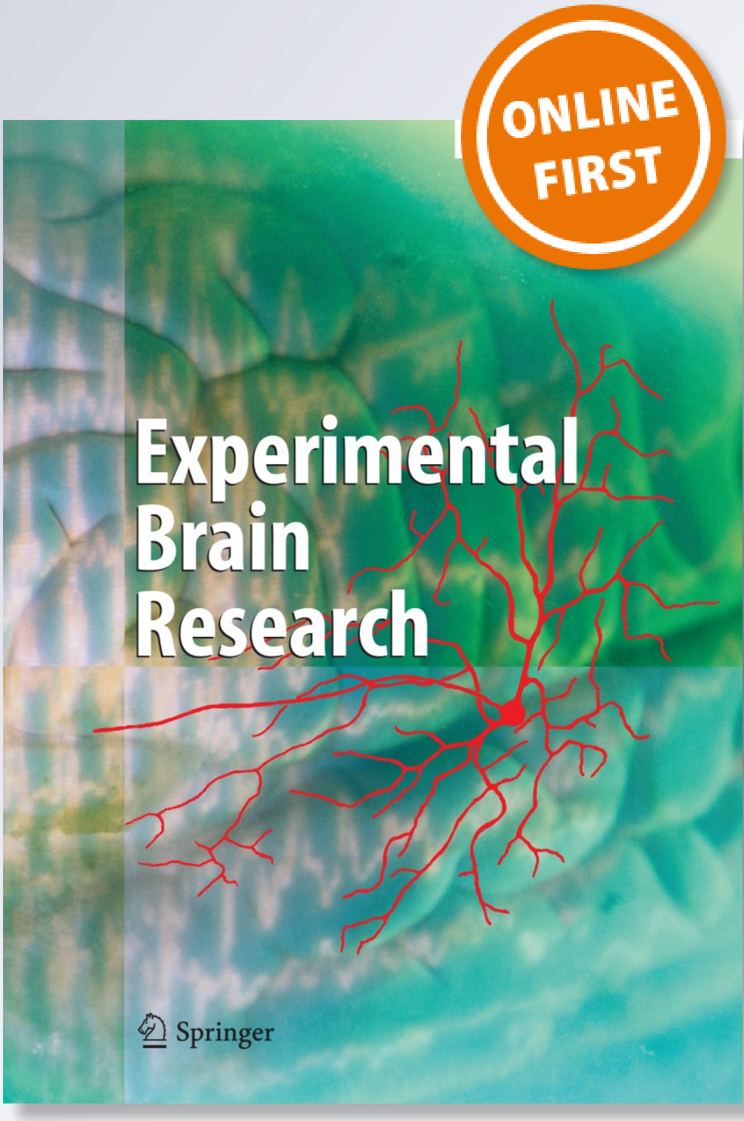
Bridging the gap between different measures of the reading speed deficit in developmental dyslexia

Marialuisa Martelli, Maria De Luca, Laura Lami, Claudia Pizzoli, Maria Pontillo, Donatella Spinelli & Pierluigi Zoccolotti

Experimental Brain Research

ISSN 0014-4819

Exp Brain Res
DOI 10.1007/s00221-013-3735-6



Exp Brain Res
DOI 10.1007/s00221-013-3735-6

RESEARCH ARTICLE

Bridging the gap between different measures of the reading speed deficit in developmental dyslexia

Marialuisa Martelli · Maria De Luca · Laura Lami · Claudia Pizzoli · Maria Pontillo · Donatella Spinelli · Pierluigi Zoccolotti

Received: 16 January 2013 / Accepted: 3 October 2013

Abstract The study assessed how decoding and pronunciation times contribute to total reading time in reading aloud and how these measures change in the presence of developmental dyslexia. Vocal reaction times (RTs), pronunciation times, and total reading times were measured while 25 children with dyslexia and 43 age-matched typically developing readers read singly presented words and non-words that varied for length. Group differences were large for vocal RTs; children with dyslexia were increasingly slower as a function of condition difficulty (over-additivity effect); lexicality and length influenced RTs even when over-additivity was controlled for by *z*-score transformation. The group differences were also large for vocal total reading times, but the effect of over-additivity was smaller than that of vocal RTs and no selective influence of lexicality and length was detected. Pronunciation times showed very small individual differences and no over-additivity effect; children with dyslexia were more sensitive to the effect of lexicality and length than controls. To assess the contribution of the cognitive and sensory–motor compartments in determining group differences, we applied the difference engine model. As for RTs, the relationship between means and standard deviations closely supported the prediction of a general cognitive delay in the slow group, with no group difference in the sensory–motor compartment. The variance in total reading times was predicted by combining the model results for RTs with the linear relationship between pronunciation times and task difficulty. The results help clarify the internal structure of reading times, a measure largely used in clinical testing to assess reading rate.

Keywords Reading speed · Developmental dyslexia · Vocal RT · Pronunciation time · Reading time

M. Martelli (✉) · P. Zoccolotti
Department of Psychology, Sapienza University of Rome, Rome, Italy
e-mail: Marialuisa.Martelli@uniroma1.it

M. Martelli · M. De Luca · D. Spinelli · P. Zoccolotti
Neuropsychology Unit, IRCCS Fondazione Santa Lucia, Rome, Italy

L. Lami · C. Pizzoli
Regional Centre for Linguistic and Cognitive Disabilities, ASL, Bologna, Italy

M. Pontillo
Child Neuropsychiatry Unit, Department of Neuroscience, I.R.C.C.S. Children Hospital Bambino Gesù, Rome, Italy

D. Spinelli
Department of Education Sciences in Sport and Physical Activity, University "Foro Italico", Rome, Italy

Introduction

Children with dyslexia show a pervasive deficit in reading speed across different orthographic materials and conditions. They are slow in reading meaningful texts as well as lists of words (and non-words); severe deficits are also revealed by laboratory tasks, such as single-word reading or lexical decision tasks. Nevertheless, it is difficult to directly compare performance across different tasks because of the use of different dependent measures. The aim of the present study was to assess the different components of reading time (i.e., sensory, motor, and cognitive) that contribute to differences between children with dyslexia and typically developing readers. For this purpose, reaction times (RTs) and pronunciation in reading were measured separately.

Published online: 17 October 2013

In functional reading, the total time to read a text (or a list of words) is divided to obtain a measure of time per item (e.g., s/word or s/syllable). The use of these measures in groups of children with dyslexia who are native speakers of Italian (the language studied here) results in delays (with respect to chronologically age-matched typically developing readers) of a factor of about 2. For example, using a standard reading test (e.g., MT Reading test, Cornoldi and Colpo 1995, 1998), children with dyslexia read a text at 0.58 s/syllable, while proficient readers read it at 0.25 s/syllable (Spinelli et al. 2005), that is, with a 132 % delay. Similar delays were obtained in later studies (De Luca et al. 2008: 0.45 vs. 0.25; 80 % delay; De Luca et al. 2010: 0.39 vs. 0.22; 77 % delay; Paizi et al. 2011: 0.38 vs. 0.21; 81 % delay; Zoccolotti et al. 2008: 0.39 vs. 0.22; 77 % delay Zoccolotti et al. 2012: 0.51 vs. 0.23; 117 % delay), averaging a reading delay of 97 % in children with dyslexia compared with typically developing children.

On the other hand, investigations into single-word reading (or lexical decision) are not usually aimed at obtaining an overall estimate of reading delay; they usually focus on measuring the effect of psycholinguistic variables (such as word frequency, lexicality, and length) on RTs and examine group differences as a function of such variables. Brinley plots can be used to obtain an overall estimate of the severity of the reading delay across experimental conditions. They report the RT means of children with dyslexia for all experimental conditions as a function of those of typically developing readers. It has been proposed that the slope (i.e., the *beta* coefficient) of the linear relationship between the means of two groups is a reliable overall estimate of the speed deficit and should be closely related to performance on standard tests (Kail and Salthouse 1994). We used this procedure with words and non-words of various lengths and obtained a beta equal to 4.0 in Italian children with dyslexia (Zoccolotti et al. 2008). This indicates that over and above the effect of specific experimental conditions children with dyslexia show a marked, multiplicative (or over-additive) RT delay across stimulus materials and conditions (also called a global factor) in dealing with orthographic materials.

In previous work, we proposed that this global factor indicates the difficulty of children with dyslexia in constructing a graphemic representation (or "graphemic description" in the terminology of Marsh and Hillis 2005). According to this view, children with dyslexia are selectively impaired when they have to operate on a string of graphemes presented visually, and this slowness cascades on all subsequent processing and amplifies group differences as a function of general condition difficulty. Consistent with Kail and Salthouse's (1994) suggestion, the *beta* derived from RTs to single-item presentation in laboratory experiments was related to reading times in a functional context (Zoccolotti et al. 2008). This supports the view that the beta captures the severity of the speed deficit in children with dyslexia.

