Psychological Research (2025) 89:93
https://doi.org/10.1007/s00426-025-02080-w

RESEARCH

Is letter position coding a unique skill for developing and adult readers in early word processing? Evidence from masked priming

Pablo Gómez[1] · Ana Marcet[2] · Francisco Rocabado[3] · Manuel Perea[2,3]

Received: 17 October 2023 / Accepted: 14 January 2025

Abstract

Reading words in alphabetic scripts requires encoding the relative order of the letters. This process of letter position coding is known to be flexible. For instance, the masked transposed-letter prime `jugde` activates the word *JUDGE* to a greater degree than a replacement-letter prime like `jupte`, a phenomenon known as the transposed-letter effect. In this study, we investigated whether the transposed-letter effect in masked priming is related to reading skills (as measured by a standardized reading test) in a sample of sixth-grade children. Targets (e.g., `RITME`: Catalan for rhythm) were preceded by identity primes (`ritme`), transposed-letter primes (`rimte`), or replacement-letter primes (`risle`) in a sandwich priming paradigm. Results showed that transposed-letter primes were more effective than replacement-letter primes but less effective than identity primes. More importantly, while the readers' reading skills modulated overall latency and accuracy, we found no evidence that the participants' reading skill modulated the size of the priming effects. This outcome prompted us to re-analyze analogous conditions in a masked priming mega-study with approximately 1000 adult participants (Adelman et al. Behav Res Methods 46(4):1052–1067, 2014), where we found a near-zero correlation between the size of transposed-letter priming and spelling and vocabulary tests. These findings suggest that if there are individual differences in the first moments of word processing, these are not detectable for neurotypical readers in laboratory tasks.

A highly robust phenomenon that has advanced our understanding of visual word recognition and reading is the approximate coding of letter position: pseudowords created by switching the positions of two letters in a word (e.g., `mohter` for "mother" or `jugde` for "judge") are perceived as quite similar to the original words (see Grainger, 2018, 2022, for reviews). This phenomenon has often been examined using the lexical decision task (e.g., deciding whether an item is a word or not) by comparing the responses to transposed-letter pseudowords (e.g., `mohter`) and replacement-letter pseudowords (e.g., `monfer`; where `th` in `mother` is replaced with `nf`). The typical finding is that lexical decision times are longer—and with a higher proportion of "word" responses—for `mohter` than for `monfer` (see Perea & Lupker, 2004, for early research; see Mirault & Grainger, 2021, for recent evidence), thus suggesting that `mohter` activates its base word to a larger degree than `monfer`. Another commonly employed procedure to study the flexibility of letter position coding during word recognition is Forster and Davis' (1984) masked priming paradigm. In this procedure, the response times to the target word `MOTHER` are collected; critically, this target is briefly preceded by a transposed-letter prime like `mother` or a replacement-letter prime `monfer`. The typical finding is that the prime-TARGET pair `mohter-MOTHER` produces shorter response times than `monfer-MOTHER` (see Perea & Lupker, 2003, for early evidence; see Hasenäcker & Schroeder, 2022; Spinelli et al., 2022, for recent evidence), revealing some uncertainty in letter position assignment in the first moments of word processing. Notably, this flexibility in letter position coding led the field to reject models with fixed slots for encoding letter positions (e.g., McClelland & Rumelhart, 1981) and motivated the implementation of more sophisticated models of the encoding of serial order information during word recognition (e.g., spatial coding model, Davis, 2010; LTRS model, Adelman, 2011;

✉ Pablo Gómez
pgomez@skidmore.edu

1 Skidmore College, 815 N Broadway, Saratoga Springs, NY 12866, USA

2 Universitat de València, Valencia, Spain

3 Centro de Investigación Nebrija en Cognición (CINC), Universidad Nebrija, Madrid, Spain

Published online: 21 April 2025

dual-route model, Grainger & Ziegler, 2011; overlap model, Gomez et al., 2008).

Critically, most previous research on letter position coding has focused on average group data, either individuals from the same age or, in developmental studies, across age groups. Thus, the implicit assumption is that there is a priming effect that researchers aim to estimate (which we term θ from now on). Most of this research has aimed to establish if the priming effect is present across various populations and types of stimuli (is $\theta > 0$?). Given the robustness of the transposed letter effects in masked priming experiments, we can, with high confidence, assume that indeed $\theta > 0$. If we assume that there are true individual differences in priming effects, then θ is the latent attribute of the person that putatively relates to letter position coding skill; hence, the different authors of this article have θ_{Pablo}, θ_{Ana}, θ_{Fran}, and θ_{Manolo}. In terms of a "lexical quality" account of individual differences in reading (e.g., Castles et al., 2007; Perfetti & Hart, 2002), one might argue that the orthographic representations of better readers may be more fine-tuned to the encoding of letter order than that of poor readers; in other words, θ covaries with reading skill. Indeed, several lexical decision experiments revealed that the process of letter position encoding might be more accurate for more proficient readers. For instance, Perea et al. (2016) conducted an unprimed lexical decision experiment that included transposed-letter pseudowords and replacement-letter pseudowords (e.g., `CHOLOCATE` vs. `CHOTONATE`). They compared a group of adult readers with extensive expertise in orthographic-lexical processing—competitive Scrabble players—and a control group. They found that the transposed-letter effect (i.e., the difference between the replacement-letter vs. transposed-letter conditions) in the error data was much smaller for the Scrabble players than for the individuals in the control group (4.2% vs. 18.2%, respectively). A similar pattern emerged for the latency data (75 ms vs. 117 ms, respectively), but the critical interaction was not significant. Note, however, that this study, while informative, simply supports the hypothesis that there are group-level differences and not that there are discernible individual differences related to other measurements.

More related to the current paper, Gómez et al. (2021) examined the role of individual differences in letter position coding with a sample of Grade 6 readers. Gómez et al. collected a measure of reading ability from a standardized reading test and conducted a lexical decision experiment. The pseudowords in the experiment were created by letter transposition or letter replacement (e.g., `MOHTER` vs. `MONFER`). As expected, they found much longer and less accurate responses to `MOHTER` than `MONFER` (141 ms slower and 33.1% less accurate). More important, Gómez et al. (2021) also found that, in the accuracy data, the transposed-letter effect was smaller for those children with better reading abilities than those with poorer reading abilities ($b = -0.42$); however, the response time data did not show any clear signs of a relationship. Thus, the above-reviewed two studies provide some hints on the role of individual differences in letter position coding during visual word recognition. However, these effects occurred mainly in the accuracy data. Furthermore, the focus was on the pseudoword stimuli, thus making it difficult to generalize to word recognition and reading (see also Pagán et al., 2021, for a similar pattern using a reading-like task).

To address the above limitations, the present study focused on the gold standard technique to investigate the earliest moments of orthographic processing during word recognition: the masked priming technique applied to the lexical decision task (Forster & Davis, 1984; see Grainger, 2008, 2018, for reviews). In this technique, the relationship between a briefly presented (and forwardly masked) lowercase prime and an uppercase target stimulus is manipulated (e.g., identity priming condition: `mother-MOTHER`; transposed-letter priming condition: `mohter-MOTHER`; replacement-letter priming condition: `molfer-MOTHER`). For pseudoword primes, there is an advantage of the transposed-letter condition over the replacement-letter condition in the response time data; in addition, the identity condition is slightly more effective than the transposed-letter condition (e.g., Forster et al., 1987; Kezilas et al., 2017; Perea & Lupker, 2004).

Two prior studies examined how individual differences influence the size of masked transposed-letter priming effects in lexical decision tasks (Andrews & Lo, 2012; Ziegler et al., 2014). In a sample of university students, Andrews and Lo (2012) investigated whether the readers' reading abilities modulated the difference between orthographically related primes (transposed-letter primes or one-letter replacement primes) and an unrelated condition. In their experiment, half of the targets were paired with pseudoword primes (e.g., [transposed-letter prime] `crue-CURE`, [one-letter different prime] `cire-CURE`, [unrelated prime] `gine-CURE`). In contrast, the other half had word primes (e.g., [transposed-letter prime] `colt-CLOT`, [one-letter different prime] `plot-CLOT`, [unrelated prime] `punt-CLOT`). The overall by-subject and by-items Analyses of Variance did not show concurrent differences between the response times in form-related priming conditions (either by transposition of two letters or by replacement of one letter) and the unrelated prime condition. For example, the lexical decision times to the target `CURE` were only 8 ms faster when preceded by the transposed-letter pseudoword prime `crue` compared to when preceded by the unrelated pseudoword `gine`—it is worth noting that the pattern was the opposite for word primes (i.e., the transposed-letter word prime `colt` produced slower responses [around 13 ms] to `CLOT` than the unrelated word prime `punt`).

Regarding individual differences analyses, Andrews and Lo (2012) found that more proficient readers exhibited faster response times. They also observed that the slower responses for transposed-letter word primes (e.g., the prime colt for the target CLOT) relative to the unrelated primes were more pronounced for the more proficient readers. However, the (small) differences between transposed-letter pseudoword primes and unrelated primes were barely influenced by reading ability (see Fig. 2 in Andrews & Lo, 2012). Thus, while the findings of Andrews and Lo (2012) suggest that there may be substantial inhibition from transposed-letter word primes for proficient readers, the overall small differences between the transposed-letter pseudoword priming and the unrelated priming conditions do not allow for solid conclusions regarding the role of reading abilities on the flexibility of letter position coding from sublexical levels at the earliest moments of word processing.

One strategy to amplify the magnitude of masked priming effects is by including a briefly presented "pre-prime" between the forward mask and the prime. This is known as the "sandwich" variant of the masked priming technique (Lupker & Davis, 2009). Using the sandwich technique, Ziegler et al. (2014) examined how children's reading abilities in Grades 1–5 modulated the size of masked transposed-letter priming. They compared transposed-letter and replacement-letter priming conditions using pseudoword primes (e.g., cousre-COURSE versus coufpe-COURSE). As lexical decision times significantly differ for younger and older groups (ranging from 2 s to 600–700 ms), Ziegler et al. utilized *z*-scores of the response time (RT) differences between the transposed-letter and replacement-letter conditions to compare the groups. Overall, reading ability was positively associated with the size of the transposed-letter priming effect ($r = 0.304$). However, as noted by Grainger and Ziegler (2011), the level of orthographic processing, whether fine-grained or coarse-grained, may considerably differ between children just starting to read (e.g., Grades 1 and 2) and more independent readers (e.g., Grades 4 and 5).

Overview of the present studies

The main goal of the present paper was to investigate whether individuals' reading abilities modulate letter position coding in early word recognition processing, operationalized in terms of the size of masked transposed-letter priming. We pre-planned Study 1, featuring original data collection from sixth-grade readers (i.e., the same population as in the Gómez et al., 2021, experiment). To gain a comprehensive understanding of the phenomenon under scrutiny, we compared the transposed-letter (TL) primes against two baselines: (1) a RL-TL (replacement-letter condition vs. transposed-letter condition, e.g., risle-RITME vs. rimte-RITME), which provided a measure of how much more effective a transposed-letter prime is relative to its orthographic control prime; (2) ID-TL (identity vs. transposed-letter condition; e.g., ritme-RITME vs. rimte-RITME; [the Catalan for *rhythm*], which measured how confusable the transposed-letter prime is with its base word. This study was conducted in Catalan, which is an orthography that lies in the middle of the transparent/opaque spectrum (Llaurado & Tolchinsky, 2015)—note that transposed-letter effects in European languages (e.g., French, English, Spanish, Italian, etc.) show a very similar pattern (e.g., Perea et al., 2018; see Lee et al., 2021, for more elusive evidence of these effects in Korean Hangul). In addition, we administered two scales (pseudoword reading, word reading) of the most widely used reading assessment test in Spain (PROLEC-Catalan; Cuetos et al., 2007)—these two scales were the same as in the Gómez et al. (2021) experiment with Grade 6 students. We pre-planned and registered the analysis for Study 1 at https://osf.io/x8tua

Considering the results of Study 1, we conducted Study 2. This second study was not pre-planned, but we performed the same analyses as in Study 1 using archival data from adult participants in a masked priming mega-study in English ($N > 900$; Adelman et al., 2014). While Adelman et al. used numerous experimental conditions, our focus was only on the three conditions that paralleled those used in Study 1 (i.e., identity condition, e.g., design-DESIGN; internal transposed-letter condition, desgin-DESIGN; internal double replacement-letter condition, dewvgn-DESIGN). Notably, Adelman et al. also collected data from spelling ability and vocabulary size tests, thus allowing us to examine the interplay between letter position coding in the first moments of word processing and reading skills. The final analyses presented here, which we term Study 3, involve a post hoc re-analysis of the reliability of the priming measurements from Study 1, and further explorations of the relationship between the reading scores and the priming effects.[1]

In sum, the present paper examined whether reading skill modulates letter position coding in the earliest moments of orthographic processing via transposed-letter priming using the sandwich variant of the masked priming technique. A detectable relationship between transposed-letter priming effects, with better readers showing smaller transposed letter effects, would be consistent with the "lexical quality" hypothesis (e.g., Castles et al., 2007; Perfetti & Hart, 2002) and would suggest that the parameters related to letter encoding flexibility vary depending on reading skill (e.g., the values concerning position uncertainty in Davis', 2010 spatial coding model and Gomez et al.'s 2008 overlap model would

[1] We thank an anonymous reviewer for suggesting some of these analyses.

be smaller for better readers). On the other hand, if there is no discernible relationship between the reading test scores, this would indicate that the variability in priming effects among individuals is not sufficiently modulated by reading skills to be detectable. This outcome would align with recent evidence using parafoveal previews during sentence reading with adult readers (Lee et al., 2024). Specifically, Lee et al. (2024) found that the size of transposed-letter parafoveal priming effects on a target word across several eye movement measures was not modulated by a composite of reading ability scores[2]—note that, similar to masked priming, this technique taps into the initial stages of word processing (see Marcet et al., 2019, for a comparison of the two techniques).

