Default spacing is the optimal spacing for word reading

van den Boer, M.; Hakvoort, B.E.

DOI
10.1080/17470218.2014.964272

Publication date
2015

Document Version
Final published version

Published in
The Quarterly Journal of Experimental Psychology

License
Article 25fa Dutch Copyright Act

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2015
Vol. 68, No. 4, 697–709, http://dx.doi.org/10.1080/17470218.2014.964272

Default spacing is the optimal spacing for word reading

Madelon van den Boer and Britt E. Hakvoort

Research Institute of Child Development and Education, University of Amsterdam, Amsterdam, The Netherlands

Increased interletter spacing is thought to reduce crowding effects and to enhance fluent reading. Several studies have shown beneficial effects of increased interletter spacing on reading speed and accuracy, especially in poor readers. Therefore, increased interletter spacing appears to be a relatively easy way to enhance reading performance. However, in adult readers reading speed was shown to be impeded with increased interletter spacing. Thus, findings on interletter spacing are still inconclusive. In the current study we examined the effect of a range of interletter spacings (−0.5, default, 0.5, 1, 1.5, 2) on naming fluency of monosyllabic and bisyllabic words in beginning (Grade 2) and more advanced (Grade 4) readers. Additionally we tested the effects of spacing in a subsample of poor readers. In contrast to previous findings, neither beginning nor advanced readers benefited from an increase in interletter spacing. However, they did show reduced reading fluency when letter spacing was smaller than the default spacing, which may be indicative of a crowding effect. Poor readers showed a similar pattern. We conclude that an increase in interletter spacing has no effect on word naming fluency.

Keywords: Interletter spacing; Reading fluency.

Evidence from recent studies suggests that reading fluency may be influenced not only by cognitive skills that underlie reading, such as phonological awareness (e.g., Vellutino, Fletcher, Snowling, & Scanlon, 2004), but also by physical properties of the presented text. Interletter spacing—that is, the spacing between two adjacent graphemes (henceforth: LS)—in particular has been widely studied (e.g., Perea, Panadero, Moret-Tatay, & Gómez, 2012; Risko, Lanthier, & Besner, 2011; Spinelli, de Luca, Judica, & Zoccolotti, 2002; Vinckier, Qiao, Pallier, Dehaene, & Cohen, 2011; Zorzi et al., 2012). The idea that a larger LS facilitates reading performance became increasingly popular, as it could be a relatively easy way to ameliorate reading difficulties. It has even been suggested that increased LS is especially effective for children with dyslexia, for whom reading difficulties are severe and persistent (Perea et al., 2012; Spinelli et al., 2002; Zorzi et al., 2012).

Theoretically, enlarging LS is thought to reduce the amount of visual crowding a reader experiences. Crowding refers to an impediment in discerning features of a particular target letter because of interference from features of peripheral letters (Bouma, 1970). When the space between two letters becomes larger, readers might have less difficulty identifying the letters, which could result in faster reading times. It has been suggested that poor readers are especially affected by crowding

Correspondence should be addressed to Madelon van den Boer, Research Institute of Child Development and Education, University of Amsterdam, P.O. Box 94208, 1090 GE Amsterdam, The Netherlands. E-mail: m.vandenboer@uva.nl

We thank Peter de Jong, Elise de Bree, and Titia van Zuijen for their valuable comments on an earlier version of this manuscript.
Both authors contributed equally.

(Spinelli et al., 2002) and that they could benefit more from increased LS than fluent readers, who experience less crowding to begin with (Martelli, Di Filippo, Spinelli, & Zoccolotti, 2009).

A substantial number of studies investigating LS are available, using a large variety of tasks, but mainly lexical decision, word naming, and sentence reading, and targeting children and adults with and without reading difficulties. Table 1 presents an overview of these studies. Overall, the effect of LS on reading performance varies. In lexical decision studies, for example, Perea and colleagues (Perea et al., 2011, 2012) presented skilled adult readers, skilled young readers, and young readers with dyslexia with words and nonwords using default LS and a +1.2 LS condition (e.g., animal) and found that all three groups benefited from an increase in LS. Lexical decision times were faster in the increased LS condition. Vinckier et al. (2011), in contrast, found negative effects of LS on both speed and accuracy for adults performing a lexical decision task.