However, a beta equal to 4 indicates a four times factor in the reading speed deficit, whereas, as summarized above, the reading delay was estimated as a factor of about 2 based on the MT standard reading test (Cornoldi and Colpo 1995, 1998). There is a mismatch between these two estimates; however, a direct comparison between them is impossible. In fact, RTs exclude the actual time taken to utter the target (here to after "pronunciation time"), which instead is included in functional measures. Although it is generally assumed that pronunciation times do not contribute to the group difference, to the best of our knowledge, this assumption has never been directly tested. A notable exception is a recent study by Davies et al. (2012) who examined reaction times and response duration in Spanish children with and without a reading deficit. Perhaps there is a lack of data because measuring pronunciation time requires time-consuming trial-by-trial analysis, whereas various software packages allow automatic evaluation of RTs. Thus, in most cases, total reading time includes a component of the response (pronunciation) that has been little studied and modeled. On the other hand, to explain group differences based on RT estimates models typically exclude the execution component from the contributing factors.

The "*difference engine model*" (DEM), proposed by Myerson et al. (2003), explains group RT differences by assuming that in the absence of a peripheral deficit, most differences between individuals are due to the amount of cognitive processing required. The DEM tries to account for individual variability in RTs to timed tasks by assuming there are two different (and independent) portions of the response (referred to as "compartments"). One compartment identifies the cognitive time differentiating two groups of individuals with a global difference in processing efficiency across different tasks (e.g., young adults *versus* elderly in Myerson et al.'s study; in the present work, this statement refers to typically developing readers compared to children with dyslexia). Here, we assessed the cognitive compartment by manipulating task difficulty through item length and lexicality. A different sensory–motor compartment identifies the portion of time required for peripheral analysis (sensory perception) and programming/beginning the motor response. This latter portion of the response is expected to be unrelated to the cognitive compartment and not different in groups varying for a global speed factor. Note that the sensory–motor compartment of RT estimates the time needed for sensory processing, motor planning, and voice onset; it excludes the execution component and is largely unaffected by manipulation of task difficulty along cognitive dimensions. For this reason, the DEM predictions cannot be extended to directly account for variability in total reading times. According to this model, linearity is expected in the plots contrasting condition means with their standard deviations; the slope of the linear relationship (*beta*) indicates the correlation between the cognitive stages involved in the task, and the intercept on the X-axis estimates the time of the sensory–motor compartment. The same *beta* for groups with different processing speeds indicates that the cognitive compartment is described by the same general factor; furthermore, the same x-intercept indicates that the two groups are comparable for sensory–motor compartment. Some evidence indicates that this pattern holds when older and younger adults are compared on a variety of timed tasks (Zheng et al. 2000). These predictions of DEM have not yet been tested in comparisons of children with dyslexia and typically developing readers.

Notably, the DEM has been applied to RT experiments where only the *incipit* of the response (provided by a key-press) is considered; by contrast, the actual time taken to carry out a given task (in the present case, reading aloud the target word or non-word) is ignored. In this vein, a difference between groups in the peripheral compartment does not indicate a delay in execution time and vice versa. An elderly person may be ten times slower than a young one in running 100 m but may be only a few seconds slower in processing the acoustic signal, planning the start, and raising their foot from the starting blocks. Indeed, the model assumes that the motor/executive component is small and constant across tasks. However, in measures of functional reading, which include word pronunciation, response times are clearly not invariant across stimuli. For example, one would expect pronunciation times to be longer for longer words; these aspects may also contribute to individual and group differences. Therefore, comparing groups for functional reading may prove somewhat more complex than the ideal situation envisaged for RTs by Myerson et al. (2003).

In the present study, we aimed to separately measure the components that contribute to reading times, that is, decoding and pronunciation, and evaluate whether (and to what extent) pronunciation times contribute to group differences. In particular, two models were used to investigate different but complementary questions. The DEM (Myerson et al. 2003) was used to characterize the nature of the global factor by disentangling the two compartments (cognitive and sensory–motor) that affect RT measures. As for the DEM, a second aim was to evaluate whether the model's assumption (i.e., no group difference in sensory–motor compartments) is appropriate in this context or whether the sensory–motor compartment has a role in generating group differences between children with dyslexia and controls. The rate and amount model (RAM; Faust et al. 1999) was used to identify the presence of specific effects of experimental variables (i.e., lexicality and length) after controlling for over-additivity.

Furthermore, we used the DEM applied to RTs to predict individual differences in total reading times by considering pronunciation times and the influence of the experimental manipulations on them. Our working hypothesis was that the DEM could be adjusted to accommodate the more complex (but closer to functional reading) case in which reading times (and not simply RTs) are predicted. Overall, we expected that the present analysis would contribute to clarifying the different size of the speed deficit shown by children with dyslexia when measured with vocal RTs or text reading.

Method

Participants

Participants included 43 typically developing readers and 25 children with dyslexia. Groups were comparable for age and gender (see Table 1). The children with dyslexia scored at least 1.65 SD below the norm for either reading time or accuracy on a standardized reading test (MT Reading test, Cornoldi and Colpo 1995). In this test, the child

Table 1 Summary statistics for the two groups of participants: mean age (in years, with range in parentheses); N, number of female and male participants; mean raw scores (and SDs in parentheses) on Raven's Coloured Matrices (these data were available for 12 children with dyslexia and all typically developing readers), WISC total scores (these data were available for 13 children with dyslexia), mean raw scores and *z*-scores (and SDs in parentheses) on reading time and accuracy measures of the MT Reading test

	Age	Female	Male	Raven test	WISC	Reading time (sec/syll)	Reading accuracy (errors)	Reading speed (*z*-score)	Reading accuracy (*z*-score)
Children with dyslexia (*N* = 25)	11.8 (10.5–13.3)	11	14	27.3 (2.6)	96.2 (10.1)	0.23 (.04)	21.7 (9.1)	−2.8 (1.9)	−3.0 (1.8)
Typically developing readers (*N* = 43)	11.6 (11.1–13.4)	23	20	28.8 (3.4)	–	0.51 (.17)	6.2 (3.6)	0.4 (0.4)	−0.1 (0.7)

reads a text passage aloud with a 4-min time limit; reading time (s/syllable) and accuracy (number of errors, adjusted for the amount of text read) are scored (data are presented in Table 1 as absolute and normalized values). Considering the reading time raw data, the average reading time of the two groups was 0.23 (typically developing readers) and 0.51 s/syllable (children with dyslexia), that is, a slowing of 122 %, which is close to the factor of 2 indicated in the introduction. To assess non-verbal IQ levels, we used the scores obtained by the typically developing readers and 12 children with dyslexia on Raven's Coloured Progressive Matrices. All children scored well within the normal limits according to Italian norms (Pruneti et al. 1996). WISC data were available for the other 12 children with dyslexia; scores were well within the normal range for both performance and verbal sub-scales. All participants had normal or corrected-to-normal visual acuity.

Stimuli, apparatus, and procedure

Stimuli were 80 words and 80 non-words. We selected 4-, 5-, 6-, and 7-letter words from the LEXVAR database (Barca et al. 2002; http://www.istc.cnr.it/grouppage/lexvar). There were 20 words per length condition, matched for frequency and initial phoneme across all lengths (1.4 mean log frequency). Non-words were derived from words by changing one (or two) letter (s).

Stimuli appeared in black lowercase Times New Roman on a white background; at the viewing distance of 57 cm, mean center-to-center letter distance was 0.4° horizontally. The stimuli were displayed singly on a PC screen controlled by DMDX software (Forster and Forster 2003). Each trial sequence consisted of a 15-ms acoustic tone, a 400-ms blank, and a 250-ms fixation cross followed by presentation of the target. The stimulus disappeared at the onset of pronunciation or after 4,000 ms. The 160 stimuli appeared in 8 blocks of 20 trials each (4 words and 4 non-word blocks) separately for the different lengths. Stimuli appeared in pseudo-randomized fixed order within each block. The order of type of stimulus (words, non-words) and length was counterbalanced across participants by keeping the four word blocks separate from the four non-word blocks. The first block of words (and non-words) was preceded by a brief practice block (five 6-letter items that were different from the experimental stimuli) and followed by a short pause. The children were instructed to read the word (or non-word) aloud as fast and accurately as possible. A voice key connected to the computer recorded vocal reaction times at the onset of pronunciation. The whole vocal response was also digitally recorded from onset.

Participants were tested individually in a quiet room at their school. The experimenter noted pronunciation errors. Only RTs, pronunciation times, and total reading times on correctly responded items were considered for the analyses.

Data analysis

Onset and offset of the vocal responses were manually detected by means of Check Vocal software (Protopapas 2007). *Vocal RT* was the time between the stimulus onset and the onset of the vocal response; *pronunciation time* was the time between the vocal onset and the end of the child's utterance; *total reading time* was defined as the time between the stimulus onset and the end of the child's utterance. Therefore, vocal RT and pronunciation time add up to constitute the reading time measure. Invalid trials due to technical failures accounted for 1.2 and 2.2 % of the responses of typically developing readers and children with dyslexia, respectively, and were discarded from the analyses. The error rate was 6.0 % for typically developing readers and 19.0 % for children with dyslexia.