Study 1. Masked priming with sixth-grade children

Methods

Participants

We recruited 84 sixth-grade children (45 self-identified females; mean age = 11.6 years, SD = 0.48, range: 11–12) from three schools in Valencian-speaking towns in the metropolitan area of Valencia. This sample size mimics that used by Gómez et al. (2021). All participants were native speakers of Valencian, the name used for the variant of Catalan spoken in the Valencian Community, where both Valencian and Spanish are official languages. Valencian had been the instruction language of the participants at all levels in primary school. All participants had normal or corrected vision. Informed consent forms were signed by their parents before the experimental session. Five individuals were excluded from the analyses because of diagnoses of learning disabilities or dyslexia, thus resulting in a total sample of seventy-nine participants.

Materials

We selected 120 words from the Catalan word-frequency database (Rafel i Fontanals, 1998). These words had a mean word frequency per million was 122.3 (range: 25.14–1102.24), an average length was 5.7 letters (range: 5–7), and an average number of orthographic neighbors (i.e., Coltheart's *N*) was 4.9 (range: 0–25). For each target word (e.g., `RITME` [rythm]), which was presented in uppercase, we created three lowercase primes: (1) an identity prime (e.g., `ritme`); (2) a transposed-letter [TL] prime in which two middle adjacent consonants from the base word were transposed (e.g., `rimte`); and (3) a replacement-letter [RL] prime where two adjacent middle consonants from the base word were replaced (e.g., `risle`). The TL and RL primes were matched on log-bigram frequencies and Coltheart's *N* (both $ps > 0.23$; TL-primes: 1.50 [range: 0.40–2.52] and 0.79 [range: 0–9], respectively; RL-primes: 1.46 [range: 0.34–2.38] and 0.78 [range: 0–5]). We created a set of 120 orthographically legal pseudoword targets for the lexical decision task. These pseudowords were matched on number or letters and syllable structure with the word targets (e.g., the word *SENYAL* [signal] served to create the pseudoword `VENYER`, the word *CAMBRA* [room] served to create the pseudoword `LAMPRE`, etc.). The mean Coltheart's *N* of the pseudowords was 2.33 (range: 0–13). For each target pseudoword, we created an identity prime, a transposed-letter prime, and a replacement-letter prime in the same manner as the word targets. The list of prime-target pairs is presented in Appendix A. We constructed three counterbalanced sets of materials of 240 trials (120 word trials: 40 in each priming condition; 120 nonword trials: 40 in each priming condition) following a Latin square manner. For instance, if the TL-pseudoword `rimte` appeared in one set, the RL-pseudoword `risle` would appear in the second set, and the identity prime `ritme` would appear in the third. We also included a practice phase of 18 trials (9 word and 9 nonword trials) with the same characteristics as the experimental trials.

Procedure

The experiment took place in groups of up to 10 participants in a quiet classroom within the school premises. To display the stimuli and register the timing and accuracy of the responses, we employed Windows computers equipped with DMDX (Forster & Forster, 2003). The sequence of events in a given trial was the following, where all stimuli appear in the same location of the center of the screen: (1) a pattern mask `(#######)` was displayed for 500 ms; (2) the target stimulus was presented in lowercase for 33 ms (i.e., the pre-prime); (3) a lowercase prime stimulus was presented for 50 ms; and (4) the target stimulus, in uppercase, was presented until response or until the 2000 ms time limit. All the stimuli were displayed in black on a white background. We used 12-pt Courier New to present the stimuli, except for the pre-prime, which was presented in a smaller size (8-pt; see Lupker & Davis, 2009, for a similar choice to minimize visual overlap). The participants' task was to decide whether the uppercase stimulus was a word or not by pressing a keyboard key with a green sticker for "yes" responses or a key with a red sticker for "no" responses with their index fingers. Instructions stressed both speed and accuracy while not

[2] The composite measure of reading ability included the Wechsler Individual Achievement Test (WIAT-II UK, Wechsler, 2005), the Nelson Denny Reading Test (Brown et al., 1993), and the LexTale, among others.

mentioning the presence of practice primes. Each participant received a random ordering of trials (18 as practice, 240 as experimental phase).

In addition, to obtain an index of reading ability, each participant was individually evaluated with the PROLEC-R reading test in Catalan (Cuetos et al., 2007). This test measures the accuracy and speed when reading aloud a set of 40 words (word sub-test) and a group of 40 pseudowords (pseudoword sub-test). Half of the participants did the experiment first, and the other half did the reading test. The whole session (experiment and reading test) lasted between 25 and 30 min.

Table 1 Summary of latency and accuracy findings in Study 1

Prime	Mean RT correct	SD RT (by participant)	Accuracy
Responses to words			
ID	853	290	0.918
RL	901	292	0.896
TL	866	295	0.909
Responses to Nonwords			
ID	1139	330	0.778
RL	1147	333	0.772
TL	1140	334	0.768

The Response Times (RTs) are presented in milliseconds

Results

This research combines confirmatory and exploratory methods, and hence, we preregistered our analysis plan at https://osf.io/x8tua. First, we analyze the sixth-grade children's performance on the standardized reading test (PROLEC-R, Cuetos et al., 2007) to assess if our sample has enough score variability. Then, we examine the results of the masked priming experiment using the sandwich priming method, comparing response times and error rates for identity, transposed-letter, and replacement-letter primes. Finally, we explore whether the transposed-letter effect size correlates with reading test scores.

Standardized test results

Figure 1 shows the distribution of scores from the reading test. The scores demonstrate a well-behaved distributional shape, although they appear on the lower end of the scale compared to the interpretation guide. For this study, we focused on two sub-scores: word reading and nonword reading. The word reading score assesses lexical processes, while the nonword reading score measures sub-lexical processes, such as letter-to-sound coding, using pseudowords. Crucially for our objectives, the distribution exhibits satisfactory variability.

Transposed-letter priming effect

To establish the presence of a transposed-letter priming effect in our sample of young readers, we performed latency and accuracy analyses. Incorrect responses and responses with durations shorter than 250ms (accounting for 52 responses, 0.27% of the total) were excluded. Following the standard practice in masked priming experiments using lexical decision, we conducted separate analyses for word and nonword trials using the Bayes factors R package (Morey & Rouder, 2018) for the statistical analyses; we also used the

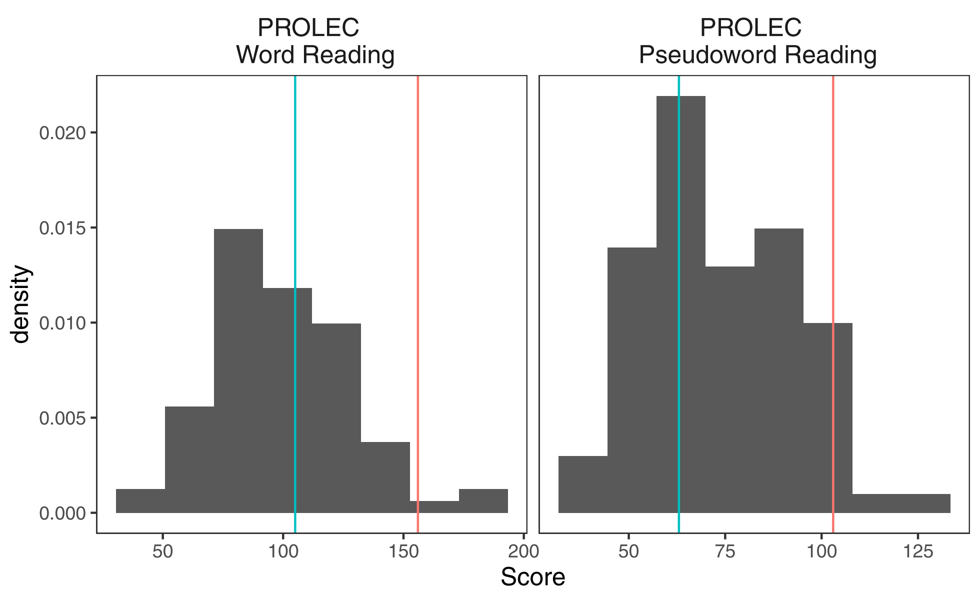

Fig. 1 Distribution of scores in the PROLEC-R test. The vertical lines represent the cut-off points for the low-performance, middle-performance, and high-performance groups as described in the test's interpretation manual

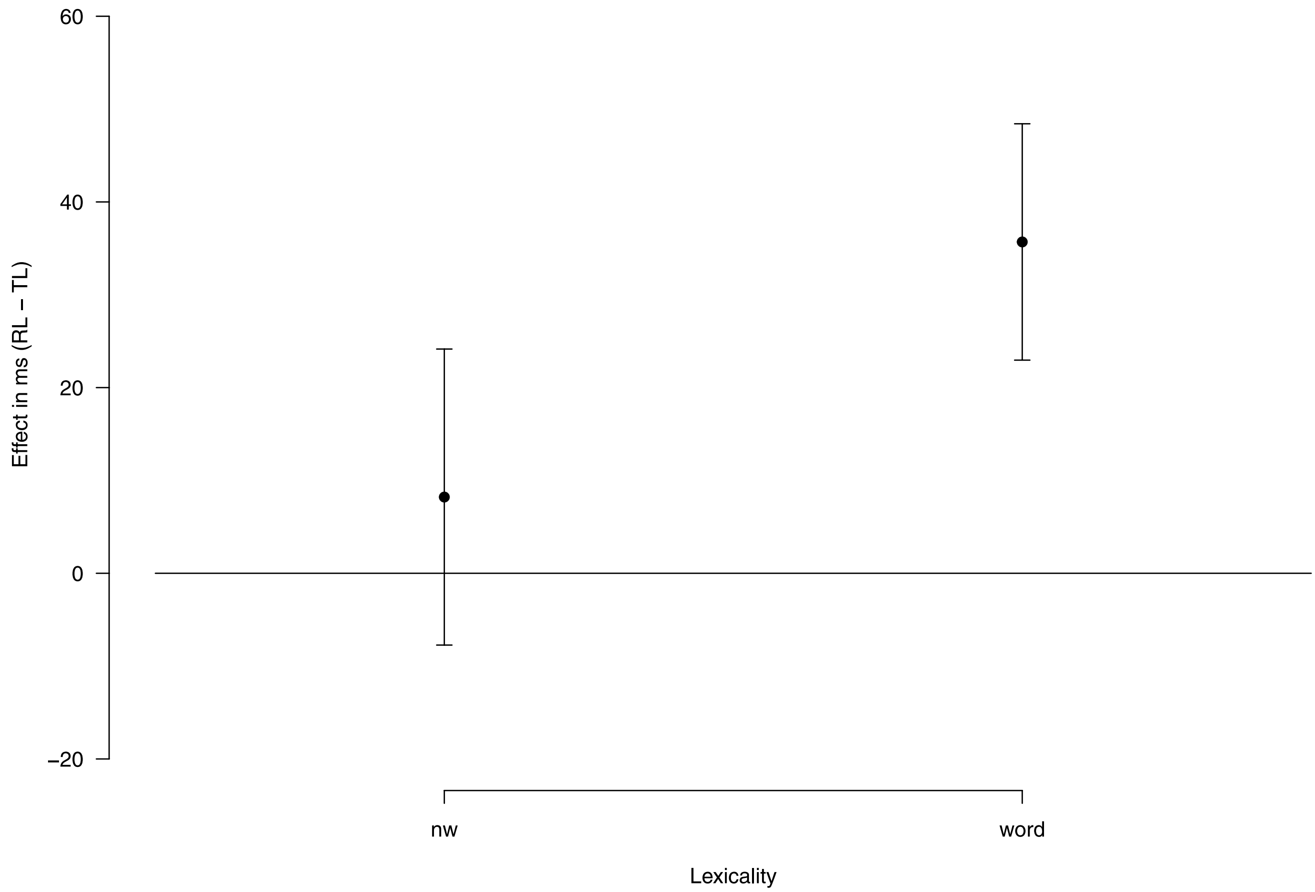

Fig. 2 Priming effects for the contrast between replacement-letter and transposed-letter primes in Study 1. The figure displays the mean effect across subjects, along with the frequentist confidence interval

lme4 package (Bates et al., 2015) with identical conclusions; those analyses are available in the Online Appendix.

