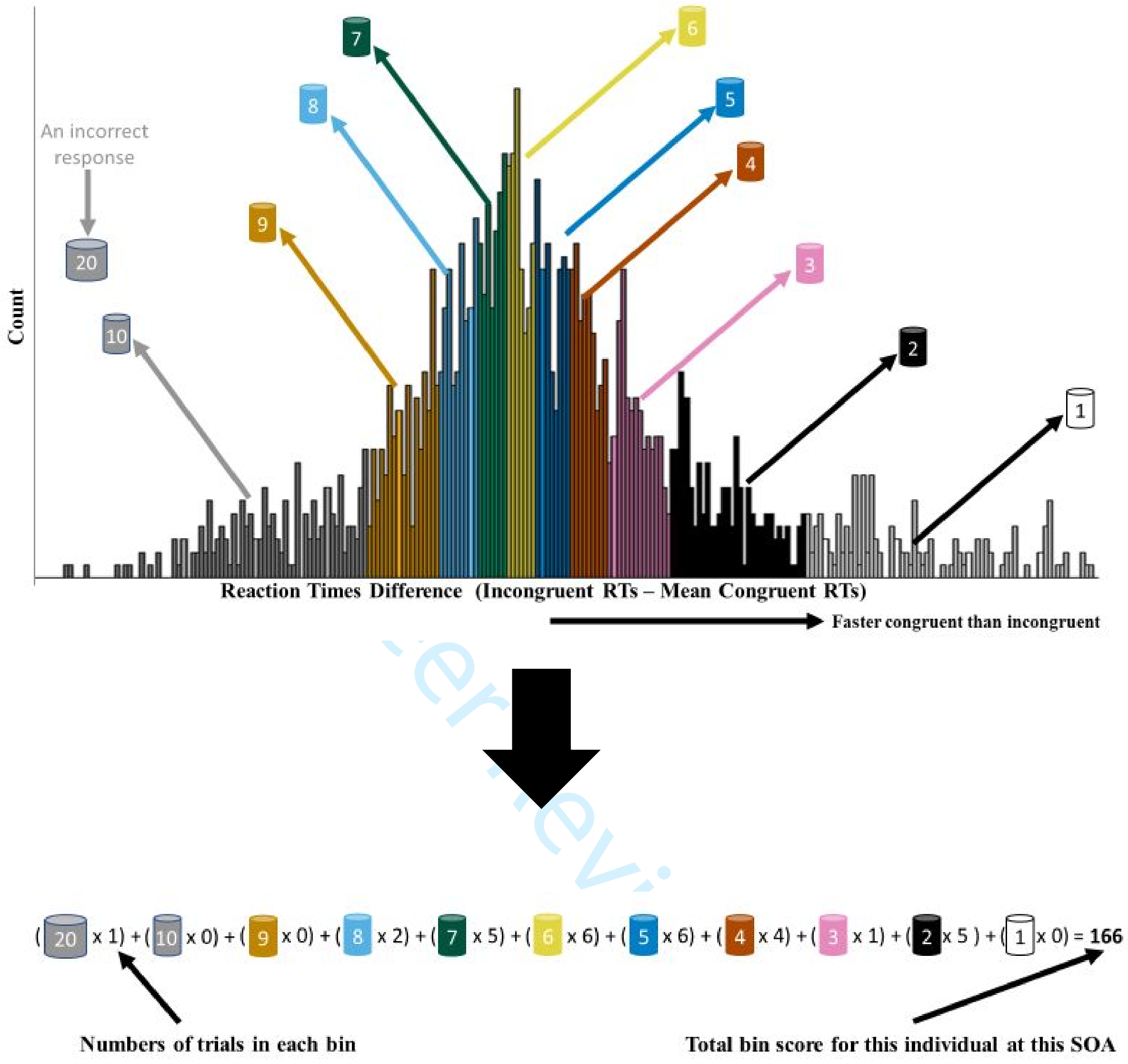


Figure S1. Schematic of the binning procedure.

Note. This histogram shows the distribution of reaction time differences for all participants in an SOA condition. Reaction time differences are calculated as the RTs in the incongruent condition minus the mean RT of the congruent condition for each individual, and this is repeated for all 30 trials per condition per individual. A bin value of 1 is associated with the largest congruency effect (child's mean RT in the congruent condition is faster than that particular trial in the incongruent condition). A bin value of 10 is associated with the smallest or no congruency effect magnitude (RTs in two conditions are equal or the incongruent condition is faster). Then, an individual's bin score can be calculated by summing the bin scores for all the trials for this SOA condition. Lower bin scores indicate greater efficiency and magnitude of congruency effect. Figure is based on Hughes et al. (2014).