To detect global components accounting for group differences, we tested the prediction of a linear relationship between the condition means of the two groups (children with dyslexia and typically developing readers) that varied in overall information-processing rate (the so-called Brinley plot). This relationship is diagnostic of the presence of an over-additivity effect, that is, the tendency to obtain greater group differences in the case of more difficult conditions over and above the effect of specific experimental manipulations (Faust et al. 1999; Myerson et al. 2003). Over and above such "large-scale" effects, selective contributions of specific variables can also be envisaged. According to the RAM (Faust et al. 1999), it is possible (by means of appropriate data transformations) to separate the contribution of global components (accounting for the presence of over-additivity) and specific components contributing to the group difference. Note that the RAM assumes there is a general linear structure in the expected values of response latencies that can be tested by looking at the Brinley functions and the linear relation between standard deviations and means across conditions. Other approaches, such as the state-trace analysis (Bamber 1979), which can be broadly applied also to closed-scale measures, or the diffusion model, which is commonly applied to binary answers such as lexical decision data (e.g., Zeguers et al. 2011), avoid many of these assumptions but are better suited for ad hoc experimental designs (Prince et al. 2012). To examine selective components in the group differences, Faust et al. (1999) suggest examining the effects of specific experimental manipulations by comparing parametric analyses (such as ANOVAs) on raw versus z-transformed data. Interactions that are significant in both the raw score and z-transformed score analysis indicate the selective influence of a given parameter; by contrast, interactions that are significant only in the raw data, and not in the z-transformed values, indicate that the interaction observed for raw data is spurious, that is, entirely explained by over-additivity.[1] Therefore, raw data were transformed into *z*-scores by taking each individual's condition means, subtracting their overall mean and dividing it by the standard deviation of their condition means. *Z*-scores indicate an individual participant's performance in a given condition relative to all other conditions based on the individual means of all conditions (therefore, each individual has an average of 0 across conditions and an SD = 1). This transformation re-scales individual performance to a common reference; hence, it allows controlling for global components while it preserves the information regarding individual variability across experimental conditions. In general, this transformation is widely applied in psychometrics to solve the problem of an idiosyncratic use of the scale, namely *response style* (Fischer 2004). This type of transformation effectively minimizes the response bias when it is performed on several measures drawn from equal variance distributions (Bartram 1996). In the RAM model, *z*-score transformation would be a standardized scale of amount of information processing, under the assumption of the absence of intrinsic variability in individuals in a specific condition. In this case, mean response latencies would vary for individual differences in processing rate and condition differences in difficulty. We carried out the *z*-score transformation separately for reading times and vocal RTs.[2]

Separate ANOVAs were carried out on raw and z-transformed data to examine the effect of group (unrepeated measure), length and lexicality (repeated measures) on reading times and vocal RTs, respectively. Only an ANOVA on raw data was carried out on pronunciation times as no global factor was detected for these measures (see below). Finally, an ANOVA was carried out on percentage of errors to evaluate the effect of group, length, and lexicality.

Last, we tested the DEM (Myerson et al. 2003) prediction of a linear relationship between the overall group means in reaction times and the standard deviations of the same conditions. As summarized in the introduction, according to the DEM analysis, this type of plot is critical to detect a global factor. The DEM makes a number of predictions in this regard (see "Results"). Note that the model assumes the sensory/motor component to be small and constant across experimental conditions and individuals. Accordingly, the DEM can be applied directly to vocal RTs. Here, we extended the predictions of the model to predict individual variability in reading times (in which pronunciation time is included and contributes variable amounts of time depending on the experimental manipulations) and compared these predictions with the data; details of these computations are given below.

Results

The plots in Fig. 1 present three different measures of reading: vocal RTs (a), pronunciation times (b), and total reading times (c). Inspection of the figures allows for a number of general observations:

- Group differences were large in the case of *vocal RTs* and *total reading times* but much smaller in the case of *pronunciation times*;
- Unlike *vocal RTs* and *total reading times*, *pronunciation times* generated very small individual differences (i.e., small SDs);

Clear effects of length were present in both groups but were more selective for typically developing readers (the slopes of the linear regressions of letter length on the different measures, in ms per letter, are reported in Table 2). Typically developing readers showed a moderate effect of length in pronunciation (for both words and non-words), a moderate effect in vocal RTs to non-words, and a negligible effect for words. Children with dyslexia showed a length influence on pronunciation times (similar to controls) and a much larger effect on vocal RTs.

Brinley plots

In Fig. 2, data were re-plotted as Brinley plots, that is, the means obtained in each condition by the group of children with dyslexia were plotted as a function of the means of the control children in the same conditions. Data for vocal RTs, pronunciation times, and reading times are presented in plots 2a, 2b, and 2c, respectively.

Group differences in *vocal RTs* increased as a function of condition difficulty and were well accounted for ($R^2 = 0.88$) by a single regression line with beta equal to 3.44. This pattern indicates an over-additivity effect, that is, children with dyslexia were progressively more impaired with respect to typically developing readers in more difficult conditions over and above the influence of the specific experimental manipulations. The slope of the linear relationship represents an estimate of the severity of the

[1] Although not explicitly predicted by the RAM (Faust et al. 1999), the opposite pattern may also occur and, indeed, it did occur in the present study; *i.e.*, an interaction may be significant in the z-transformed values analysis but not in the raw data analysis. Apparently, controlling for the influence of global components can enhance the sensitivity of statistical comparisons and allow detecting differences that are masked in the raw data analysis.

[2] Note that these transformations may be applied to open scales, such as time, but they are inappropriate in the case of closed scales, such as accuracy.

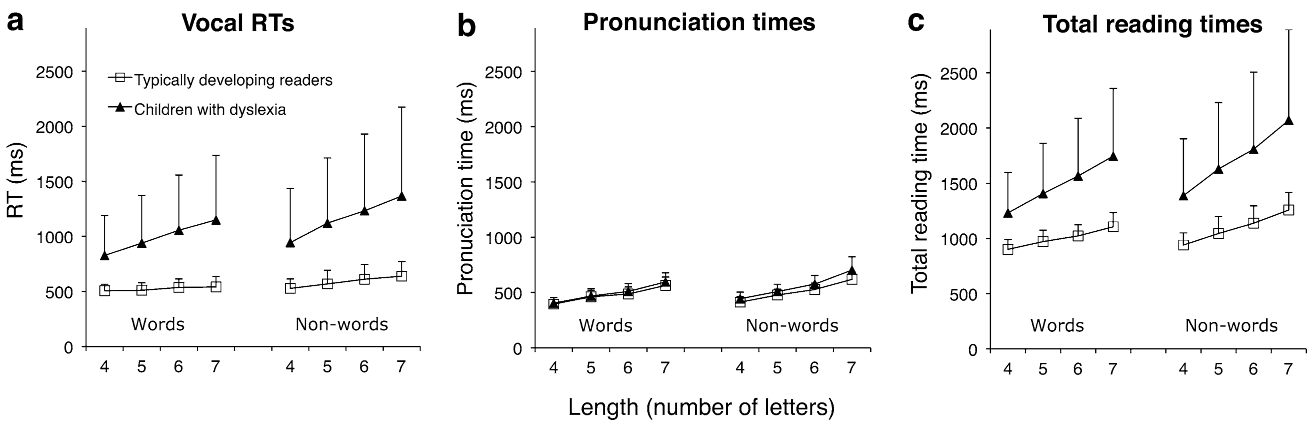


Fig. 1 Means and standard deviations of vocal reaction times (**a**), pronunciation times (**b**), and total reading times (**c**) as a function of stimulus length for typically developing readers (*open squares*) and children with dyslexia (*filled triangles*). The slopes (*bs*) and the R^2 of the linear regressions for the various conditions (and groups) are presented in Table 2

Table 2 Dependence of processing time on stimulus length: values indicate ms per letter and correspond to the slope (b) of the linear regression analyses

	Typically developing readers		Children with dyslexia	
	Words	Non-words	Words	Non-words
Vocal RTs	13 [452] (0.91)	38 [380] (0.99)	109 [396] (0.998)	138 [406] (0.99)
Pronunciation times	54 [182] (0.97)	67 [141] (0.99)	62 [154] (0.98)	84 [93] (0.97)
Total reading times	67 [635] (0.99)	105 [521] (0.998)	170 [550] (0.999)	223 [499] (0.995)

The intercept on the ordinate axes is reported between square brackets, and the R^2 of the linear regressions in parentheses

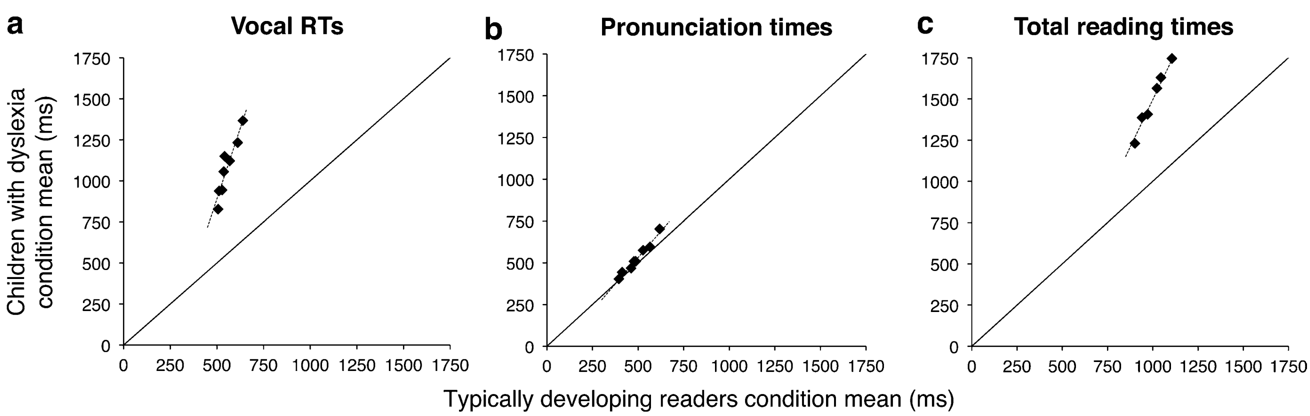


Fig. 2 Brinley plots: condition means of children with dyslexia are plotted against the respective conditions of control children. The equality line is presented in *solid gray*; the regression line appears as *black dots*. The slope of the linear regression was 3.44 for vocal RTs ($R^2 = 0.88$), 1.26 for pronunciation times ($R^2 = 0.97$), and 2.3 for total reading times ($R^2 = 0.99$)

impairment; accordingly, children with dyslexia were ca. 3–4 times slower than typically developing readers. It must be noted that although the R^2 is high it is not equal to 1, leaving room for biases in the *z*-score analysis based on RAM (see below).