The overall analyses produced straightforward findings, replicating previous research (see Table 1 and Figs. 2 and 3). As discussed in the Introduction, the effect of prime type in our dataset can be assessed in two complementary ways: comparing the RTs for the items with replaced letter primes to those with transposed letter primes (RL-TL from here on) or comparing the RTs to items with transposed letter primes to those with identity primes (TL-ID from here on). For word trials, we observed a significant 36-ms advantage of the transposed-letter condition over the replacement-letter condition, indicating the presence of a transposed-letter priming effect. Additionally, there was a 12-ms advantage of the identity priming condition over the transposed-letter condition. The Bayesian ANOVA with items and subjects as random effects yielded a Bayes Factor (BF_{10}) of 7e + 12 when comparing the model that includes the type of prime factor vs not (prior: rscale = 0.3; this corresponds to a narrow prior for all analyses because masked priming effects tend to be small), providing overwhelming evidence favoring a prime effect. Importantly, this effect is not solely driven by identity primes. If we compare transposed-letter vs. replacement-letter primes (the standard way to measure the transposed-letter priming effect), we obtain a substantial BF_{10} of 7e + 6, supporting the presence of the transposed-letter priming effect.

For nonword trials, the results indicate little to no effect of the prime type, with a Bayes Factor of $BF_{01} = 163$, suggesting that the data is consistent with the null model. Similarly, the effects on accuracy are numerically small for word trials, with a Bayes Factor of $BF_{01} = 3$ (anecdotal evidence for the null model). For nonword trials, the Bayes Factor is $BF_{01} = 518$ (indicating substantial evidence for the null model). These findings suggest that the prime type had minimal impact on accuracy for both word and nonword trials.[3]

[3] For all BF ANOVAs in Study 1 and Study 2, the code in the BayesFactor R package: BayesFactor::anovaBF(Dep Variable ~ prime + subject + item, data, whichRandom = c("subject", "item"), rscaleFixed = .3).
We compare the model with prime as a factor versus the model without.

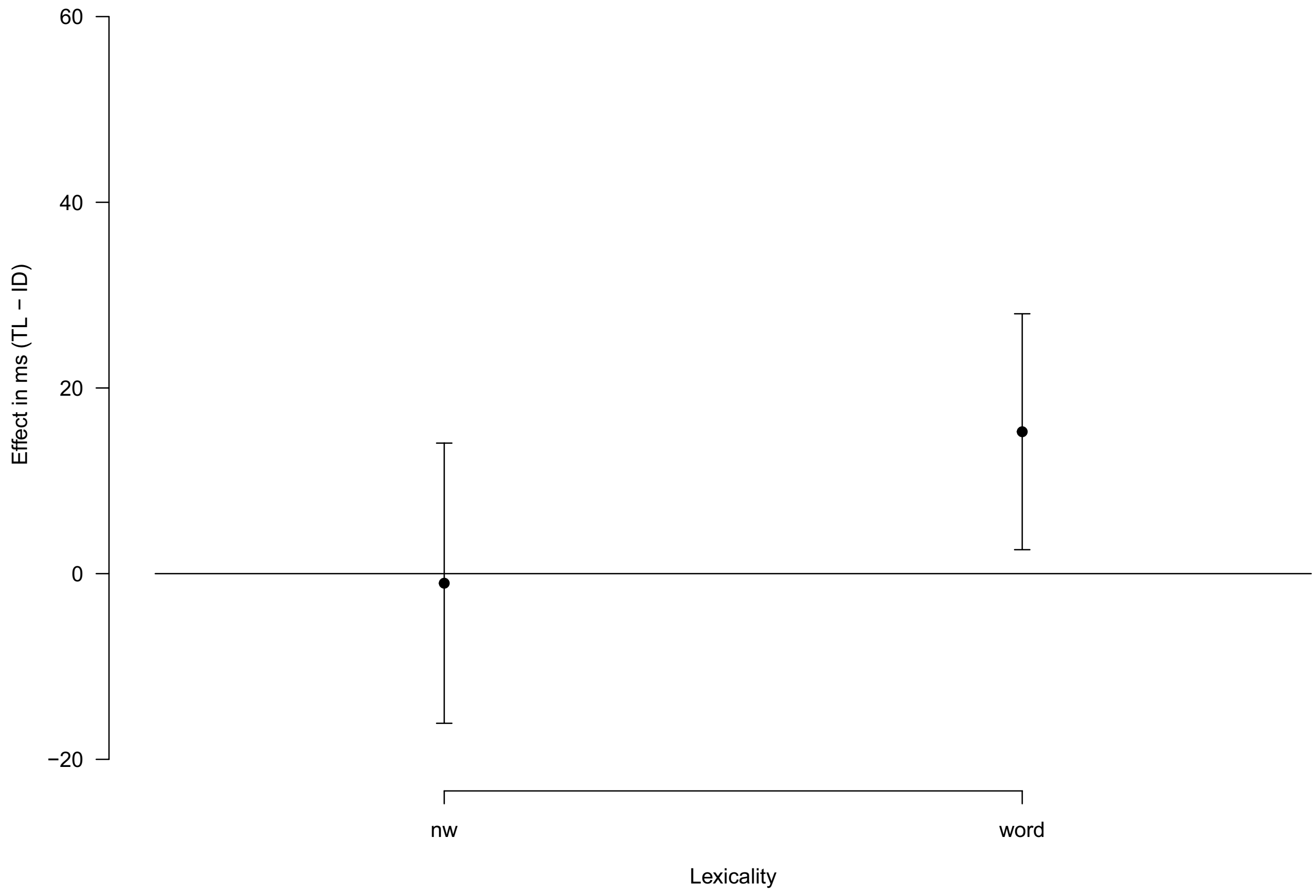

Fig. 3 Priming effects for the contrast between transposed-letter and identity primes in Study 1. The figure displays the mean effect across subjects, along with the frequentist confidence interval

Exploratory data analysis (delta plots)

The results described above point towards evidence favoring a null model for RT to nonword trials and for accuracy to both words and nonwords. In other words, there is an effect of prime time only evident in the RTs to words. Because of this reason, we present here exploratory data analyses (Tukey, 1977) only for latency data for word trials (the parallel analyses for nonwords are shown in the online Appendix). Exploratory data analyses, broadly defined, allow us to explore the distributional features of the data via statistical graphs. Specifically, for response time data, the temporal dynamic of an effect can be diagnostic of different cognitive processes at play.

Delta plots are a simple way to display how a latency effect (in our case, the effect of type of prime) evolves across time. They have been widely used in cognitive control tasks (e.g., De Jong et al., 1994) and have been used to describe the effects of identity primes versus unrelated primes in adults by Fernández-López et al. (2022). The plots were built as follows:

Step 1. For the correct responses to word trials, we first calculated the RTs at the 0.1, 0.3, 0.5, 0.7, and 0.9 quantiles for each participant and each priming condition; hence, we had five sets of RTs for the three prime types: Identity primes Transposed-letter primes, and replacement-letter primes for each of the participants.
Step 2. The RTs at each quantile from *Step 1* were averaged across participants so that we have the average at the quantile and the prime type level (vincentiles); hence, we have five vincentiles for each of the three prime types.
Step 3. The difference in each of the quantiles between the two conditions to be compared is obtained (hence the name of this type of graph: Delta plot); in addition, the average for each of the quantiles is calculated. The deltas are mapped into the *y*-axis of the plot, and the averages are mapped into the *x*-axis.

Figure 4 shows the two relevant comparisons. As can be seen in the Figure, the delta plots are relatively flat. The contrast that yields the highest delta is between Replaced Letter primes and Transposed-Letter primes. A slightly different

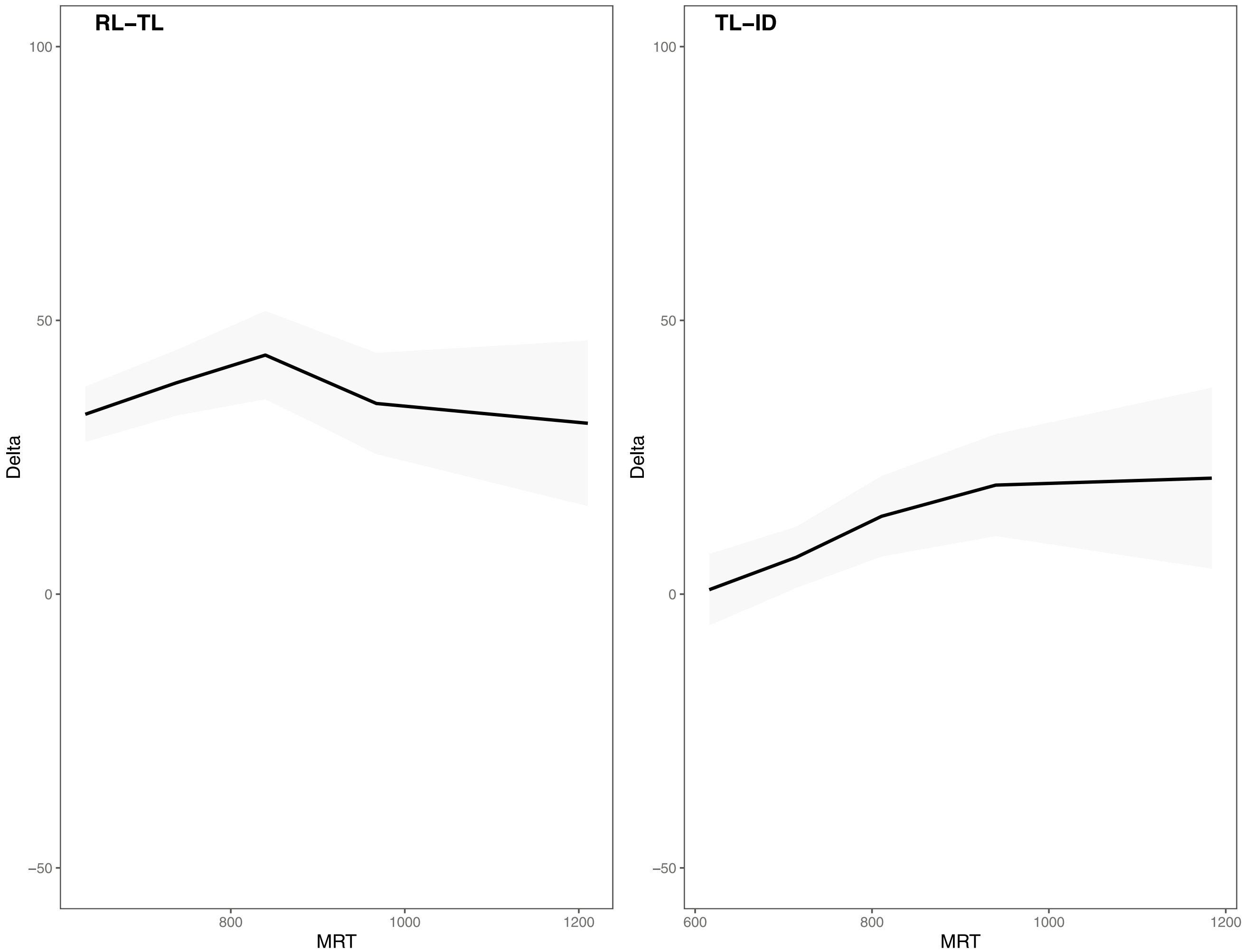

Fig. 4 Delta plot illustrating the difference in response between the conditions labeled on the top of each panel for Study 1. The *x*-axis represents time intervals of latency, while the *y*-axis represents the delta in response time between the conditions. The flat nature of the plot suggests that the effect of the type of prime on RTs is consistent regardless of the relative speed of the responses

pattern can be seen in the contrast between Transposed Letter primes and IDentity primes, with a difference between the 0.1 and the higher quantiles being about 20 ms, with an overall effect of 13 ms.