Studies on sentence and text reading have indicated that the effects of LS might be specific to struggling readers, such that children with dyslexia benefit from an increase in LS (Perea et al., 2012; Zorzi et al., 2012), but that typically reading adults (Chung, 2002; Cohen et al., 2008; Paterson & Jordan, 2010; Perea & Gomez, 2012; Slattery & Rayner, 2013; Yu et al., 2007), and even beginning average readers (Reynolds & Walker, 2004) do not. Similar findings have been reported by studies investigating word naming. Spinelli et al. (2002), for example, reported positive effects of LS on word naming speed in children with dyslexia. In adults, however, LS was found to have no effect (Tai et al., 2006), or even a negative effect (Risko et al., 2011), on word naming.

Taken together, evidence from the abovementioned studies appears to suggest that LS fosters reading performance in poor readers, but has no effect on, or even interferes with, fluent reading in normal readers. Risko et al. (2011) suggest that effects of LS are not found in fluent readers, because increased LS hampers the parallel processing of letters necessary for fluent reading (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). Increasing LS thus forces readers to resolve to slower serial analysis of letters. In line with this idea, Tai et al. (2006) showed that fixation durations decrease as a result of LS, but that at the same time the number of fixations within words increases, which is an indication of slower serial processing. Furthermore, in a functional magnetic resonance imaging (fMRI) study, Cohen et al. (2008) found that increasing word degradation, measured by manipulations in spacing, rotation, and position, induced activation of a second neural pathway in addition to activation in the visual word form area, which is typically associated with and specialized in processing print. When text is manipulated, a more serial reading process needs to be employed involving posterior parietal areas associated with attention. The authors found significant increases in activation of this dorsal route with LS of +2.25 and +3. Theoretically, the effects of hampered parallel processing on reading fluency can be explained with the bigram coding hypothesis (Vinckier et al., 2011). According to this hypothesis, word recognition is invariant across small changes in LS. Larger LS, however, interferes with parallel processing of letter strings, because bigram detectors—that is, neurons that are sensitive to combinations of letters over a range of two or three letter positions—are disrupted (Dehaene, Cohen, Sigman, & Vinckier, 2005).

In all, the results on, as well as the theoretical accounts of, the effects of LS vary. On the one hand, reading performance has been argued to increase with larger LS due to a decrease in crowding, especially so for poor readers. On the other hand, reading speed has been shown to be impeded with increased LS, because parallel letter processing is disrupted, especially so in adults.

Legge and Bigelow (2011) have shown that there is a range of print sizes that allows for fluent reading. It can thus be expected that the reading system is also invariant to small changes in LS. Too little or too much LS, however, could interfere with fluent reading as a result of crowding or serial letter processing, respectively. In the current study, we aimed to examine the effects of LS in children in a word naming task. To date,

Table 1. *Overview of studies examining the effect of letter spacing by type of task in children and adults with and without reading difficulties*