The data on *pronunciation times* (Fig. 2b) had a good fit ($R^2 = .97$). The two groups performed similarly; the over-additivity effect was present but very small, and the beta of the regression was close to unity (1.26). The 99 % confidence interval of the regression line was 1.07–1.45; as the lower bound of the interval was above unity, this indicates a small but detectable difference between groups. At any rate, interpolation of a line with a slope of 1 reduced the goodness of the fit only minimally ($R^2 = .94$).

Table 3 Results of the ANOVAs on vocal RTs, pronunciation times, and total reading times

	df	Vocal RTs			Pronunciation times			Reading times		
		MSe	*F*	*p*	MSe	*F*	*p*	MSe	*F*	*p*
Raw data										
Group	1,66	859,340	40.4	.0000	27,846	5.0	.028	941,426	41.7	.0000
Lexicality	1,66	43,618	40.3	.0000	2,785	103.8	.0000	44,370	78.3	.0000
Length	3,198	25,020	46.7	.0000	1,209	790.9	.0000	24,393	172.2	.0000
Group × *lexicality*	1,66	43,618	8.8	.0042	2,785	11.3	.0013	44,370	14.3	.0003
Group × *length*	3,198	25,020	20.4	.0000	1,209	8.4	.0000	24,393	26.7	.0000
Lexicality × *length*	3,198	11,680	3.6	.014	619	31.4	.0000	11,207	9.7	.0000
Group × *lexicality* × *length*	3,198	11,680	.14	–	619	2.7	.049	11,207	.42	–
Z-transformed data										
Group	1,66	0	.00	–				0	.00	–
Lexicality	1,66	.64	125.4	.0000				.30	209.6	.0000
Length	3,198	.38	133.6	.0000				.12	804.7	.0000
Group × *lexicality*	1,66	.64	8.8	.0041				.30	.3	–
Group × *length*	3,198	.38	2.4	.068				.12	.9	–
Lexicality × *length*	3,198	.37	9.7	.0000				.14	17.6	.0000
Group × *lexicality* × *length*	3,198	.37	5.9	.0007				.14	2.8	.043

Analyses are presented separately for raw and z-transformed data

Group differences in *total reading times* increased as a function of condition difficulty and were well accounted for ($R^2 = 0.99$) by a regression line with a beta of 2.30. Note that the slope of the linear relationship was smaller than that of vocal RTs.

The effects of group, lexicality, and length were analyzed with ANOVAs on raw and z-transformed data. This transformation enables exploring the effects of the experimental manipulations by controlling for the global factor that relates the means of the two groups (i.e., controlling for the over-additivity effect). A synthesis of these analyses is presented in Table 3 (Please refer to the table for the significance of main effects and interactions); note that analyses on z-transformed values were carried out for total reading times and vocal RTs, but not pronunciation times because in the latter case, the value of the global factor relating means of dyslexic and typically developing readers was close to unity (see below).

As for *vocal RTs*, the effect of group was significant: children with dyslexia were much slower (1,079 ms) than typically developing readers (555 ms), that is, there was a 524-ms difference. The main effect of the group factor is inherently nil in the *z*-score analyses because of the data transformation. Vocal RTs were longer for non-words (876 ms) than words (758 ms; that is, lexicality effect) and they increased steadily as a function of the number of letters in the stimulus (length effect). The lexicality and length factors interacted so that a larger length effect was present for non-words (all comparisons were significant: *p*s at least <.001, HSD Tukey's test) compared with words (5 out of 6 comparisons were significant: *p*s at least <.05). Although the group by lexicality interaction was significant in both raw and *z*-score analyses (see Fig. 3a, b), it indicated effects in different directions. In the raw data analysis, the difference between words and non-words was greater in children with dyslexia than in typically developing readers (diff 173 ms, $p < .0005$, and 63 ms, $p < .05$, respectively). When the global factor was controlled for by the *z*-score transformation, however, this difference was actually smaller in children with dyslexia than in typically developing readers (both *p*s < .0005). The group by length interaction in the raw data analysis (see Fig. 1a) indicated that children with dyslexia were more affected by length (all *p*s at least < .01) than typically developing readers (all differences n.s.). In the *z*-score analysis, the interaction was marginally significant; however, in this case, the difference in the length effect in the two groups emerged most clearly as a function of lexicality. Thus, the group by lexicality by length interaction was significant (but this was not so in the raw data analysis). This interaction is presented in Fig. 3c: typically developing readers showed a much smaller effect of length ($p < 01$ for two comparisons, n.s. for the remaining four) than children with dyslexia (*p* at least <.01 for four comparisons, n.s. for the remaining two) in the case of word stimuli. By contrast, both groups were similarly affected in the case of non-words (most *p*s at least <.0001). However, there was also a slight tendency for controls to be more affected by

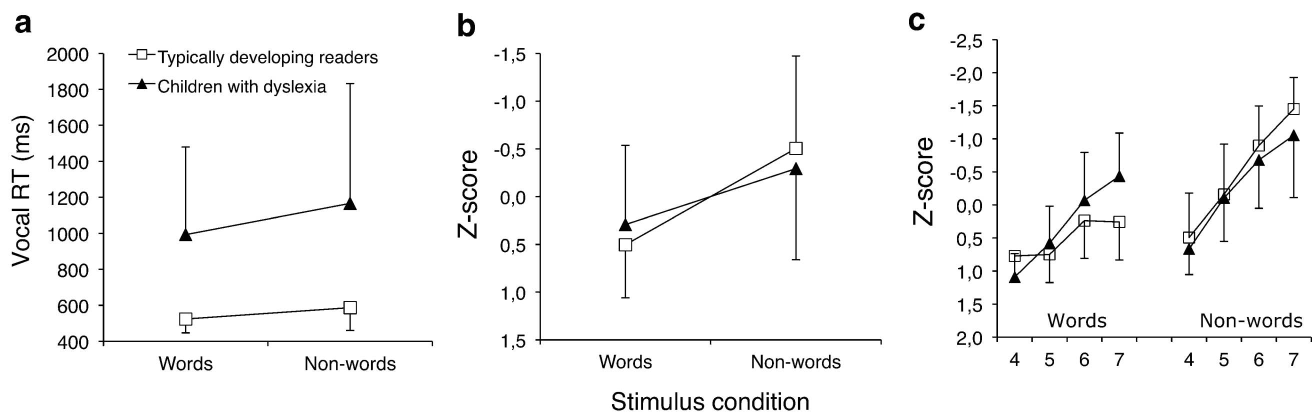


Fig. 3 The left plot **a** presents data for vocal RTs and the central plot **b** *z*-score data on word and non-word (lexicality effect) performance of children with dyslexia and typically developing readers. Note that children with dyslexia show a greater difference between words and non-words in the case of raw data and a smaller one in the case of *z*-score data. The rightmost plot **c** shows the effect of length for words and non-words in children with dyslexia and typically developing readers in *z*-score data (for the related means in raw vocal RTs please refer to the left plot (a) in Fig. 1)

length, which was limited to the longest non-words; this effect was difficult to interpret.

In the analysis of *pronunciation times*, the error terms (MSes) associated with the various main effects and interactions were considerably smaller (in fact, at least 15 or more times smaller) than those in the vocal RTs and reading times (see Table 3). Accordingly, the analysis of pronunciation times is quite sensitive and captures effects even though they are small in absolute terms. Thus, a significant main effect of group was detected (33 ms group difference). The main effect of length indicated slower pronunciation times for longer stimuli. The main effect of lexicality indicated slower pronunciation times for non-words (533 ms) compared with words (485 ms; diff 48 ms). The length and lexicality factors interacted, in that length effects were more marked in the case of non-words. The group by length interaction indicated that children with dyslexia were slower than typically developing readers for 6-letter (diff 36 ms; $p < .00005$, HSD Tukey's test) and 7-letter (diff 58 ms; $p < .00005$) stimuli but not for 4- or 5-letter stimuli. The group by lexicality interaction revealed a group difference for non-words (diff 49 ms; $p < .0005$) but not for words (18 ms, n.s.). As for the group by lexicality by length interaction, the two groups were different in the case of non-words and the size of the difference grew as a function of length (4-letter NWs = 32 ms; 5-letter NWs = 32 ms; 6-letter NWs = 48 ms; 7-letter NWs = 84 ms); in the case of words, the two groups were different only for 7-letter words (31 ms; $p < .005$).