The flat delta plot for the RL-TL contrast is consistent with perceptual accumulation modeling by Gomez and Perea in Grade 2 and 4 children (2020; see also Gomez et al., 2013, for evidence with adult readers). According to their account (based on the drift–diffusion model of the lexical decision task, Ratcliff, et al., 2004), masked priming produce a shift in the RT distribution because the prime head-starts the encoding of letters and other orthographic features.[4]

[4] This shift in the RT distribution is consistent with an additive process; for this reason, we decided not to perform reciprocal transformations (see for discussion of this issue, Gellman & Hill, 2007; Gomez & Perea, 2020; Tan & Yap, 2016).

Relationship between performance in the priming experiment and reading test scores

As expected, the priming effects are quite clear in the response times to words. The Bayes factors provide overwhelming evidence in support of the presence of transposed-letter effects in masked priming. These findings align with the existing literature, as the transposed-letter priming effect is considered one of the most robust phenomena in the word recognition literature.

Given the robustness of the transposed-letter effect and the hypothesis that it may be linked to the precision of letter position coding in orthographic processing—which may vary across individuals—it becomes natural to ask if there is a relationship between the transposed-letter effect and measurements of reading skill. To examine the relationship between the standardized scores and the results of the experiment, we summarized the data from the experiments with six measurements: the overall average RT for words, the overall RT for nonwords, the total accuracy for words, the total accuracy for nonwords, the TL effect for words on RT (RT RL-TL), the TL effect for words on RT (RT TL-ID). Additionally, we considered the pseudoword reading score in the Prolec test and the word reading score in the Prolec test.

We begin by presenting the pair-wise graph to summarize the data from the experiment (see Fig. 5). Pairs-graphs have the advantage of providing multiple ways to visualize the data. The scatter plots are shown on the bottom half of the figure (each point represents a participant). A numerical summary of the pair-wise relationships (Pearson's *r*) is shown in the top half of the figure. In addition, the distribution of the measurements' scores is shown in the diagonal (well-behaved distributions in all measurements can be seen).

Notably, the strongest relationships are among the related measurements (e.g., overall RTs to words and overall RTs to nonwords, or the two scores of the standardized reading test), but the TL priming effects on RT and sizes yield weak relationships with the standardized tests ($p > 0.24$). Pair-wise correlations might obscure more complex covariations; so, as in the Gómez et al. (2021) study, we conducted a path analysis to test the critical question of the relationship between test scores and the experimental results. Path modeling is a method used to describe the dependencies between exogenous variables and dependent variables. We selected four dependent variables: total accuracy and latency for words, and TL-effects measured as RL-TL and as TL-ID. Additionally, we chose two exogenous variables: test scores for PROLEC-words and PROLEC-pseudowords. This model is depicted in Fig. 6 and Table 2. The arrows in the graph represent dependencies, and the coefficients on the lines are measurements of statistical relationships (also graphically illustrated by the width of the line). The results of these analyses are clear when examining the paths that start from each of the test scores:

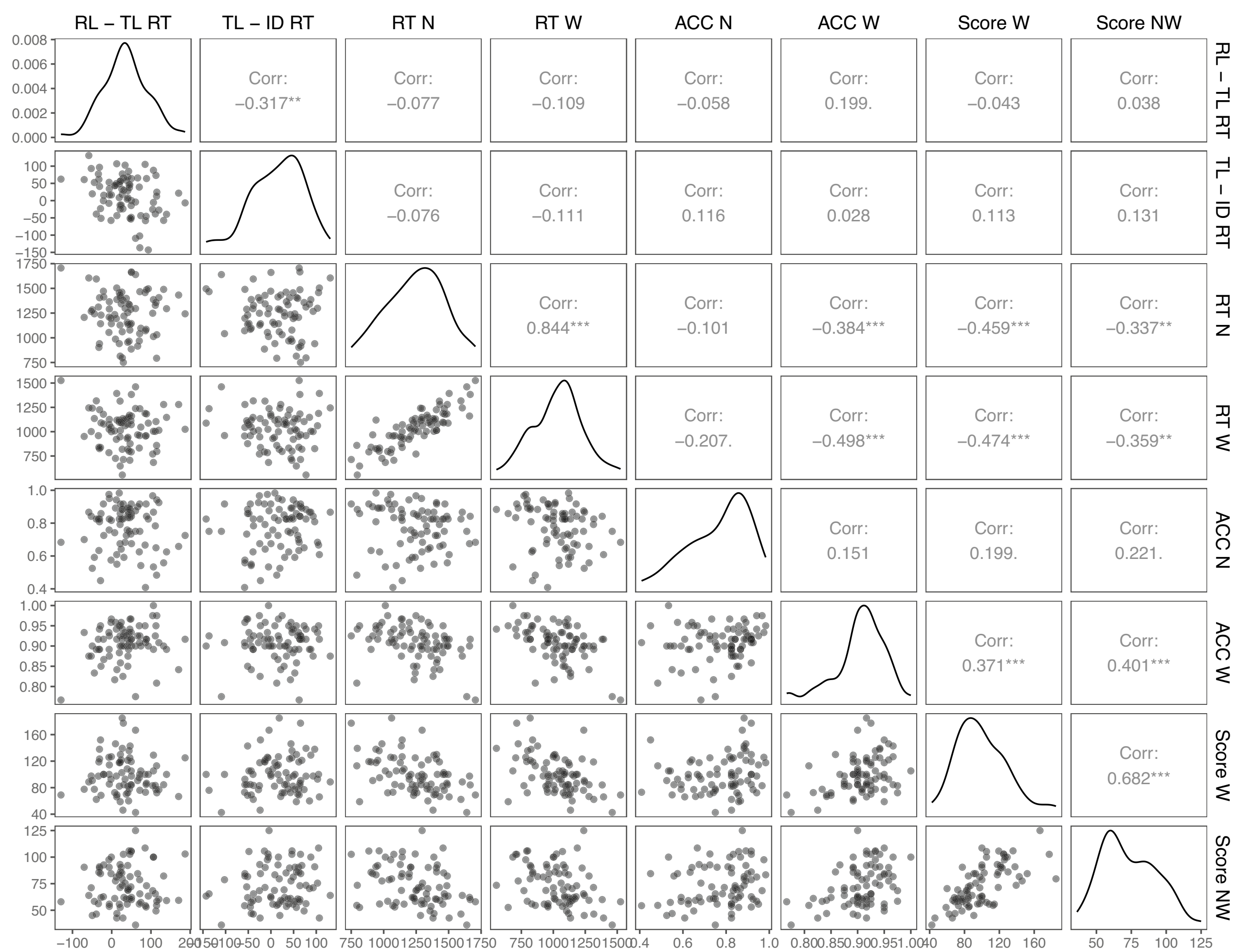

Fig. 5 Pairs plot among the priming effect sizes, overall performance measurements, and reading scores in Study 1. The plot illustrates the relationships among these variables, providing insights into their correlations and distributions

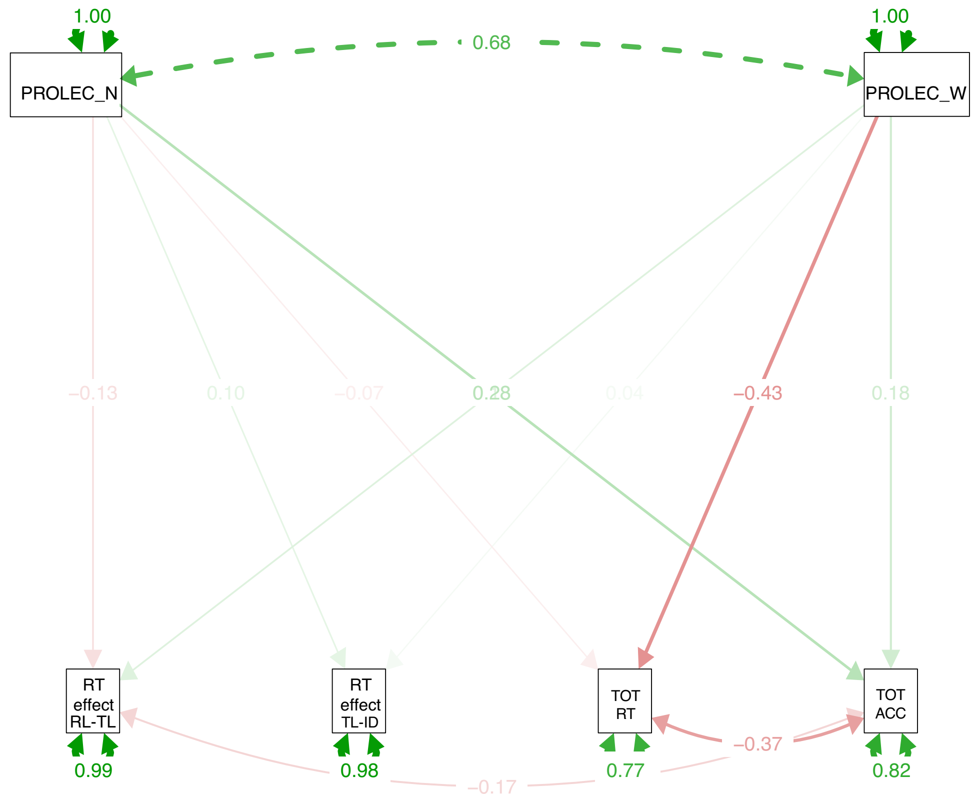

Fig. 6 Path model for Study 1 (see Table 2 for the details in the implementation)

1. For both the PROLEC-word reading and the PROLEC-pseudoword reading, there are only negligible dependencies with the TL effect sizes (all $|z| < 1.3$, $ps > 0.1$).
2. For the PROLEC-word reading, the dependency with the total RT for words is strong ($z = -3.111$, $p = 0.002$).
3. For the PROLEC-pseudoword reading, the only modestly strong dependency is with the total accuracy ($z = 1.967$, $p = 0.049$).

Summary of findings

We found a sizable priming effect for word targets, with identity primes resulting in the best performance in terms of latency and accuracy, followed by transposed-letter primes and, finally, replacement-letter primes. More importantly, the path analysis revealed that the word reading score strongly correlated with total accuracy and total RT for both words and nonwords (see Fig. 6). However, there were negligible dependencies for the TL effect sizes. The pseudoword reading score had a modestly strong dependency only on the total accuracy of words.

In sum, while we found a relationship between overall performance and reading skill, we found very little evidence that the size of transposed-letter priming (i.e., a proxy for letter position coding in the first processing stages) was modulated by reading skills in Grade 6 children.

Study 2. Masked priming with adult readers

To assess the generalizability of our findings from Study 1, we conducted a parallel analysis using data from the Adelman et al. (2014) mega-study of masked form priming on adult readers. We selected three priming conditions from their study that matched those in Study 1: medial transposition, central double substitution, and identity primes. Although the analysis plan was not preregistered, we followed the same procedures as in Study 1, adapting them to the archival data. For brevity, the complete data processing and graphical representations can be found in the online Appendix. Here, we focus on the main point: examining the relationship between the priming effects and reading skills in the Adelman et al. (2014) data. Importantly, Adelman et al. (2014) collected reading scores from two different tests: a vocabulary test based on Shipley (1940) and a spelling test (Burt & Tate, 2002). Thus, together with spelling abilities, the current study will also help us establish the potential role of vocabulary size on the magnitude of the transposed-letter priming effect.

Method

For details on the methods, we refer the reader to Adelman et al.'s (2014) paper; here, we present only the aspects of the study that are relevant to our research question.