Authors	*Language*	*Type of task*	*Participants*	n	*Font/Size*	*Conditions*	*Speed*	*Accuracy*
Martelli et al. (2009)	Italian	Letter identification	11-year-old children 11-year-old dyslexics	32 29	Courier, n/a	The centre to centre distance between trigrams was varied to determine the critical spacing for letter identification		Critical spacing to correctly identify letters is significantly larger for dyslexic children
Marzouki and Grainger (2014)	French	Letter identification (within 5-letter strings)	Adults	21	Courier New, n/a	Spacing −6 pixels Spacing default Spacing +6 Spacing +12 Spacing +18 Spacing +24 Spacing +30 Spacing +36 Spacing +42		Correct identification of letters in Positions 1 and 5 decreases as spacing increases, whereas accuracy for Position 3 increases.
Moll and Jones (2013)	English	Letter naming (Eyetracking)	Adults	17	Courier, n/a	Spacing Foveal (1°), Parafoveal (2.5°), or Peripheral (5°) Spacing default Spacing +1.2 high- and low-frequency words and nonwords	Gaze duration was longer for dyslexic participants, especially in the parafoveal condition.	
			Dyslexic adults	17			Naming times were not influenced by the spacings.	
Perea, Moret-Tatay, and Gómez (2011)	Spanish	Lexical decision	Adults	38	Times New Roman, 14		Words with wider spacing had faster reaction times. There were no effects of frequency. There were no effects of lexicality.	No effects were found of spacing on accuracy of words. On nonwords, increased spacing led to a decrease in accuracy.
Perea et al. (2011)	Spanish	Lexical decision	Adults	16	Times New Roman, 14	Spacing default Spacing +1.2 5-letter and 8-letter words and nonwords	Words with wider spacing had faster reaction times. There were no effects of length. Eight-letter nonwords had longer response times with wider spacing.	No significant effects on accuracy for words and nonwords.
Perea et al. (2012)	Spanish	Lexical decision	Adults 7–8-year-olds 9–10-year-olds 11–13-year-old dyslexics	24 24 24 8	Times New Roman, 14	Spacing default Spacing +1.2 long and short words and nonwords	All groups had significantly lower reaction times in the larger spacing condition with longer words.	Dyslexics' accuracy on long words also significantly improved with larger spacing.

(*Continued overleaf*)

Table 1. Continued.

Authors	*Language*	*Type of task*	*Participants*	n	*Font/Size*	*Conditions*	*Speed*	*Accuracy*
Vinckier et al. (2011)	French	Lexical decision	Adults	12	Arial, 7	Spacing default Spacing +1 space Spacing +2 spaces Spacing +3 spaces Spacing +4 spaces 4, 6, 8 letters	Reaction times increased as a function of spacing and of word length.	Error rates increased significantly with spacing.
Spinelli et al. (2002)	Italian	Word naming	12-year-old dyslexics 12-year-old controls	27 99	Geneva, 12	Spacing default Spacing +0.32 Spacing +0.41 Spacing +0.595 (centre to centre distance)	For controls, reading times for spacing +0.595 were significantly slower than those for the default spacing. Dyslexics benefited from an increased spacing of +0.32 but were slower for spacing +0.595.	
Risko et al. (2011)	English	Word naming	Adults	56	Courier, 18	Spacing default Spacing 2 spaces between letters	Longer reaction times were found for the larger spacing condition.	No effects were found of spacing on accuracy.
Risko et al. (2011)	English	Word and nonword naming	Adults	64	Courier, 18	Spacing default Spacing 2 spaces between letters 3- and 4-letter words	Increased spacing yielded longer reaction times in 3-letter words and nonwords, and even more so in 4-letter words and nonwords.	No effects were found of spacing on accuracy.
Tai, Sheedy, and Hayes (2006) (abstract)	English	Word naming	Adults	41	Verdana, 10	Spacing default, spacing condensed (−0.5, −1, −1.5, −1.75) Spacing expanded (0.5, 1, 1.5, 2)	No effects were found of spacing on reading speed. The number of fixations increased, but fixation durations decreased across the spacing continuum.	
Perea and Gomez (2012)	Spanish	Sentence reading (Eyetracking)	Adults	24	Times New Roman, 14	Spacing default Spacing +1 Spacing +1.5	Gaze duration was significantly lower in spaced conditions but the number of forward fixations in the largest spacing was significantly larger.	
Chung (2002)	English	Sentence reading (Eyetracking)	Adults	6	Courier, n/a	Spacing −0.5 Spacing −0.707 Spacing default Spacing +1.414 Spacing +2	Reading speed significantly decreased as a function of spacing for the smaller than for default spacings. Reading speed in the wider spacings did not differ from that for the default spacing.	