The analysis of *reading times* revealed a large main effect of group: children with dyslexia (1,605 ms) were slower than typically developing readers (1,048 ms) by 557 ms. The main effects of length and lexicality were significant, indicating slower reading times for non-words and

Table 4 Results of the ANOVA on reading errors

	df	Reading errors		
		MSe	*F*	*p*
Group	1,66	21,590	89.2	.0000
Lexicality	1,66	15,107	161.4	.0000
Length	3,198	3,670	62.2	.0000
Group × *lexicality*	1,66	1,750	18.7	.0000
Group × *length*	3,198	273	4.6	.0037
Lexicality × *length*	3,198	1,941	40.7	.0000
Group × *lexicality* × *length*	3,198	78	1.6	–

for longer stimuli. These two factors interacted indicating larger effects of length in the case of non-words. The group by length and group by lexicality interactions were significant in the case of raw data but vanished in the *z*-score analysis, indicating that they can be accounted for in terms of over-additivity. The group by lexicality by length interaction was significant only in the *z*-score analysis, but was difficult to interpret. In fact, the interaction was entirely due to the tendency of the typically developing readers to be more affected by length in the longest non-words than the children with dyslexia.

Finally, an analysis was carried out on the *percentage of errors*. Results are summarized in Table 4. The analysis revealed a significant main effect of group (6 % controls, 19 % dyslexics), length (going from 5 % for the shortest to 19 % for the longest stimuli), and lexicality (7 % words, 18 % non-words). The length and lexicality factors interacted in that length effects were more marked in the case of non-words. The group by length interaction indicated that adding letters significantly increased the error rate for children with dyslexia (comparisons between 4 and 5, as

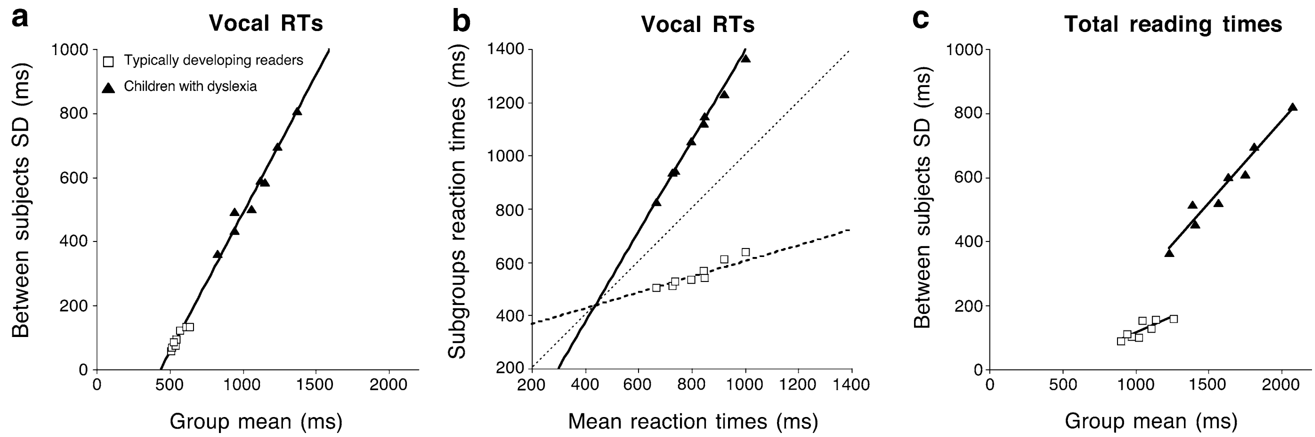


Fig. 4 **a** SDs of children with dyslexia and typically developing readers are plotted as a function of the related condition means in vocal RTs. **b** Vocal RT data for controls and dyslexics are fit with DEM; fitting the data from the two groups independently yields an $R^2 = 0.99$ and an $R^2 = 0.88$ for children with dyslexia and typically developing readers, respectively. **c** SDs of children with dyslexia and typically developing readers are plotted as a function of the related condition means in total reading times and are fit with two linear regressions, with a slope of 0.52 ($R^2 = 0.95$) for children with dyslexia, and a slope of 0.20 ($R^2 = 0.70$) for typically developing readers, respectively

well as 5 and 6 letters all *ps* < .00003, HSD Tukey's test) and only marginally increased errors in the case of controls (comparison between 5 and 6 letters $p < .01$; all other *ps* n.s.). The group by lexicality interaction revealed a greater increase in errors when reading non-words in the dyslexic group relative to the control group (all *ps* < 0.0001); the group by length by lexicality interaction was not significant.

Test of the difference engine model (DEM)

This model characterizes the nature of the global factor by separating the influence of different (cognitive and sensory–motor) compartments. Fig. 4a plots the means of vocal RTs against SDs for the two groups of children.

The variability grows linearly with increasing condition means for both typically developing readers and children with dyslexia. Note that a single line fits the two groups well. Thus, the solid line in Fig. 4a represents the DEM prediction calculated on all the observers using the following equation (Eq. 1):

$$\mathrm{SD} = \left(r - \frac{\sigma_c}{\alpha}\right)(\mathrm{RT} - t_e) \quad (1)$$

where σ_c, α, t_e and r are parameters that are free to vary and represent the variance and amplitude of the effects, the time required by the sensory–motor compartment, and the theoretical correlation between the cognitive stages, respectively. Note that it is assumed here that children with dyslexia and typically developing readers share a similarly efficient peripheral compartment. In the figure (Fig. 4a), the sensory–motor compartment is represented by the x-intercept of the regression line. The model explains the variance in the data ($R^2 = 0.99$) with an estimated time for the execution compartment of 438 ms and a slope of $m = \left(r - \frac{\sigma_c}{\alpha}\right)$ of 0.87.

This indicates that, although the condition means of the two groups show no overlap, the slope relating means and SDs allows placing the two groups on the same line. Thus, the deficit of children with dyslexia in dealing with orthographic materials is entirely captured by a general cognitive factor (indicated by the same slope of the relationship between SDs and mean RTs) with no group difference in the sensory–motor compartment (measured by the x-intercept).

Using the modeled relationship between means and SDs (the solid line in Fig. 4a), it is possible to predict the observed differences among groups in the condition means. Based on the parameters estimated by DEM, we predicted group differences with Eq. 2:

$$TR_i = (1 + z_i m)TR_{ave} - z_i m t_e, \quad (2)$$

where z_i is the mean RT in *z*-scores for the group *i*. The *z*-scores are calculated on the entire sample; thus, slow observers have positive scores.

Figure 4b presents the experimental data of the two groups (filled and open symbols for children with dyslexia and typically developing readers, respectively) and the predictions of DEM for the two groups (represented by solid and dashed lines, respectively). The DEM accurately predicts the differences in mean RTs with an R^2 of 0.99 and 0.83 for children with dyslexia and typically developing readers, respectively. Thus, consistent with the model prediction, RT group differences are explained well by

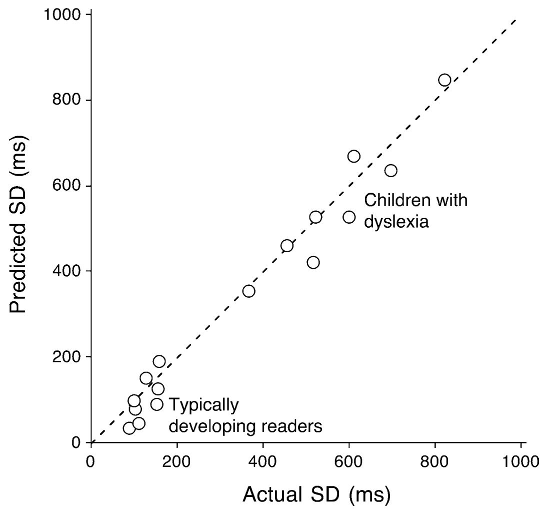


Fig. 5 DEM prediction of the individual differences in total reading times: predicted SDs as a function of the measured SDs (based on Eq. 3). The *dashed line* represents the equality line with a slope of 1 and an intercept of 0 (i.e., identity between the measured and predicted SDs); this line explains the most of the variance ($R^2 = 0.99$)

combining a single global cognitive slowing factor with the relationship between mean RTs and SDs.

Although we have considered only RT data up until now, Fig. 4c shows the relationship between SDs and total reading times. In this case, the DEM model cannot be directly applied because it does not include predictions for an execution compartment, such as pronunciation time that varies as a function of length and lexicality (as shown above). Indeed, Fig. 4c is clearly different from Fig. 4a, and the data of the two groups cannot be fit by a single line.