Table 2 The covariance matrix, means, and SDs for the path model variables in Study 1 ($N = 79$)

Covariance

	Eff_RLTL	Eff_TLID	RT_N	RT_W	ACC_N	ACC_W	PROLEC_W	PROLEC_N
Eff_RLTL	3229.96774	1022.18060	975.77267	1185.53736	0.44964	-0.49949	68.76068	-41.53340
Eff_TLID	1022.18060	3217.96782	-957.44560	-1205.49217	0.89959	0.07133	183.00762	143.07680
RT_N	975.77267	-957.44560	49671.14098	36139.21361	-3.06788	-3.78033	-2910.70007	-1443.06859
RT_W	1185.53736	-1205.49217	36139.21361	36920.76877	-5.43046	-4.22126	-2589.61242	-1322.78580
ACC_N	0.44964	0.89959	-3.06788	-5.43046	0.01869	0.00091	0.77252	0.58019
ACC_W	-0.49949	0.07133	-3.78033	-4.22126	0.00091	0.00195	0.46543	0.34002
PROLEC_W	68.76068	183.00762	-2910.70007	-2589.61242	0.77252	0.46543	807.96465	372.30485
PROLEC_N	-41.53340	143.07680	-1443.06859	-1322.78580	0.58019	0.34002	372.30485	368.34003

Selected model fit measurements

Npar	Chi^2	df	p value	RFI	NFI	AIC	RMSEA	ntotal
15	10.961	3	0.012	0.185	0.825	2389.763	0.187	79

Selected parameter estimates

	lhs	op	rhs	est	se	z	pvalue
1	Eff_RLTL	~	PROLEC_N	-0.372	0.462	-0.805	0.421
2	Eff_RLTL	~	PROLEC_W	0.257	0.312	0.822	0.411
3	Eff_TLID	~	PROLEC_N	0.299	0.460	0.650	0.516
4	Eff_TLID	~	PROLEC_W	0.089	0.310	0.287	0.774
5	RT_W	~	PROLEC_N	-0.658	1.381	-0.476	0.634
6	RT_W	~	PROLEC_W	-2.902	0.933	-3.111	0.002
7	ACC_W	~	PROLEC_N	0.001	0.000	1.967	0.049
8	ACC_W	~	PROLEC_W	0.000	0.000	1.288	0.198

Model

```
Eff_RLTL ~ PROLEC_N + PROLEC_W
Eff_TLID ~  PROLEC_N + PROLEC_W
RT_W ~ PROLEC_N + PROLEC_W
ACC_W ~ PROLEC_N + PROLEC_W
Eff_RLTL  ~~ 0* Eff_TLID
Eff_RLTL  ~~ 0* RT_W
Eff_TLID  ~~ 0* RT_W
```

Participants

We used 964 of the more than 1000 participants in the megastudy. We followed the authors' selection criteria, which involved removing participants with an overall accuracy lower than 75% in the experiment and participants from one of the sites (Nebraska); however, we used all participants instead of aiming for a balanced number of participants per stimuli list.

Materials and procedure

Adelman et al. (2014) used two measurements of reading ability. The first measurement was a spelling test, which presented items with inconsistent or unusual spelling, based on Burt and Tate (2002). The target stimuli consisted of 42 words and 40 nonwords, and participants had to decide if each target stimulus was spelled correctly or misspelled. The spelling score used for our analyses was the number of correct responses.

The second measurement was a vocabulary test based on Shipley's (1940) test, which included 40 target words of "increasing difficulty (from *TALK* to *PRISTINE*)". Participants were asked to select a synonym for the target word from four options (one correct option and three foils). The vocabulary test score used for our analyses was the number of correct responses.

In the lexical decision priming study by Adelman et al. (2014), the number of target words was 420, all six letters long. Each participant responded to 840 trials (420 target words and a set of 420 target nonword foils). The conditions selected for our analyses were the "central letter substitution" (labeled as DSN-M in their data) and the "medial letter transposition" (TL-M). For both words and nonwords, there were 15 trials per type of prime.

The trial structure followed a typical masked priming procedure: a fixation point was presented for 300 ms, followed by a 200 ms blank display, then a 500-ms mask (######), and finally, the prime in lowercase and 5/8 size for 50 ms. This was followed by a target string in uppercase, which remained on the screen until the participant responded, with a response deadline of 2 s.

Results

Figure 7 shows that the vocabulary and spelling tests have a good spread and are reasonably normal; hence, both scores can be used to explore the relationships among these scores and the masked priming effects.

Adelman et al. reported a 17-ms difference in the response time (RT) for transposed-letter vs. replacement-letter primes. Our re-analysis matched their finding, yielding a BF_{10} of 2e + 9 when contrasting these two conditions. When contrasting the identity (ID) vs. the transposed-letter (TL) condition, there was an advantage of 12 ms in favor of the identity priming condition. This result is fully consistent with the model that assumes a difference, with a BF_{10} of 9e + 4. More importantly, when using the entire data set (encompassing all priming conditions), Adelman et al. reported higher correlations between overall speed and both vocabulary and spelling tests than between such tests and any of the priming sizes. It is worth noting that they based all their analyses using unrelated primes as nonwords.

In terms of the Delta plots (Fig. 8), the Adelman et al. data show the flat character that Gomez et al. (2013), Gomez and Perea (2020) and Study 1 also feature. Note that this has become a highly replicated pattern of data.

Pertaining to the issue at hand, specifically, whether there is an association between an individual's reading skills and the transposed-letter priming effect, the correlations are similar to those found in Study 1. Figure 9 illustrates a robust correlation between the two components of the reading test (Vocabulary and Spelling) and the overall average response time (RT). However, there is virtually no correlation between the transposed-letter priming effect and either of the two reading tests. In conclusion, the correlation between individuals' scores on the reading tests and the magnitude of transposed-letter priming effects is nearly zero, a finding consistent across a large sample of almost one thousand participants. Similarly, the path analysis (Table 3 and Fig. 10) shows a very similar pattern to that in Study 1: sizable relationships among the tests and the overall measurements of speed and accuracy, but no relationship of the priming effect sizes with either test.

Study 3. Post hoc analysis of the reliability of TL effects

In the studies presented here, and in other word recognition paradigms (visual and non-visual; see e.g., Baciero et al., 2022, for evidence in braille) and sentence reading paradigms (e.g., parafoveal previews, as in the Lee et al., 2024, experiment), the transposed-letter effect has been

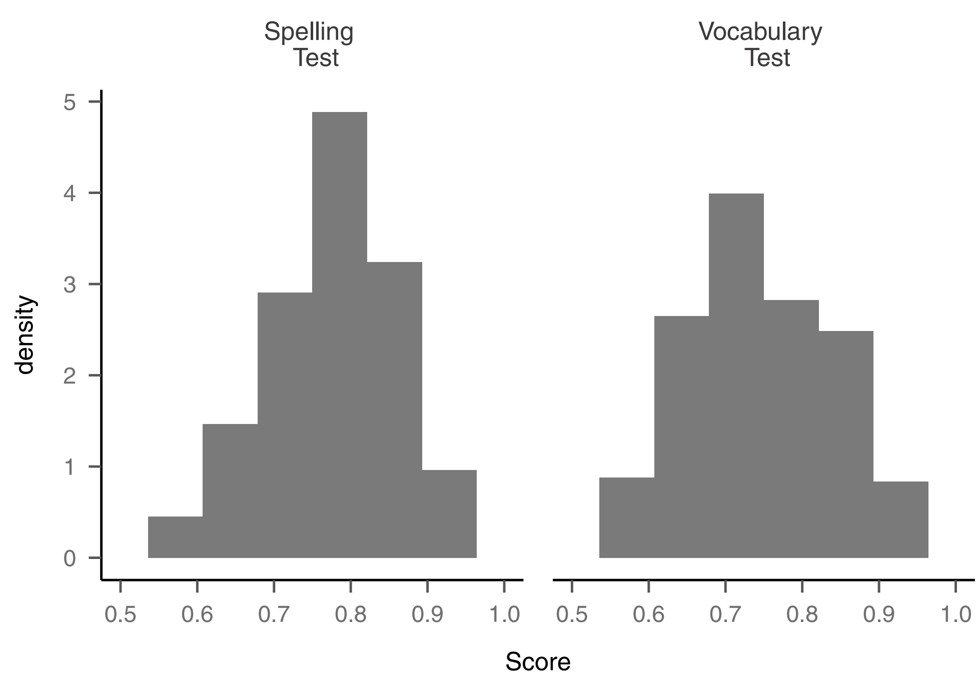

Fig. 7 Distribution of scores in the spelling and vocabulary test used in the Adelman et al. (2014) study

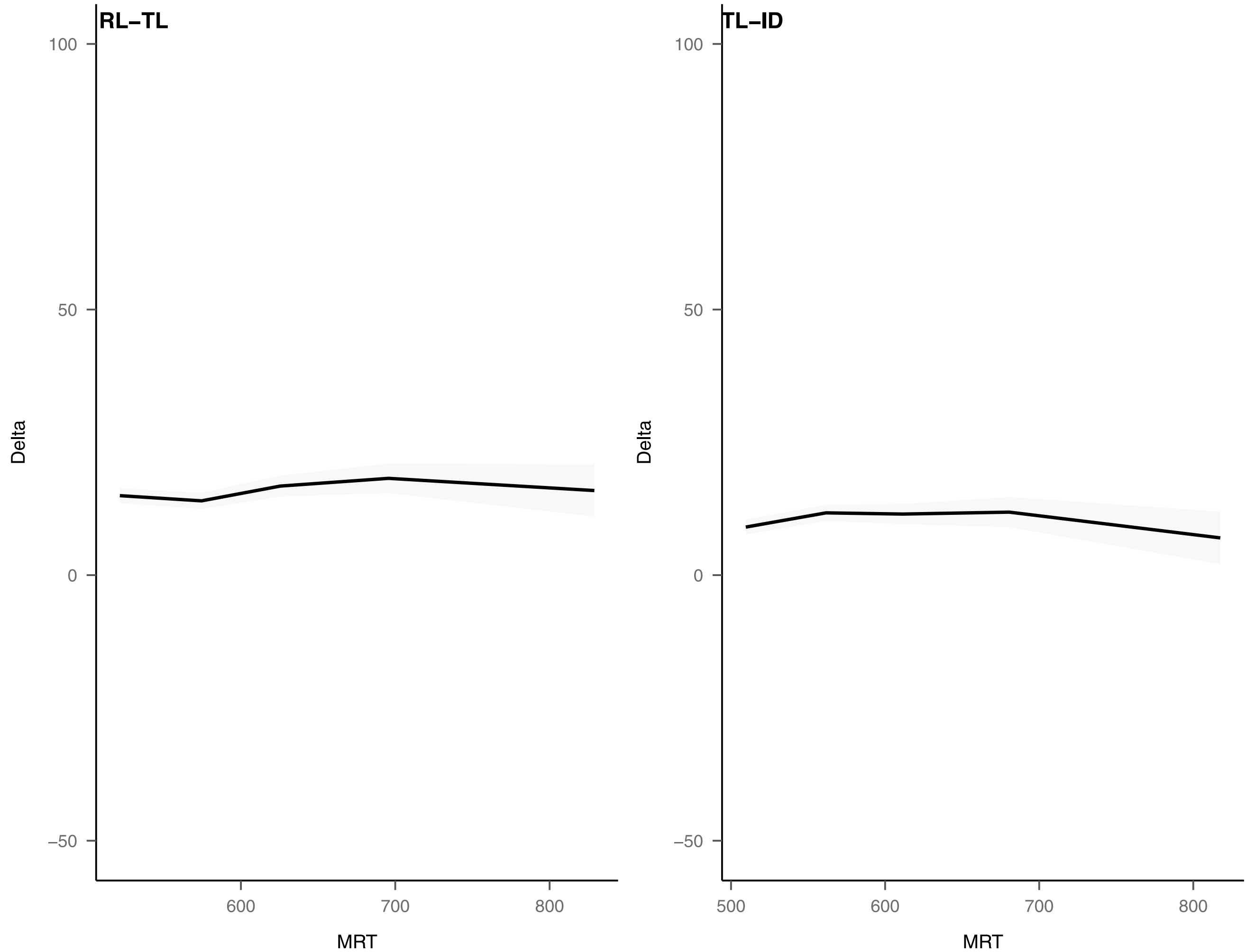

Fig. 8 Delta plot illustrating the difference in response between the conditions labeled on the top of each panel for Study 2. The *x*-axis represents time intervals of latency, while the *y*-axis represents the delta in response time between the conditions. The flat nature of the plot suggests that the effect of type of prime on RTs is consistent regardless of the relative speed of the responses

highly robust and not easily susceptible to efforts to reduce it or eliminate it (e.g., see Marcet et al., 2019). Yet, it does not seem to reliably relate to measurements of reading skill in the initial moments of processing—as tapped by masked priming. One possibility is that the cross-participant effects are small and drowned by trial variability. To explore this possibility, we performed post hoc and relatively simple reliability analyses by performing split-half correlations like in the work by Stolz et al. (2005). The results are presented in Table 4: the split-half correlations tended to be rather low—the only exception is the Gomez and Perea (2020) dataset in which the transposed-letter effect was measured as the difference in accuracy for the TL-nonwords versus the RL-nonwords in a lexical decision task. Notably, this is the dataset that yields the most sizable relationship with the reading score, of those analyzed in Table 4.