Slattery and Rayner (2013)	English	Sentence reading	Adults	32	Times New Roman/ Cambria	−0.5 pixel Default +0.5 pixel +1 pixel	Default spacing was optimal for reading speed.	Accuracy was unaffected by spacing manipulation.
Zorzi et al. (2012)	Italian/French	Sentence reading	French dyslexics Italian dyslexics Italian reading-age-matched controls (mean age 10 years)	40 34 30	Times New Roman, 14	Spacing default Spacing +2.5	Reading times improved significantly in the spaced condition for dyslexics.	Dyslexics' accuracy significantly improved in the spaced condition
Yu, Cheung, Legge, and Chung (2007)	English	Sentence reading/ Letter identification	Adults	5	Courier, small print size	Spacing −0.5 Spacing −0.707 Spacing default Spacing +2	Reading times were fastest for the default spacing. Increases in reading times were found for both smaller and larger than default spacings.	Accuracy for letter identification was optimal in the default condition. Accuracy decreased for both smaller and larger than default spacings.
Paterson and Jordan (2010)	English	Sentence reading	Adults	16	Courier, n/a	Spacing default Spacing increased (details n/a) Interword spacing increased by 1, 2 or 3 spaces	Reading rates were significantly slower and fixation and gaze durations significantly larger in increased spacing with 1 interword space, but not with 2 or 3 interword spaces.	
Reynolds and Walker (2004)	English	Text reading	5–7-year-olds	24	Century Educational,19	Spacing −20 Spacing default Spacing +15 Spacing +25	The effect of letter spacing on reading rate was not significant.	
Perea et al. (2012)	Spanish	Text reading	9–10-year-olds 11–13-year-old dyslexics	20 20	Information, n/a	Spacing default Spacing "slightly wider"—details n/a	Dyslexics read significantly more words per minute in the wider spacing condition.	Dyslexics were significantly more accurate in the wider spacing condition.
Cohen, Dehaene, Vinckier, Jobert, and Montavont (2008)	English	Silent reading (Neuroimaging)	Adults	12	Courier, n/a	Spacing default Spacing +0.75 Spacing +1.5 Spacing +2.25 Spacing +3	Reading times increased for spacings +2.25 and +3.	

word naming has been mostly addressed in adults and somewhat older children. Therefore, we included young readers: one group of beginning (Grade 2) and one group of more advanced (Grade 4) readers. Previous studies have shown that especially poor readers can benefit from increased letter spacing (e.g., Perea et al., 2012; Spinelli et al., 2002; Zorzi et al., 2012). We aimed to examine whether this effect extends to very young beginning readers in Grade 2, similar in reading level to children with dyslexia. In addition, we focused specifically on the effect of LS for the poor readers within the sample, to determine whether effects of LS might be specific to struggling readers since word identification is assumed to be the core deficit in children with dyslexia (e.g., Vellutino et al., 2004). Six LS conditions were examined, ranging from −0.5 to +2. We included a decreased spacing condition to examine possible effects of increased crowding. If the effects of LS can indeed be ascribed to crowding, a reduced LS can be expected to increase crowding and thus to interfere with reading fluency. A number of increased LS conditions was included to examine whether in children increased LS benefits or interferes with fluent reading. The focus was on reaction times and accuracy rates, as is done in most studies, as well as on word reading fluency, to account for possible speed–accuracy trade-off effects. Monosyllabic and bisyllabic words were studied as effects of LS have been shown to be larger for longer words (e.g., Perea et al., 2012; Risko et al., 2011).

EXPERIMENTAL STUDY

Method

Participants

One hundred and five children from Grade 2 (54 boys, mean age = 7 years 10 months, *SD* = 4.94 months) and 92 children from Grade 4 (46 boys, mean age = 9 years 11 months, *SD* = 4.97 months) participated in this study. All children spoke Dutch at home as their native language (98.5%), or second language (1.5%). Scores on the One Minute Test (*Eén Minuut Test;* Brus & Voeten, 1995), a standardized test of word reading fluency with an average of 10 and a standard deviation of 3, indicated that the samples included a representative range of reading abilities (Grade 2: *M* = 11.40, *SD* = 3.05; Grade 4: *M* = 9.97, *SD* = 2.87).