However, we used the model already applied to the RTs (Fig. 4a) to predict the individual differences expressed by the SDs in the total reading times (Fig. 5). The hypothesis is that the relation between SDs and total reading times (which appears different in the two groups) can be fully accounted for by the same cognitive and motor compartments estimated in the RT measures plus an execution component that varied with task difficulty (i.e., pronunciation time in reading). We assumed that the SDs of this execution component would not increase as a function of condition mean difficulty. Failure in predicting SDs in total reading times would indicate that, rather than the general cognitive factor, other components, possibly related to the time taken to utter the words, would contribute to the group differences in total reading times.

In total reading time, the execution component represented by the pronunciation times varies as a function of the experimental manipulation used (i.e., the number of letters in the string, see Fig. 1b), as does the decoding component measured by RTs. Thus, to compute the prediction of total reading time, we considered the two components (RTs and pronunciation times) of total reading times by analyzing the linear relationship between pronunciation times and RTs. The regression lines of RTs on pronunciation time (not reported in the figure) were calculated separately for the two groups and yielded an R^2 of 0.93 (slope 0.53, intercept −42 ms) and 0.62 (slope 1.24, intercept −195 ms) for children with dyslexia and typically developing readers, respectively. The difference in the slopes obtained by the two groups is due to the non-homogeneous range of RTs between the two groups (around 500 ms for typically developing readers and between 800 and 1,300 ms for children with dyslexia; see Fig. 1a) regardless of a similar range of pronunciation times (Fig. 1b). Taken together, the parameters obtained from the DEM (Eq. 1) and the relationship between RTs and pronunciation times allow predicting the SDs in total reading times according to the following equation:

$$\mathrm{SD} = m\left(\frac{T_p - a_i}{b_i} - t_e\right), \quad (3)$$

where m and t_e are derived from the model (Eq. 1), and a and b are the intercept and the slope of the regression line of task difficulty on pronunciation times (T_p) of the group i.

Figure 5 presents the predicted SDs as a function of the measured SDs of total reading times. The dashed line represents the equality line with a slope of 1 and an intercept of 0 (i.e., identity between the measured and the predicted SDs); this line explains most of the variance ($R^2 = 0.99$). Consistent with the hypothesis proposed above, although RTs and total reading times have very different patterns (Fig. 4a, c), the same cognitive factor predicts a very large proportion of variance in both measures.

Discussion

In clinical settings, reading speed is calculated by measuring reading times for lists of words or text passages, and in laboratory settings, vocal RTs are often used. As indicated in the introduction, the critical difference between these two measures is the contribution of pronunciation times (absent in the vocal RTs). Results showed that the slope of the Brinley plots relating the means of the two groups (typically developing readers and children with dyslexia) was less steep for reading times (2.3, indicating a 130 % delay) than for vocal RTs (3.4, indicating a 240 % delay). Thus, the contribution of pronunciation times reduced the difference between groups. The rationale for this change is that the overall reading times are due to both decoding time, measured by RTs, which markedly distinguishes between the two groups, and pronunciation time, which shows only a very small (though significant) group difference. Overall, we concluded that the deficit in reading single words is less severe if reading times (rather than RTs) are considered. Below we detail the pattern of findings for the different measures considered (vocal RTs, pronunciation times, and reading times).

Vocal RTs

The data closely confirmed the predictions of both the rate and amount model (Faust et al. 1999) and the difference engine model (Myerson et al. 2003) that a single global factor accounts for a large proportion of the group difference between children with dyslexia and typically developing readers. As predicted by the DEM, the variability of the two groups was well fit by a single regression line. The x-intercept of the regression measures the time spent in early perceptual processing and motor planning up to voice onset, which represent the sensory–motor compartment of the task. In keeping with the model, children with dyslexia and typically developing children were not different with regard to this processing. As to the cognitive compartment, the children with dyslexia were severely impaired (with a slope *beta* of 3.4) across all task conditions. Therefore, vocal RTs are particularly sensitive in capturing the reading deficit. This sensitivity derives from a) the possibility of isolating the decoding component from the execution time and b) the possibility (applying models such as the DEM) of distinguishing the cognitive part of the response, which identifies the largest group difference from the portion of the task performed equivalently by the two groups (and from the actual execution time).

The presence of a single global factor indicates that more difficult conditions produce *proportionally greater* group differences (over-additivity effect). This pattern (as well as the size of the multiplicative factor) confirms previous evidence obtained from both reading and lexical decision tasks (e.g., Di Filippo et al. 2006; Paizi et al. 2013). Previous studies have shown that this factor was not present in the case of very short orthographic stimuli (one- or two-letter stimuli; De Luca et al. 2010), pictures (Zoccolotti et al. 2008) or when stimuli are presented in the acoustic modality (Marinelli et al. 2011). Thus, the slowness of children with dyslexia a) regards graphemic processing of strings of letters (either words or non-words) and b) is proportionally amplified by the difficulty of letter string decoding. Based on this pattern of findings, we have proposed that the global factor marks a deficit in pre-lexical graphemic processing (Zoccolotti et al. 2008). Notably, this is a robust effect and accounts for a large proportion of the difference between typically developing readers and children with dyslexia. Alternative views have also been proposed. For example, van den Broeck and Geudens (2012) discussed the possibility that the slowing of dyslexic children is due to a deficit in orthographic–phonological binding. Evidence in this direction comes from imaging studies indicating close associations between letters and speech sounds early in development (for a review see Blomert 2011); accordingly, effective letter–speech sound integration is necessary for the development of reliable letter recognition. Thus, a deficit in orthographic–phonological binding may represent a proximal cause of the reading slowness in dyslexia (Blomert 2011). Additional research seems necessary to distinguish between these two different options.

Taking into account group differences in global graphemic processing (i.e., once over-additivity effects were controlled by following the indications of the RAM, Faust et al. 1999), also specific small effects of length and lexicality were detected in the vocal RT analyses. In particular, children with dyslexia were sensitive to length in the case of both words and non-words; by contrast, typically developing readers showed a length effect for non-words and a much more limited influence of lexical stimuli. This pattern of findings confirms previous evidence obtained from reading (Paizi et al. 2011) and lexical decision (Di Filippo et al. 2006) tasks. In general, a deficit in decoding long stimuli marks the difficulty of children with dyslexia, and this effect has been found to cut across regular and irregular orthographies (for a comparison between German and English children with dyslexia, see Ziegler et al. 2003).

With regard to lexicality, opposite results emerged from raw and *z*-score analyses. In the former case, the children with dyslexia showed a *larger* lexicality effect than typically developing readers. In the latter case, the children with dyslexia showed a *smaller* lexicality effect; in other words, children with dyslexia were selectively impaired in reading words compared with non-words. Note that the detection of a selective deficit in reading words compared with non-words is usually masked by the generally greater difficulty of non-lexical stimuli and the resulting amplification of the group difference related to general task difficulty; therefore, to detect such an effect, it is necessary to control for over-additivity. Notably, this finding confirms previous evidence obtained with similar stimulus materials (De Luca et al. 2010); we proposed that in terms of lexicality, the pattern of results might be explained on the basis of two partially independent deficits: one in the early stage of graphemic analysis and one in forming (or accessing) lexical entries (De Luca et al. 2010). The deficit in reading non-words would be accounted for by the deficit in graphemic analysis alone. The greater deficit in reading words would be explained by both a deficit in graphemic analysis and a deficit in the expansion of the orthographic lexicon. There is already evidence that Italian children with dyslexia have a limited orthographic lexicon. In particular, it was found that these children show delayed RTs in reading low-frequency words over and above their general slowness in dealing with orthographic materials (Paizi et al. 2013).

Furthermore, the frequency effect interacted with context: typically developing children showed a list context effect for high- and low-frequency words, whereas children with dyslexia showed a list context effect only for high-frequency words (Paizi et al. 2011).

Similar results, based on the state-trace approach (STA), were recently presented by Van den Broeck and Geudens (2012). Unlike RAM, this model does not require that assumptions be made about the general linear structure on the expected values of the measures. The authors designed an experiment based on STA by functionally matching the stimuli on the state factor (lexicality) so that the selected items in the easy dimension (words) varied in length between groups and could be read at the same speed by normal and disabled readers. Then, a non-word was produced for each item. A comparison of the effect of lexicality revealed the absence of a selective deficit for non-word reading and the presence of word reading disability in children with dyslexia, which was evident after controlling for letter decoding. The authors interpreted these findings within the "decoding stagnation hypothesis" (see also Van den Broeck et al. 2010): inexperienced readers (and children with dyslexia) rely more on phonological codes as they lack word-specific knowledge in comparison with typically developing readers. As a consequence, they show a smaller difference between words and non-words than typical developing readers because these latter have a larger word-specific knowledge base (Van den Broeck et al. 2010). Notably, the decoding stagnation hypothesis is also able to account for the difference in the length effect between dyslexic and typically developing children described above. Accordingly, the large length effect in children with dyslexia would express their greater tendency to rely on smaller grapheme–phoneme units (and inability to activate whole word-specific knowledge).