The question then is why do the data in Study 1 and Study 2 show such low split-half variability? Briefly put, the answer is that, in the true priming effects (what we termed θ_{Pablo}, θ_{Ana}, θ_{Fran}, and θ_{Manolo} in the introduction for the authors of this paper), there is trial and item variability that is much larger than the variability due to the individual differences. Of course, we cannot claim that this variability in true effects is zero, but it seems to be too small to be detectable with the number of trials we can realistically obtain from children in this study, or with the number of trials reanalyzed in Study 2; what we can say is that the variability across participants is too small, relative to the across-trial variability, to be detectable. In Appendix B, we present this explanation in more formal terms.

For completeness' sake, and to explore if there might be distributional features of the RT that hide the relationship between reading test and priming effects, we generated

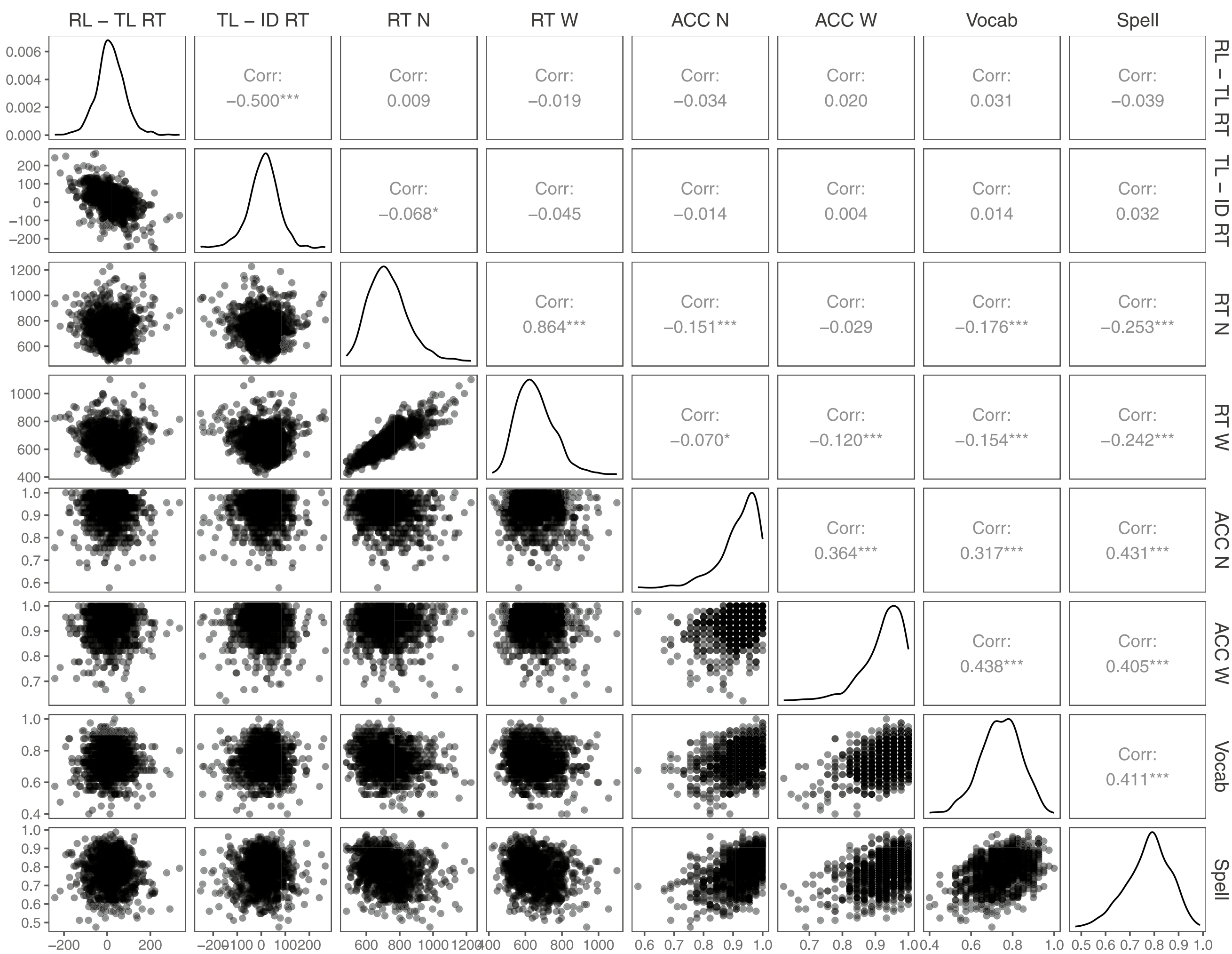

Fig. 9 Pairs plot among the priming effect sizes, the overall performance measurements, and the reading scores in Study 2

separate Delta plots for different groups of participants in Study 1: those that scored below the median in both the Word and the Pseudoword scales of the PROLEC ($n = 33$), those that scored below the median in one of the scales but not the other ($n = 14$), and those that scored above the median in both tests ($n = 32$). Figure 11 shows the TL-RL delta-plot for the three groups; there is no evident distributional difference among the three groups. This pattern of data supports the idea that, as far as we can measure it, the transposed-letter priming effect does not seem to be different among these three groups.

General discussion

One significant limitation in current computational models of visual word recognition is their lack of concern for the impact of individual differences. This oversight stems from insufficient empirical evidence to substantiate which processes are particularly susceptible to these influences (i.e., which processes may be mostly invariant across participants and which processes are subject to individual differences; see Rouder et al., 2009). In this context, a key research issue in the domain of reading research involves exploring the potential modulation of orthographic markers, such as the encoding of letter position (via the transposed-letter effect), by individuals' reading abilities during the initial phases of word identification.

To shed light on this issue, we employed the masked priming technique within the most commonly utilized word recognition paradigm—lexical decision. In Study 1, we recruited a moderately large cohort of Grade 6 children in Spain's final year of mandatory primary education under the premise that they would exhibit substantial individual differences in reading. Additionally, we utilized the 'sandwich' variant of the masked priming technique, as prior research has consistently demonstrated its ability to enhance the size of priming effects. By using this paradigm, we aimed to

Table 3 The covariance matrix, means, and SDs for the path model variables in Study 2 ($N=964$)

Covariance

	Eff_RLTL	Eff_TLID	RT_N	RT_W	ACC_N	ACC_W	Vocab	Spell
Eff_RLTL	4335.79931	-2139.86865	68.04664	-122.66070	-0.14158	0.07952	0.19991	-0.22808
Eff_TLID	-2139.86865	4220.75036	-520.62673	-291.93114	-0.06027	0.01608	0.08597	0.18032
RT_N	68.04664	-520.62673	13710.56789	10104.11386	-1.12799	-0.20619	-2.00109	-2.61479
RT_W	-122.66070	-291.93114	10104.11386	9981.37776	-0.44564	-0.72661	-1.49554	-2.13267
ACC_N	-0.14158	-0.06027	-1.12799	-0.44564	0.00409	0.00141	0.00197	0.00243
ACC_W	0.07952	0.01608	-0.20619	-0.72661	0.00141	0.00368	0.00259	0.00216
Vocab	0.19991	0.08597	-2.00109	-1.49554	0.00197	0.00259	0.00945	0.00352
Spell	-0.22808	0.18032	-2.61479	-2.13267	0.00243	0.00216	0.00352	0.00776

Selected model fit measurements

Npar	Chi^2	df	p value	RFI	NFI	AIC	RMSEA	ntotal
15	281.817	3	0.000	1.000	0.623	30215.658	0.310	964

Selected parameter estimates

	lhs	op	rhs	est	se	z	pvalue
1	Eff_RLTL	~	Vocab	38.658	23.889	1.618	0.106
2	Eff_RLTL	~	Spell	-46.928	26.354	-1.781	0.075
3	Eff_TLID	~	Vocab	0.531	23.608	0.022	0.982
4	Eff_TLID	~	Spell	22.995	26.045	0.883	0.377
5	RT_W	~	Vocab	-67.281	35.174	-1.913	0.056
6	RT_W	~	Spell	-244.284	38.804	-6.295	0.000
7	ACC_W	~	Vocab	0.204	0.019	10.719	0.000
8	ACC_W	~	Spell	0.186	0.021	8.845	0.000

Model

```
Eff_RLTL ~ Vocab + Spell
Eff_TLID ~  Vocab + Spell
RT_W ~ Vocab + Spell
ACC_W ~ Vocab + Spell
Eff_RLTL  ~~ 0* Eff_TLID
Eff_RLTL  ~~ 0* RT_W
Eff_TLID  ~~ 0* RT_W
```

maximize the likelihood of detecting any potential association between observed transposed-letter priming effects and individual reading scores. In Study 2, we re-analyzed data from a mega-study involving over 900 adult participants using the standard masked priming method (Adelman et al., 2014). We hoped the large sample size would reveal a stronger relationship between reading skills and priming effects (if real).

The results of Study 1 revealed a reasonably large transposed-letter effect (RL priming condition vs. TL priming condition), which can be characterized as a shift in the response time (RT) distributions (see the left panel of the delta plots in Fig. 4). This outcome aligns with the view that the locus of the obtained masked priming effects is at the encoding stage (see Gomez et al., 2013; Gomez & Perea, 2020, for similar evidence with masked repetition priming in both adult and developing reader groups, respectively). Regarding the relationship between reading scores and performance, there was an overall advantage, with better test scores being associated with faster response times in both experiments. However, more critically, we found no evidence of a relationship between reading scores and the size of the transposed-letter effect. This same null interaction occurred when we examined whether the difference between the transposed-letter priming condition and the identity condition was modulated by reading skill.

Study 2 utilized the same three critical priming conditions (identity, transposed letter, replacement letter) from a masked form-priming lexical decision mega-study (Adelman et al., 2014). This study used the conventional masked priming paradigm with a large number of individuals and also collected measures of reading skill, namely, scores in spelling abilities and vocabulary. The results from Study 2 are almost identical to those of Study 1: there is a

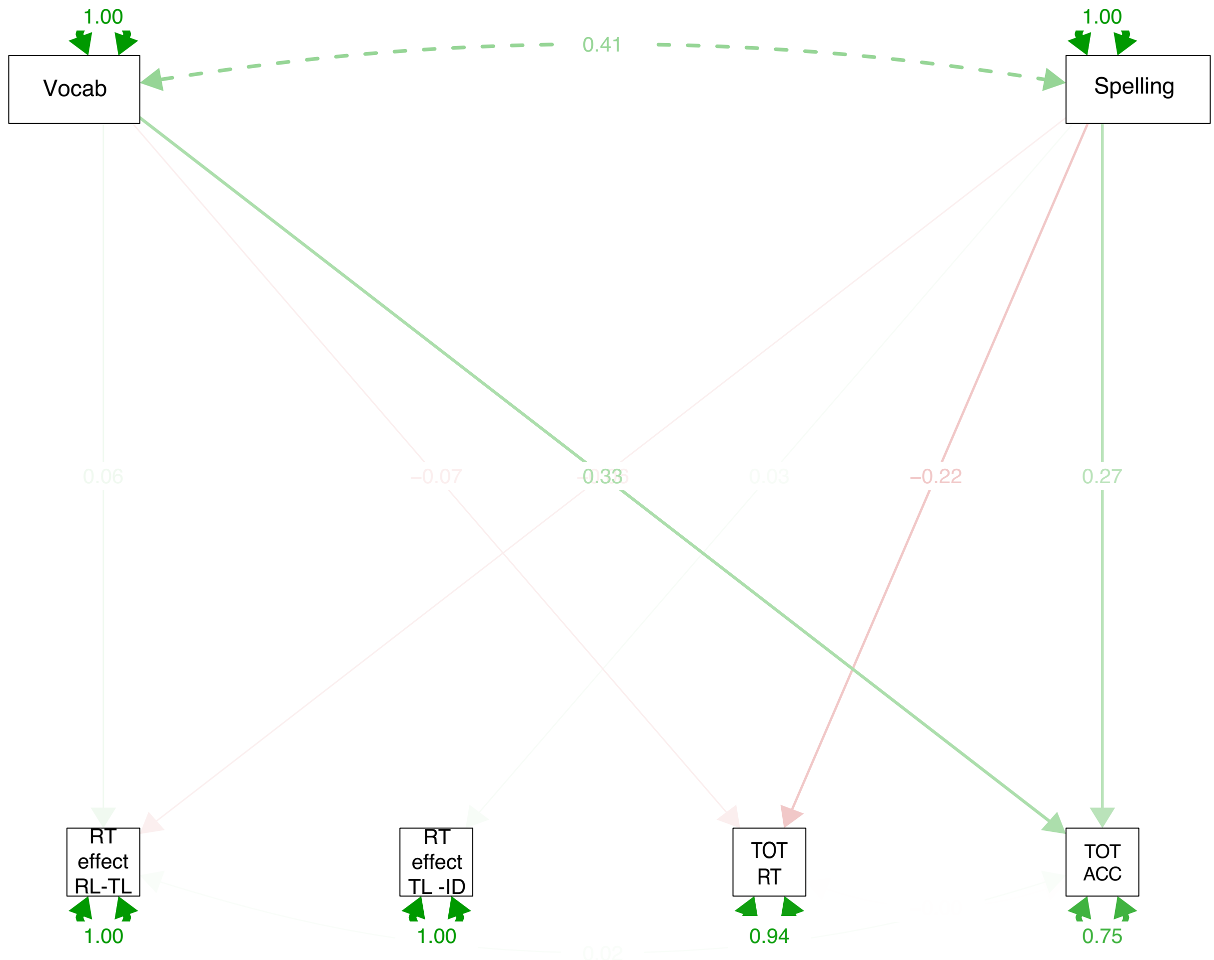

Fig. 10 Path model for Study 2 (see Table 3 for the details in the implementation)