Materials and procedure

Reaction times and accuracy rates as a consequence of LS were measured using a naming task. A total of 144 monosyllabic words were selected from the CELEX database (Baayen, Piepenbrock, & van Rijn, 1993). These words were divided over six lists of 24 words each. The lists were matched on onset, frequency, and neighbourhood size. Each list was presented in a different LS condition (−0.5, default, +0.5, +1, +1.5, +2). To control for list-specific effects, six versions of the naming tasks were designed, such that each wordlist occurred in each LS condition. Following the same procedure, 144 bisyllabic words were selected, which were also divided over six lists, and were presented in each LS condition across six versions of the task. An example of a monosyllabic and a bisyllabic word as presented in the different versions of the task is presented in Figure 1.

Children were randomly assigned to reading either monosyllabic or bisyllabic words and were

- 0.5	berg	bliksem
0	berg	bliksem
+ 0.5	berg	bliksem
+ 1	berg	bliksem
+ 1.5	berg	bliksem
+ 2	berg	bliksem

Figure 1. *Example of the monosyllabic word berg (mountain) and the bisyllabic word bliksem (lightning) as presented in the different interletter spacing (LS) conditions.*

Table 2. *Reading scores of second and fourth graders reading monosyllabic and bisyllabic words*

	Monosyllabic words		*Bisyllabic words*			
Grade	N	M *(*SD*)*	N	M *(*SD*)*	t *statistic*	p*-value*
Grade 2	52	42.62 (16.65)	53	45.26 (13.51)	0.894	.373
Grade 4	46	66.70 (13.89)	46	62.3 (13.09)	1.337	.185

then again randomly assigned to one of the six versions of the task. Thus, each child was presented with a total of 144 words, which were either monosyllabic or bisyllabic. For both Grade 2 and Grade 4, scores on the One Minute Test (Brus & Voeten, 1995) did not differ for children reading monosyllabic and children reading bisyllabic words (see Table 2).

The naming task was programmed in E-prime (Schneider, Eschman, & Zuccolotto, 2002). Words were presented on a computer screen one at a time in 14-point Times New Roman font. A plus sign was presented for 750 ms to focus attention. Then the word appeared and children were asked to read the word aloud as fast and accurately as possible. A voice key registered naming latencies from the onset of stimulus presentation until the onset of the response. The experimenter coded the child's responses (correct, incorrect, or invalid) using a response box. Responses were coded invalid when the word had disappeared from the screen by a sound other than the child's pronunciation of the word. The order of the word lists within the task (and therefore the LS conditions) and of the words within each list were random for each child. The experiment was preceded by six practice trials. In addition, each list started with two practice items, which were not included in the analyses.

Data analysis

The data were analysed in two ways. First, reaction times and accuracy rates were analysed with multilevel models using MLwiN 2.24 (Rasbash, Steele, Browne, & Goldstein, 2008). Reaction times and accuracy rates are embedded in a hierarchical structure, because responses to items (Level 1) are nested under individuals (Level 2). Within a multilevel model, effects from items and participants can be captured within a single model (e.g., Quené & van den Bergh, 2004).

Reaction times were rescaled as items read per second to normalize scores. These inverted scores reflect reading rate, with higher scores reflecting better performance. Reaction times were modelled via dummy variables that were specified for every LS condition. Parameters for these dummy variables reflect the condition means. Separate dummy variables were specified for Grade 2 and Grade 4 and for monosyllabic and bisyllabic words, amounting to a total of 24 fixed parameters. The accuracy data were dummy coded (0 is incorrect, 1 is correct). To analyse these data, a binomial rather than normal distribution was assumed, but the model contained the same parameters as did the model for the reaction time data. The main effects and interaction of grade and word length were tested with a single contrast, denoting the difference between Grades 2 and 4, and between monosyllabic and bisyllabic words, respectively. The main effect of and interactions with LS were tested with five orthogonal contrasts denoting the differences between all six LS conditions. Furthermore, the effect of each LS condition was tested against the default LS with a single contrast.