One might wonder whether the differential effects of length and lexicality in the two groups can be explained on the basis of a trade-off between RT accuracy. In fact, if children with dyslexia make more errors in the more difficult conditions (such as non-word or longer targets), group means in these conditions might be predominantly due to the easier items and RTs might indicate a smaller than real deficit. Consistently, the group by lexicality interaction in the error analysis indicated that, as expected, the children with dyslexia were more affected in reading non-words than controls. In the case of length, however, the group by length interaction indicated that children with dyslexia made more errors on long items. Therefore, if the group means on the longer items were based prevalently on the easier items, one would expect the effect of length on RTs to be smaller (rather than larger, as actually found) in children with dyslexia than in typically developing children. Thus, it seems unlikely that the weaker effect of lexicality and the marked effect of length on RTs in children with dyslexia relative to controls are due to easier items being sampled to estimate RTs in the former group.

Analyses based on the predictions of the DEM allowed separating the cognitive and sensory–motor compartments. Thus, the present results formally demonstrate that the group difference in reading tasks is specific for the cognitive analysis and that the two groups of children are equivalent in the early sensory and motor planning processes (sensory–motor compartment). A similar conclusion was reached in a lexical decision study by Zeguers et al. (2011) who, based on a diffusion model analyses, observed no difference between dyslexic and control readers in the non-decision components of the reaction times. Research on dyslexia has a long tradition of hypotheses, including several in which early sensory processes have been envisaged as the underlying cause of the disturbance. In this context, the magnocellular visual theory of dyslexia was one influential hypothesis (Stein and Walsh 1997). General slowness, regardless of the task, has also been found in children with dyslexia (Breznitz and Meyler 2003; Nicolson and Fawcett 1994; but for contrasting results see Bonifacci and Snowling 2008). The present results are not in keeping with these approaches. Rather, they are consistent with the idea that the delay of children with dyslexia is entirely linked to the central processes associated with reading.

Pronunciation times

As might be expected, longer stimuli required more time to be uttered. Further, a clear lexicality effect was detected; that is, non-words had longer pronunciation times than words. Pronunciation times have been largely neglected in previous research in reading. To the best of our knowledge, they have been directly measured only in a recent study in Spanish children (Davies et al. 2012), which showed the effect of length and lexicality as well as word frequency (not examined here). Together the two studies indicate the factors that affect duration of pronunciation.

When interpreting these effects, it should be considered that specific classes of stimuli require more time to be uttered, that is, the effect itself has a motor origin. This interpretation fits well with the case in which long (compared with short) stimuli have to be uttered. In this vein, one might interpret slowness in uttering non-words (compared with words) as indicating reduced fluency when articulating stimuli that are orthographically legal but are not present in the lexicon. Indeed, reading times drop dramatically with age/reading experience (Italian: Zoccolotti et al. 2009; Spanish: Davies et al. 2007), and this developmental trajectory may well be associated with the development of selective expertise in fluently articulating highly practiced items.

An alternative view is that some influence of the variables affecting the decoding of the stimulus (e.g., word frequency) spill over from the decoding phase of processing to the execution phase. In this vein, Davis et al. (in press) interpreted their findings by proposing that "phonological coding processes may continue after response onset, that is, that the phonological specification for a word's pronunciation may not be fully prepared at response onset." Note that according to this view, RTs do not exactly measure the time necessary for decoding, but underestimate it; the size of the underestimation varies with the length and frequency of the words. In the same vein, pronunciation time overestimates the time necessary for uttering the target word. Evidence for a partial overlap between cognitive and motor components also comes from studies in which the role of cognitive factors was examined by keeping constant the response output. For example, Balota and Abrams (1995) reported that when the subject was requested to make the same arbitrary response in a lexical decision task, the duration of the utterance varied for stimuli of different word frequency. More generally, several recent models explored the interactions between decision and execution and suggested that examining the characteristics of movement (such as variation in direction, acceleration) might provide more information on the timing of decisions than recording only when it starts (e.g., Spivey 2007; Resulaj et al. 2009). At any rate, a clear separation between these two interpretations (purely motor vs. influence of cognitive processing on motor processing) seems premature on the basis of pronunciation times only. In particular, the possibility should be kept in mind that both interpretations apply to a varying degree depending on stimulus type and task instructions (e.g., emphasizing either speed or accuracy). Accordingly, in extending the DEM model to accommodate pronunciation times, we made no assumption relative to the independence of cognitive and sensory–motor compartments on one side and a "motor" compartment on the other. Thus, the presence of a relationship between RTs and pronunciation times as a function of the experimental manipulations (lexicality and length) may indicate either that the two measures capture partially overlapping stages or that the processes are different in nature (i.e., it takes longer to utter longer words and to decode more letters) but act similarly on the dependent measure.

Our purpose in examining pronunciation times was to understand their internal structure and contribution to total reading times. In this context, a first question concerns the structure of the relationship between condition means and standard deviations. Our data indicate that variability increased minimally in more difficult conditions in both groups. Furthermore, the slope of the Brinley plot relating the pronunciation times of the two groups was just above unity. Therefore, experimental manipulations did not produce a proportional increase in the difference between groups in the case of pronunciation times, that is, the small difference in pronunciation times between groups was not explained by the global factor previously observed for RTs.

A second aspect of the results concerns the extremely small inter-individual variability in pronunciation times. From a statistical standpoint, this makes the statistical analyses of this parameter considerably more powerful than those on vocal RTs. Indeed, although the pronunciation times of the two groups of children were different, the group difference was minimal. Furthermore, a small but selective influence of the experimental variables on the two groups of children was detected. Thus, children with dyslexia were slower in uttering non-words, an effect that grew with increasing stimulus length (but was always significant). In the case of words, the two groups were different only for the longest stimuli, and no difference was observed for 4- to 6-letter words. Therefore, the present data do not indicate general basic difficulty in the articulation of these children. Rather, children with dyslexia seem selectively slower when they have to assemble for pronunciation items that are not present in their lexicon. These results generally agree with the observations of Davies et al. (2012); in fact, these authors found no overall group difference in pronunciation times but found selective influences of length, frequency, and lexicality, and children with dyslexia were slower in naming long stimuli, low-frequency words, and non-words, respectively.

Interpreting this pattern of results presents the same difficulty as that mentioned above in distinguishing purely motor components from the residual influence of coding processes on execution times. In the first perspective, the present findings are consistent with the idea that children with dyslexia have difficulty in assembling the phonological output of stimuli they have had no previous experience with. Although several models have emphasized that phonological factors constitute the core deficit of the reading disturbance, it is not clear how these interpretations deal with the productive components of reading. In the second perspective (preferred by Davies et al. 2012), the greater influence of lexicality on the pronunciation times of dyslexic children indicates a carry over from their slowness in decoding to their slowness in motor execution. It should be recalled that children with dyslexia have slower vocal RTs to non-words than controls. We proposed above that this effect depends on the (large) difference in graphemic decoding; by contrast, the effect of lexicality per se goes in the opposite direction (when general differences in performance are controlled for) and children are with dyslexia more impaired on words than non-words. If the slowness in pronouncing non-words is seen as a residual effect of general decoding difficulties, it might express difficulty in graphemic decoding more than selective difficulty in phonological processing. In any case, further research is necessary to discover the origin of the slowness in pronouncing non-words shown by children with dyslexia.

Total reading times

Total reading time is the measure that most closely approximates normal, functional reading. As already pointed out in our discussion of the Brinley plots, the beta for reading times of the children with dyslexia was lower than for vocal RTs, thus predicting a less severe deficit when reading times, rather than RTs, were considered. The present data contribute to understanding and modeling reading times.

In general, the pattern of findings on reading times can be understood as reflecting the additive effects of the vocal RTs and pronunciation times. In this respect, note that the two groups were very different as to vocal RTs (on average, ca 500 ms), and much less different (33 ms) as to pronunciation times. Moreover, decoding and pronunciation times were about the same in the typically developing readers, but were very different in the children with dyslexia. Thus, the reading times of the children with dyslexia were determined much more by the decoding than the pronunciation time, whereas pronunciation and decoding weighed similarly on the reading times of typically developing readers. For example, consider the influence of length. Typically developing readers showed a clear effect of length in reading times for words, and this is explained almost entirely by the pronunciation component. By contrast, the large length effect shown by children with dyslexia was predominantly due to the decoding component and only slightly to the pronunciation component.

At variance with the predictions of DEM (Myerson et al. 2003), the slope and the x-intercept of the curve relating SDs to means of reading times was different for the two groups considered. This can be understood if we consider that the model is defined by the simplest case in which execution adds an invariant time constant. However, in reading (like many other real-life visual-motor tasks, which do not rely on a simple key-press response), execution times vary as a function of task complexity. By taking into account a third, motor, compartment, which varied as a function of task difficulty (i.e., length and lexicality) but was very similar in the two groups (i.e., pronunciation means), we were able to show that the DEM can be adjusted to accommodate the more complex case at hand, at least for the predicted individual variability. In fact, the latter was quite close to the actual variability measured (see Fig. 5). The accuracy of the prediction-based global cognitive factor is understandable if the difference in the magnitude of the estimated reading deficit among RTs and total reading times is taken into account. Thus, decoding is the major source of the group differences, but they are attenuated in total reading estimates due to the contribution of a component (pronunciation) that is nearly invariant among groups.