Table 4 Split-half correlations for published transposed-letter effects

Data Set	Description	Measurement	TL-RL split half correlation	ID-TL split half correlation
Study 1	Sandwich priming with children (ID, TL and RL primes in internal letter positions)	RT to target words	0.053 ($p=0.64$)	−0.083 ($p=0.47$)
Study 2 (subset of Adelman et al., 2014)	Masked priming with adults (ID, TL and RL primes in internal letter positions)	RT to target words	0.049 ($p=0.13$)	0.0009 ($p=0.98$)
Gómez et al. (2021)	Lexical decision task (single-presentation), TL vs RL nonwords	RT to nonwords Error rate to nonwords	0.099 ($p=0.40$) 0.597 ($p<0.001$)	

transposed-letter priming effect (smaller in this case, likely due to the use of the conventional variant of the masked priming technique), an advantage of the identity condition over the transposed-letter priming condition, and a substantial correlation between test scores and overall response time (RT). More critically, we again found no evidence of a correlation between the reading scores and the observed priming effect. It should be noted that in this study, there were two distinct reading skill tests: vocabulary and spelling (their association was $r=0.441$), whereas in Study 1, there were two subcomponents of a standardized reading test: word reading and pseudoword reading (their association was $r=0.682$).

Taken together, we have established several important positive empirical effects, such as a transposed-letter effect (with transposed-letter primes being more effective than

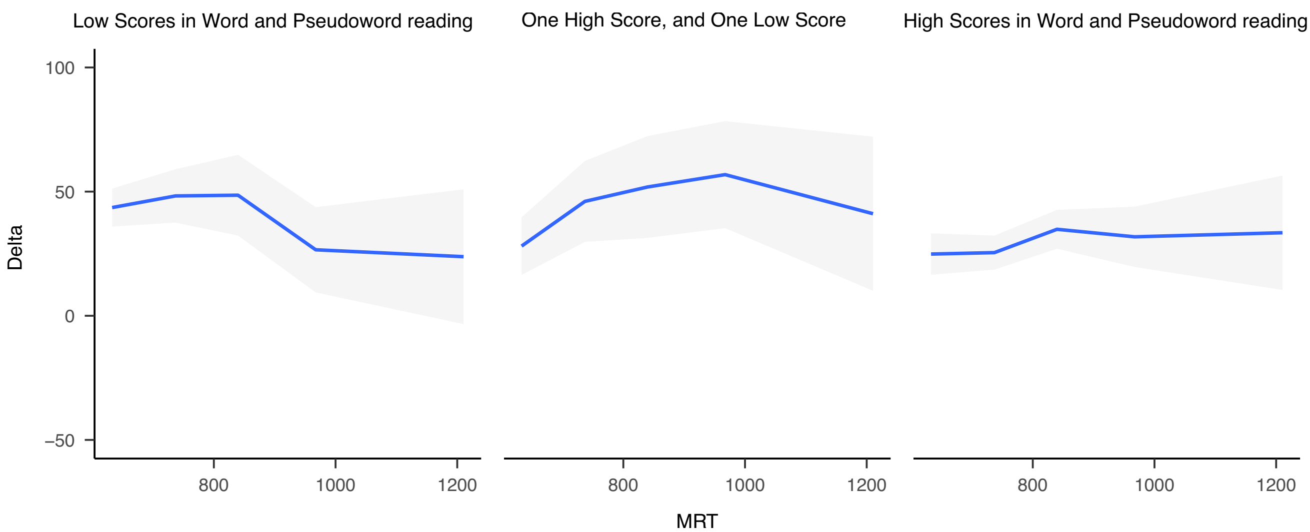

Fig. 11 Delta plot illustrating the difference in response between the RL and TL conditions in Study 1, for three different groups of participants: those that scored below the median in both the Word and the Pseudoword scaled of the PROLEC (left panel), those that scored below the median in one of the scales but not the other (middle panel), and those that scored above the median in both tests (right panel). The *x*-axis represents time intervals of latency, while the *y*-axis represents the delta in response time between the conditions. Just like in the global delta-plot in Fig. 4, the flat nature of the plot suggests that the effect of type of prime on RTs is consistent regardless of the relative speed of the responses, and of the group membership

replacement-letter primes) and an advantage of the identity primes over replacement-letter primes. In addition, we have established a clear relationship between performance in the lexical decision task (as measured by the overall response time, RT) and the measurements of reading skill (in Study 1: $r > |0.4|$, and in Study 2: $r > |0.14|$). This finding is, of course, what we would hope for if we believe that our lexical decision experiments can provide insights into reading.

The main questions under consideration for this article yielded a null finding that is considered in this section. When considering the difference in RTs between each participant's transposed-letter and replacement-letter conditions, we found that this effect did not correlate with the measurements in either experiment. This was also the case when considering the differences between the transposed-letter condition against the identity condition. One might argue that, in Study 1, we did not have enough participants to find a relationship; however, in Study 2, with a mega-study, this concern should be alleviated. Overall, these findings are consistent with the evidence that Andrews and Lo (2012) reported, who found a negligible relation between reading skills and transposed-letter priming effects for pseudoword primes (measured against an unrelated control) in adult readers. Andrews and Lo only found an association with reading skills when restricted to prime words with transposed-letter neighbors (e.g., `clot-CLOT` vs. `punt-CLOT`). This latter finding may suggest that any potential modulation from individual differences in masked priming may arise from higher processing levels (e.g., via inhibition from neighboring word primes)—unfortunately, the small number of transposed-letter word pairs in European languages (e.g., `clot` and `colt`) limits the generality of this approach. Importantly, the present findings are consistent with recent evidence during sentence reading using Rayner's (1975) gaze-contingent boundary change technique (i.e., a technique that explores early word processing in the parafovea rather than the fovea). In a group of 100 adult readers, Lee et al. (2024) found that eye fixation times on a target word (e.g., `monkey`) were longer when preceded by a transposed-letter parafoveal preview (e.g., `mnokey`) than a replacement-letter parafoveal preview (e.g., `markey`), and that eye fixation times were shorter for the better readers. Critically, as occurred here, Lee et al. (2024) found no signs that the transposed-letter priming effect was shaped by the participants' reading ability.

Nonetheless, prior work has shown some relationship between reading and masked priming effects, so it is worth addressing the particularities of those studies. Before doing that, as indicated in the Introduction, our group reported a relationship between transposed-letter effects and reading skill (Gómez et al., 2021) using single-presentation lexical decision, but this was limited to the error rate: the difference between the error rate for replacement-letter pseudowords and transposed-letter pseudowords was computed, and it correlated with the reading scores. However, it is important to note that single-presentation and masked priming experiments involve different processes; masked priming may relate to early encoding processes with minimal individual variation, whereas performance in the single-presentation lexical decision relates to wordlikeness in the stimulus string, which can be influenced by later-occurring lexical (or post-lexical) effects in which reading skill may play an important role. In the context of masked priming experiments, when this technique has been used to argue for "individual differences"—typically in the context of morphological priming, what researchers have tended to do is variations of median splits, which divide participants into two groups (e.g., Kahraman & Kırkıcı, 2021; Medeiros & Duñabeitia, 2016; but also see Beyersmann et al., 2015 for an approach using linear mixed-effects). This approach is rather coarse, and we would argue that group-level differences (interesting in their own right) are not the same as individual differences.

We conceptualize individual difference research as locating the true effect for each individual, which might be nearly impossible to do when variability in the process is small and variability in the data is large, particularly when looking at effects that compare two conditions that do not produce the most extreme effects (see, for example, an analysis by Tan & Yap, 2016 comparing masked repetition priming to masked semantic priming). In a recent study by Hasenäcker and Schroeder (2022) that used a longitudinal approach, the findings related to the transposed-letter effect and orthographic knowledge were inconclusive. The analyses of the two studies presented here suggest that masked priming effects, both with the sandwich and conventional versions, are related to the encoding process. This is demonstrated by the delta plots in Study 1, in which the sample is split into three groups based on their performance in the reading test (see Fig. 11). These groups show very similar delta plots, which are also quite similar to previous masked priming findings with the lexical decision task (i.e., an approximate flat line; see Gomez et al., 2013, for modeling). This encoding process may have relatively low cross-individual variability, thus making any correlation with reading skill difficult to extract and, in practice, quite close to zero. This reasoning is consistent with roughly similar sizes of masked repetition effects in lexical decisions for Grade 2 and Grade 4 children compared to the adult data, despite the vast differences in the overall response times (see Gomez & Perea, 2020). The questions and methodologies that have underpinned cognitive psychology since its beginnings primarily focus on establishing whether an effect, such as the transposed-letter effect, is present at a population level (Rouder & Haaf, 2021), (Rouder andMehrvarz2024), (Hedge et al. 2018). By identifying which components of a cognitive task are functionally or statistically invariant across individuals, we can develop more nuanced models of human cognition. In the context of this study, such an approach advances our understanding of the initial stages of word recognition.

Importantly, the focus of this paper is on the experimental task and on the variability of its effect across individuals. There is yet another source of noise that interferes with our goal: Reading tests can vary widely in what they measure, encompassing aspects such as decoding skills, comprehension, fluency, and vocabulary. Establishing content validity is particularly difficult because reading is a multifaceted skill, and different tests may prioritize different components (see Lee et al., 2024, for discussion). Having said this, it is worth noting that the Prolec test did show modulation of transposed-letter effects in a single-presentation (unprimed) lexical decision task with Grade 6 children (Gomez & Perea, 2020). This suggests that the test itself may not be the key factor, but rather the paradigm used. Masked priming effects seem to show minimal individual differences among neurotypical participants, as seen in similar masked repetition priming results for children in Grade 2, Grade 4, and adults (Gomez & Perea, 2020), despite differences in overall reaction times and reading skills. In addition, in Study 2, we also did not find a relationship between masked transposed-letter priming effects and reading skill in adult individuals using other tests (vocabulary and spelling) (see Lee et al., 2024, for a similar pattern with a parafoveal preview paradigm using other tests: the Wechsler Individual Achievement Test, the Nelson Denny Reading Test, and the LexTale). We acknowledge, however, that future reading tests specifically focused on letter order information could be more sensitive than the standard reading tests employed in the above-cited studies.

In conclusion, our findings provide compelling evidence that while transposed-letter priming effects and individual reading skills are significant elements in visual word recognition, their interrelation is minimal to non-existent in the initial moments of word processing as measured by masked priming (see Lee et al., 2024, for a similar conclusion regarding parafoveal priming during sentence reading). Despite a wide range of reading abilities, this underscores that the flexibility in the initial encoding of letter position is a universally robust phenomenon in alphabetic orthographies, largely independent of individual reading ability, and with relatively low across-participant variability. This is, after all, in agreement with the seminal article by Kohnen and Castles (2013), which reports a minimal improvement in letter position coding (relative to other reading skills) from Grade 2 to Grade 4, leading them to postulate that loose letter position coding is a "fixed feature of the reading system" (p. 102). Accounting for the variability in these processes should be an agenda for future computational models of cognition. In the overlap model, the noisy locations of objects (e.g., letters) in beginning readers would be reflected as larger s values. Note that in an extreme case, an impairment of the processes underlying letter position coding may lead to letter position dyslexia (see Friedmann & Gvion, 2001, 2005); however, for most readers, the results presented in this paper suggest that letter position coding is encapsulated from word reading, spelling, and vocabulary.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00426-025-02080-w.