Second, because skilled reading is characterized by both rapid and accurate word identification, fluency scores were calculated, in which both speed and accuracy can be taken into account simultaneously. Average reading latencies per LS condition were calculated for each child and were transformed to the number of items read per second. The proportion of items correct per LS condition was calculated over valid trials. Fluency scores were calculated by multiplying the number of items read per second by the proportion of items correct. Since fluency scores can only be calculated for a set of items, not for each item, the fluency scores were analysed with a repeated measures analysis of variance (RM-ANOVA), rather than a multilevel model.

To examine whether the effect of LS on reading performance is specific to struggling readers, we ran the analyses both for the entire group of participants and for a selection of the poor readers in the sample. Poor readers were identified based on the scores on the One Minute Test (standard scores of 7 or lower; Brus & Voeten, 1995). Our sample included a total of 29 poor readers (15 second graders, 14 fourth graders; 15 children who read monosyllabic words, 14 who read bisyllabic words). For this group we only examined the effect of LS condition, not of grade and word length.

Results

Data cleaning

Reading latencies were excluded from analysis if the response was incorrect (8.2%), if the voice key was not validly triggered (5.5%), if latencies were less than 250 ms (2.3%), and if latencies were more than three standard deviations from a participant's mean (1.5%). Accuracy was not available for the invalid trials (5.5%). The mean reaction times, accuracy rates, and fluency scores separated by grade and word length are presented in Table 3.

Accuracy rates

The main effect of grade was significant, $\chi^2(1) = 15.922$, $p < .001$. Fourth graders read more accurately than second graders. Monosyllabic words were read more accurately than bisyllabic words, $\chi^2(1) = 14.553$, $p < .001$. The interaction between grade and word length did not reach significance, $\chi^2(1) = 2.953$, $p = .086$.

The main effect of LS was significant, $\chi^2(5) = 31.757$, $p < .001$. Planned contrasts, however, indicated that accuracy was significantly lower only for words in the −0.5 LS condition as compared to the default spacing, $\chi^2(1) = 12.174$, $p < .001$. This effect was not significantly different for Grades 2 and 4, $\chi^2(1) = 0.164$, $p = .686$, nor for monosyllabic and bisyllabic words, $\chi^2(1) = 0.132$,

Table 3. *Reaction times and accuracy rates for each spacing, separated by grade and word length*

	Accuracy				*Reaction times (in ms)*				*Fluency (per second)*			
	Monosyllabic		*Bisyllabic*		*Monosyllabic*		*Bisyllabic*		*Monosyllabic*		*Bisyllabic*	
LS	*Grade 2*	*Grade 4*	*Grade 2*	*Grade 4*	*Grade 2*	*Grade 4*	*Grade 2*	*Grade 4*	*Grade 2*	*Grade 4*	*Grade 2*	*Grade 4*
−0.5	89.3	95.0	84.4	90.3	974	620	1167	722	1.12	1.57	.93	1.36
	(12.1)	(5.1)	(17.1)	(11.4)	(708)	(88)	(884)	(272)	(0.41)	(0.25)	(0.40)	(0.36)
Default	91.3	96.3	88.4	91.2	957	602	1183	693	1.19	1.63	.98	1.41
	(11.9)	(4.5)	(12.9)	(10.5)	(702)	(83)	(874)	(215)	(0.43)	(0.26)	(0.41)	(0.38)
+0.5	90.5	97.0	88.5	92.7	900	590	1119	687	1.19	1.69	.99	1.44
	(12.7)	(4.0)	(12.7)	(9.2)	(550)	(98)	(769)	(223)	(0.41)	(0.28)	(0.39)	(0.37)
+1.0	91.2	96.3	89.1	91.4	945	588	1120	695	1.19	1.68	1.00	1.40
	(11.3)	(4.8)	(14.0)	(10.0)	(704)	(92)	(822)	(199)	(0.43)	(0.29)	(0.40)	(0.35)
+1.5	92.6	96.3	88.9	91.9	979	593	1184	694	1.20	1.67	1.00	1.42
	(9.3)	(5.1)	(12.4)	(9.7)	(790)	(90)	(963)	(244)	(0.42)	(0.29)	(0.42)	(0.36)
+2.0	92.2	96.5	89.0	91.0	930	607	1213	709	1.20	1.62	.99	1.39
	(11.7)	(5.0)	(12.4)	(10.3)	(619)	(92)	(943)	(247)	(0.41)	(0.27)	(0.43)	(0.39)