A last point regards the specific effects of lexicality and length. Interestingly, the reading time measure proved less sensitive in capturing specific effects of lexicality and length than both vocal RTs and pronunciation times. In fact, all between-group interactions and these factors vanished in the ANOVA on *z*-scores, that is, when over-additivity was taken into account.[3] With regard to the length effect, the contribution of the pronunciation component (which has a small length effect) to reading times presumably dilutes the effect measured with RTs; thus, the effect on reading times is no longer different from that expected based on over-additivity. With regard to lexicality, note that the children with dyslexia showed a smaller difference in vocal RTs between words and non-words than the typically developing readers, an effect that was masked by over-additivity in the case of raw data and that emerges only when this is controlled for. As for pronunciation times, children with dyslexia showed greater difficulty in uttering non-words than typically developing readers. Even though the latter effect was comparatively small, it was apparently sufficient to mask the opposite effect found in vocal RTs.

Overall, the pattern of reading time data can be understood as the weighed sum of decoding and execution times, two components with markedly different response profiles. The reading time measure reveals only large group differences (and less marked than with vocal RTs) and does not allow detecting specific effects of length and lexicality. The group difference seems dominated by a large different global ability in graphemic decoding. Although the reading time measure is less sensitive than RTs, it is closer to functional reading and is likely to be used in clinical settings; thus, it is important to know which processes are represented or masked in this measure.

Conclusions

The present results are in keeping with a very large number of studies using RTs as a measure because of its high sensitivity in detecting cognitive components of responding in timed tasks. Therefore, in the particular case of reading in dyslexia, *vocal RTs* are particularly suited for detecting both global (graphemic decoding) and specific (length and lexicality) effects that contribute to the developmental reading disorder. Analysis of RTs on the DEM also allowed demonstrating that children with dyslexia are not impaired on the sensory–motor components

[3] The group by lexicality by length interaction was significant in the *z*-score analysis on reading times but was difficult to interpret.

of the reading task. Research has generally focused on this approach and has generally neglected examining the characteristics of the motor speech output associated with reading. Analysis of pronunciation times indicated that they were modulated by stimulus length and lexicality, raising the question as to whether stimulus processing is entirely completed within the RT window. Consideration of *pronunciation times* allowed us to bridge the gap between the RT measure and the more functionally meaningful reading time measure. *Total reading times* are less sensitive to specific effects than subcomponents. However, this measure is closer to functional reading; therefore, considering its internal structure may be crucial to understand the interplay between the different components in reading behavior.

Acknowledgments This work was supported by grants from the Department of Health and Sapienza University. We would like to thank Athanassios Protopapas for useful comments on the manuscript.

References

Balota DA, Abrams RA (1995) Mental chronometry: beyond onset latencies in the lexical decision task. J Exp Psychol Learn Mem Cogn 21:1289–1302

Bamber D (1979) State-trace analysis: a method of testing simple theories of causation. J Math Psychol 19:137–181

Barca L, Burani C, Arduino LS (2002) Word naming times and psycholinguistic norms for Italian nouns. Behav Res Meth Ins C 34:424–434

Bartram D (1996) The relationship between ipsatized and normative measures of personality. J Occup Organ Psych 69:25–39

Blomert L (2011) The neural signature of orthographic–phonological binding in successful and failing reading development. Neuroimage 57:695–703

Bonifacci P, Snowling MJ (2008) Speed of processing and reading disability: a cross-linguistic investigation of dyslexia and borderline intellectual functioning. Cognition 107:999–1017

Breznitz Z, Meyler A (2003) Speed of lower-level auditory and visual processing as a basic factor in dyslexia: electrophysiological evidence. Brain Lang 85:166–184

Cornoldi C, Colpo G (1995) Nuove prove di lettura MT per la scuola media inferiore, manuale. O.S. Organizzazioni Speciali, Firenze

Cornoldi C, Colpo G (1998) Prove MT per la scuola elementare 2. O.S. Organizzazioni Speciali, Firenze

Davies R, Cuetos F, Glez-Seijas RM (2007) Reading development and dyslexia in a transparent orthography: a survey of Spanish children. Ann Dyslexia 57:179–198

Davies R, Rodríguez-Ferreiro J, Suárez P, Cuetos F (2012) Lexical and sub-lexical effects on accuracy, reaction time and response duration: impaired and typical word and pseudoword reading in a transparent orthography. Read Writ. doi:10.1007/s11145-012-9388-1

De Luca M, Barca L, Burani C, Zoccolotti P (2008) The effect of word length and other sublexical, lexical and semantic variables on developmental reading deficit. Cogn Behav Neurol 21:227–235

De Luca M, Burani C, Paizi D, Spinelli D, Zoccolotti P (2010) Letter and letter-string processing in developmental dyslexia. Cortex 46:1272–1283

Di Filippo G, De Luca M, Judica A, Spinelli D, Zoccolotti P (2006) Lexicality and stimulus length effects in Italian dyslexics: role of over-additivity effect. Child Neuropsychol 12:141–149

Faust ME, Balota DA, Spieler HD, Ferraro FR (1999) Individual differences in information-processing rate amount: implications for group differences in response latency. Psychol Bull 125:777–799

Fischer R (2004) Standardization to account for cross-cultural response bias a classification of score adjustment procedures and review of research in JCCP. J Cross Cult Psychol 35:263–282

Forster KI, Forster JC (2003) DMDX: a windows display program with millisecond accuracy. Behav Res Meth Ins C 35:116–124

Kail R, Salthouse TA (1994) Processing speed as a mental capacity. Acta Psychol 86:199–225

Marinelli CV, Angelelli P, Di Filippo G, Zoccolotti P (2011) Is developmental dyslexia modality specific? A visual-acoustic comparison on Italian dyslexics. Neuropsychologia 49:1718–1729

Marsh EB, Hillis AE (2005) Cognitive and neural mechanisms underlying reading and naming: evidence from letter-by-letter reading and optic aphasia. Neurocase 11:325–337

Myerson J, Hale S, Zheng Y, Jenkins L, Widaman KF (2003) The difference engine: a model of diversity in speeded cognition. Psychon Bull Rev 10:262–288

Nicolson RI, Fawcett AJ (1994) Reaction time and dyslexia. Q J Exp Psychol [A] 47(1):29–48

Paizi D, De Luca M, Burani C, Zoccolotti P (2011) List context manipulation reveals orthographic deficits in Italian readers with developmental dyslexia. Child Neuropsychol 17:459–482

Paizi D, De Luca M, Zoccolotti P, Burani C (2013) A comprehensive evaluation of lexical reading in Italian developmental dyslexics. J Res Read 36:303–329

Prince M, Brown S, Heathcote A (2012) The design and analysis of state-trace experiments. Psychol Methods 17:78

Protopapas A (2007) CheckVocal: a program to facilitate checking the accuracy and response time of vocal responses from DMDX. Behav Res Meth 39:859–862

Pruneti CA, Fenu A, Freschi G, Rota S, Cocci D, Marchionni M, Rossi S, Baracchini Muratorio G (1996) Aggiornamento della standardizzazione italiana del test delle Matrici Progressive Colorate di Raven (CPM). Bollettino di Psicologia Applicata 217:51–57

Resulaj A, Kiani R, Wolpert DM, Shadlen MN (2009) Changes of mind in decision-making. Nature 461:263–266

Spinelli D, De Luca M, Di Filippo G, Mancini M, Martelli M, Zoccolotti P (2005) Length effect in word naming latencies: role of reading experience and reading deficit. Dev Neuropsychol 27:217–235

Spivey M (2007) The continuity of mind. Oxford University Press, New York

Stein J, Walsh V (1997) To see but not to read; the magnocellular theory of dyslexia. Trends Neurosci 20:147–152

Van den Broeck W, Geudens A (2012) Old and new ways to study characteristics of reading disability: the case of the nonword-reading deficit. Cogn Psychol 65:414–456

Van den Broeck W, Geudens A, van den Bos KP (2010) The nonword reading deficit of disabled readers: a developmental interpretation. Dev Psychol 46:717–734

Zeguers MH, Snellings P, Tijms J, Weeda WD, Tamboer P, Bexkens A, Huizenga HM (2011) Specifying theories of developmental dyslexia: a diffusion model analysis of word recognition. Dev Sci 14:1340–1354

Zheng Y, Myerson J, Hale S (2000) Age and individual differences in visuospatial processing speed: testing the magnification hypothesis. Psychon Bull Rev 7:113–120

Ziegler JC, Perry C, Ma-Wyatt A, Ladner D, Schulte-Korne G (2003) Developmental dyslexia in different languages: language-specific or universal? J Exp Child Psychol 86:169–193

Zoccolotti P, De Luca M, Judica A, Spinelli D (2008) Isolating global and specific factors in developmental dyslexia: a study based on the rate and amount model (RAM). Exp Brain Res 186:551–560

Zoccolotti P, De Luca M, Di Filippo G, Judica A, Martelli M (2009) Reading development in an orthographically regular language: effects of length, frequency, lexicality and global processing ability. Read Writ 22:1053–1079

Zoccolotti P, De Luca M, Lami L, Pizzoli C, Pontillo M, Spinelli D (2012) Multiple stimulus presentation yields larger deficits in children with developmental dyslexia: A study with reading and RAN-type tasks. Child Neuropsychol. 2012 Aug 28. [Epub ahead of print]