Author contributions All authors contributed to the study's conception and design. Material preparation was performed by M.P., A.M., and F.R. Data collection was performed by A.M. and F.R. Statistical analyses were performed by M.P. and P.G. The first draft of the manuscript was written by P.G., and all authors commented on previous versions of the manuscript. P.G. and M.P. made the changes in R1. All authors read and approved the final manuscript.

Funding The research reported in this article has been partially supported by Grant PID2020-116740GB-I00 funded by the MCIN/AEI/https://doi.org/10.13039/501100011033 (Manuel Perea) from the Spanish Ministry of Science and Innovation to Manuel Perea, Grant CIAICO/2021/172 from the Department of Innovation, Universities, Science and Digital Society of the Valencian Government to Manuel Perea, and NSF grant SMA-2127135 to Pablo Gomez.

Data availability The stimuli, data, scripts, and outputs are available at https://osf.io/6dvbt/. Study 1 was pre-registered at https://osf.io/x8tua.

Declarations

Conflict of interest The authors declare no competing interests.

Ethical approval The procedures involving human participants in this study were approved by the Experimental Research Ethics Committee of the Universitat de València (ref # 1894511) and were following the Declaration of Helsinki. The participants' parents provided written informed consent before starting the experiments.

References

Adelman, J. S. (2011). Letters in time and retinotopic space. *Psychological Review, 118*, 570–582. https://doi.org/10.1037/a0024811

Adelman, J. S., Johnson, R. L., McCormick, S. F., McKague, M., Kinoshita, S., Bowers, J. S., Perry, J. R., Lupker, S. J., Forster, K. I., Cortese, M. J., Scaltritti, M., Aschenbrenner, A. J., Coane, J. H., White, L., Yap, M. J., Davis, C., Kim, J., & Davis, C. J. (2014). A behavioral database for masked form priming. *Behavior Research Methods, 46*(4), 1052–1067. https://doi.org/10.3758/s13428-013-0442-y

Andrews, S., & Lo, S. (2012). Not all skilled readers have cracked the code: Individual differences in masked form priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(1), 152–163. https://doi.org/10.1037/a0024953

Baciero, A., Gomez, P., Duñabeitia, J. A., & Perea, M. (2022). Reading with the fingers: Towards a universal model of letter position coding. *Psychonomic Bulletin & Review, 29*, 2275–2283. https://doi.org/10.3758/s13423-022-02078-0

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Beyersmann, E., Grainger, J., Casalis, S., & Ziegler, J. C. (2015). Effects of reading proficiency on embedded stem priming in primary school children. *Journal of Experimental Child Psychology, 139*, 115–126. https://doi.org/10.1016/j.jecp.2015.06.001

Brown, J. A., Fishco, V. V., & Hanna, G. (1993). *Nelson-Denny reading test: Manual for scoring and interpretation, forms G & H*. Riverside Publishing.

Burt, J. S., & Tate, H. (2002). Does a reading lexicon provide orthographic representations for spelling? *Journal of Memory and Language, 46*, 518–543. https://doi.org/10.1006/jmla.2001.2818

Castles, A., Davis, C., Cavalot, P., & Forster, K. (2007). Tracking the acquisition of orthographic skills in developing readers: Masked priming effects. *Journal of Experimental Child Psychology, 97*, 165–182. https://doi.org/10.1016/j.jecp.2007.01.006

Cuetos, F., Rodríguez, B., Ruano, E., & Arribas, D. (2007). *PROLEC-R: Català—Bateria d'avaluació dels processos lectors* [PROLEC-R: Catalan—Battery of assessment of reading processes]. TEA Ediciones.

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review, 117*(3), 713–758. https://doi.org/10.1037/a0019738

De Jong, R., Liang, C.-C., & Lauber, E. (1994). Conditional and unconditional automaticity: A dual-process model of effects of spatial stimulus-response correspondence. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 731–750. https://doi.org/10.1037/0096-1523.20.4.731

Fernández-López, M., Davis, C. J., Perea, M., Marcet, A., & Gómez, P. (2022). Unveiling the boost in the sandwich priming technique. *Quarterly Journal of Experimental Psychology, 75*, 1382–1393. https://doi.org/10.1177/17470218211055097

Forster, K., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(4), 680–698. https://doi.org/10.1037/0278-7393.10.4.680

Forster, K. I., Davis, C., Schoknecht, C., & Carter, R. (1987). Masked priming with graphemically related forms: Repetition or partial activation? *Quarterly Journal of Experimental Psychology, 39*(2), 211–251. https://doi.org/10.1080/14640748708401785

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*, 116–124. https://doi.org/10.3758/bf03195503

Friedmann, N., & Gvion, A. (2001). Letter position dyslexia. *Cognitive Neuropsychology, 18*(8), 673–696. https://doi.org/10.1080/02643290143000051

Friedmann, N., & Gvion, A. (2005). Letter form as a constraint for errors in neglect dyslexia and letter position dyslexia. *Behavioural Neurology, 16*(2–3), 145–158. https://doi.org/10.1155/2005/635634

Gómez, P., Marcet, A., & Perea, M. (2021). Are better young readers more likely to confuse their mother with their mohter? *Quarterly Journal of Experimental Psychology, 74*(9), 1542–1552. https://doi.org/10.1177/17470218211012960

Gomez, P., & Perea, M. (2020). Masked identity priming reflects an encoding advantage in developing readers. *Journal of Experimental Child Psychology, 199*, 104911. https://doi.org/10.1016/j.jecp.2020.104911

Gomez, P., Perea, M., & Ratcliff, R. (2013). A diffusion model account of masked vs. unmasked priming: Are they qualitatively different? *Journal of Experimental Psychology: Human Perception and Performance, 39*, 1731–1740. https://doi.org/10.1037/a0032333

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review, 115*(3), 577–600. https://doi.org/10.1037/a0012667

Grainger, J. (2008). Cracking the orthographic code: An introduction. *Language and Cognitive Processes, 23*(1), 1–35. https://doi.org/10.1080/01690960701578013

Grainger, J. (2018). Orthographic processing: A 'mid-level' vision of reading. *Quarterly Journal of Experimental Psychology, 71*(2), 335–359. https://doi.org/10.1080/17470218.2017.1314515

Grainger, J. (2022). Word recognition I: Visual and orthographic processing. In M. J. Snowling, C. Hulme, & K. Nation (Eds.), *The science of reading: A handbook* (2nd ed., pp. 60–78). John Wiley & Sons Ltd.

Grainger, J., & Ziegler, J. (2011). A dual-route approach to orthographic processing. *Frontiers in Psychology, 2*, 54. https://doi.org/10.3389/fpsyg.2011.00054

Hasenäcker, J., & Schroeder, S. (2022). Transposed and substituted letter effects across reading development: A longitudinal study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 48*, 1202–1218. https://doi.org/10.1037/xlm0001064

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods, 50*(3), 1166–1186. https://doi.org/10.3758/s13428-017-0935-1

Kahraman, H., & Kırkıcı, B. (2021). Letter transpositions and morphemic boundaries in the second language processing of derived words: An exploratory study of individual differences. *Applied Psycholinguistics, 42*(2), 417–446. https://doi.org/10.1017/S0142716420000673

Kezilas, Y., McKague, M., Kohnen, S., Badcock, N. A., & Castles, A. (2017). Disentangling the developmental trajectories of letter position and letter identity coding using masked priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*, 250–258. https://doi.org/10.1037/xlm0000293

Kohnen, S., & Castles, A. (2013). Pirates at parties: Letter position processing in developing readers. *Journal of Experimental Child Psychology, 115*(1), 91–107. https://doi.org/10.1016/j.jecp.2012.08.011

Lee, C. H., Lally, C., & Rastle, K. (2021). Masked transposition priming effects are observed in Korean in the same–different task. *Quarterly Journal of Experimental Psychology, 74*, 1439–1450. https://doi.org/10.1177/1747021821997336

Lee, C. E., Pagán, A., Godwin, H. J., & Drieghe, D. (2024). Individual differences and the transposed letter effect during reading. *PLoS ONE, 19*(2), e0298351. https://doi.org/10.1371/journal.pone.0298351

Llaurado, A., & Tolchinsky, L. (2015). The developmental pattern of spelling in Catalan from first to fifth school grade. *Writing Systems Research, 8*(1), 64–83. https://doi.org/10.1080/17586801.2014.1000812

Lupker, S. J., & Davis, C. J. (2009). Sandwich priming: A method for overcoming the limitations of masked priming by reducing lexical competitor effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 618–639. https://doi.org/10.1037/a0015278

Marcet, A., Perea, M., Baciero, A., & Gomez, P. (2019). Can letter position encoding be modified by visual perceptual elements? *Quarterly Journal of Experimental Psychology, 72*, 1344–1353. https://doi.org/10.1177/1747021818789876

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88*(5), 375–407. https://doi.org/10.1037/0033-295X.88.5.375

Medeiros, J., & Duñabeitia, J. A. (2016). Not everybody sees the ness in the darkness: Individual differences in masked suffix priming. *Frontiers in Psychology, 7*, 1585. https://doi.org/10.3389/fpsyg.2016.01585

Mirault, J., & Grainger, J. (2021). Single word reading in the "real" world: Effects of transposed-letters. *Journal of Cognition, 4*(1), 27. https://doi.org/10.5334/joc.160

Morey, R. D., & Rouder (2018). *BayesFactor: Computation of Bayes factors for common designs*. R package version 0.9.12-4.4. https://CRAN.R-project.org/package=BayesFactor

Pagán, A., Blythe, H. I., & Liversedge, S. P. (2021). The influence of children's reading ability on initial letter position encoding during a reading-like task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 47*(7), 1186–1203. https://doi.org/10.1037/xlm0000989

Perea, M., & Lupker, S. J. (2003). Does jugde activate COURT? Transposed-letter similarity effects in masked associative priming. *Memory & Cognition, 31*, 829–841. https://doi.org/10.3758/BF03196438

Perea, M., & Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *Journal of Memory and Language, 51*, 231–246. https://doi.org/10.1016/j.jml.2004.05.005

Perea, M., Marcet, A., & Gomez, P. (2016). How do Scrabble players encode letter position during *reading?*. *Psicothema, 28*, 7–12. https://doi.org/10.7334/psicothema2015.167

Perea, M., Winskel, H., & Gomez, P. (2018). How orthographic-specific characteristics shape letter position coding: The case of Thai script. *Psychonomic Bulletin and Review, 25*, 416–422. https://doi.org/10.3758/s13423-017-1279-7

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–214). John Benjamins.

Rafel i Fontanals, J. (1998). *Diccionari de Freqüencies* [Dictionary of frequencies]. Institut d'Estudis Catalans.

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review, 111*(1), 159–182. https://doi.org/10.1037/0033-295X.111.1.159

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7*, 65–81. https://doi.org/10.1016/0010-0285(75)90005-5

Rouder, J. N., & Mehrvarz, M. (2024). Hierarchical-model insights for planning and interpreting individual-difference studies of cognitive abilities. *Current Directions in Psychological Science, 33*(2), 128–135. https://doi.org/10.1177/09637214231220923

Rouder, J. N., & Haaf, J. M. (2021). Are There Reliable Qualitative Individual Differences in Cognition? Journal ofCognition, *4*(1): 46, 1–16. https://doi.org/10.5334/joc.131

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*, 225–237. https://doi.org/10.3758/PBR.16.2.225

Shipley, W. C. (1940). A self-administering scale for measuring intellectual impairment and deterioration. *The Journal of Psychology: Interdisciplinary and Applied, 9*, 371–377. https://doi.org/10.1080/00223980.1940.9917704

Spinelli, G., Colombo, L., & Lupker, S. J. (2022). Consonant and vowel transposition effects during reading development: A study on Italian children and adults. *Quarterly Journal of Experimental Psychology, 75*(11), 2023–2042. https://doi.org/10.1177/17470218211066301

Stolz, J. A., Besner, D., & Carr, T. H. (2005). Implications of measures of reliability for theories of priming: Activity in semantic memory is inherently noisy and uncoordinated. *Visual Cognition, 12*, 284–336. https://doi.org/10.1080/13506280444000030A

Tan, L. C., & Yap, M. J. (2016). Are individual differences in masked repetition and semantic priming reliable? *Visual Cognition, 24*(2), 182–200. https://doi.org/10.1080/13506285.2016.1214201

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

Wechsler, D. (2005). *Wechsler individual achievement test (WIAT-II-UK)* (2nd ed.). The Psychological Corp.

Ziegler, J. C., Bertrand, D., Lété, B., & Grainger, J. (2014). Orthographic and phonological contributions to reading development: Tracking developmental trajectories using masked priming. *Developmental Psychology, 50*(4), 1026. https://doi.org/10.1037/a0035187