Note: LS = Interletter spacing. Standard deviations in parentheses.

$p = .716$. Accuracy rates for the other LS conditions did not differ significantly from those for the default spacing. The interactions between LS and grade, $\chi^2(5) = 6.184$, $p = .289$, and between LS and word length, $\chi^2(5) = 0.927$, $p = .968$, were not significant.

For the poor readers, the effects of letter spacing were the same. The main effect of LS was significant, $\chi^2(5) = 23.648$, $p < .001$. Planned contrasts, however, again indicated that accuracy was significantly lower only for words in the −0.5 LS condition as compared to the default spacing, $\chi^2(1) = 9.433$, $p = .002$. Accuracy rates for the other LS conditions did not differ significantly from those for the default spacing.

Reading rates

The main effects of grade, $\chi^2(1) = 85.288$, $p < .001$, and word length, $\chi^2(1) = 12.502$, $p < .001$, were both significant. Reading rates were higher in Grade 4 than in Grade 2, and higher for monosyllabic than for bisyllabic words. The interaction between grade and word length did not reach significance, $\chi^2(1) = 0.045$, $p = .832$.

The main effect of LS was significant, $\chi^2(5) = 28.005$, $p < .001$. Planned contrasts, however, indicated that words in the −0.5 LS condition were read significantly slower than words in the default spacing, $\chi^2(1) = 6.308$, $p = .012$. This effect was not significantly different for Grades 2 and 4, $\chi^2(1) = 1.718$, $p = .190$, nor for monosyllabic and bisyllabic words, $\chi^2(1) = 0.642$, $p = .423$. Furthermore, words in the +0.5 LS condition were read significantly faster than words in the default spacing, $\chi^2(1) = 6.257$, $p = .012$. Again, this effect was neither significantly different for Grades 2 and 4, $\chi^2(1) = 0.478$, $p = .489$, nor significantly different for monosyllabic and bisyllabic words, $\chi^2(1) = 0.874$, $p = .350$. Reading rates for the other LS conditions did not differ significantly from the default spacing. The interactions between LS and grade, $\chi^2(5) = 2.715$, $p = .744$, and between LS and word length, $\chi^2(5) = 4.911$, $p = .427$, were not significant.

For the poor readers, the results were very similar. The main effect of LS was significant, $\chi^2(5) = 11.083$, $p = .050$. Planned contrasts indicated again that words in the +0.5 LS condition, $\chi^2(1) = 5.660$, $p = .017$, were read significantly faster than words in the default spacing. Reading rates for the other LS conditions, including LS −0.5, however, did not differ significantly from those for the default spacing.

Reading fluency

Results on the fluency scores were only slightly different from those obtained in the separate analyses of reading accuracy and reading rate (based on reaction times). The main effects of grade, $F(1, 193) = 77.018$, $p < .001$, $\eta_p^2 = .285$, and word length, $F(1, 193) = 18.892$, $p < .001$, $\eta_p^2 = .089$, were both significant. Reading fluency was higher in Grade 4 than in Grade 2, and higher for monosyllabic than for bisyllabic words. The main effect of LS was also significant, $F(5, 189) = 11.261$, $p < .001$, $\eta_p^2 = .230$. Planned contrasts, however, indicated that only LS −0.5 differed significantly from the default spacing, $F(1, 193) = 23.835$, $p < .001$, $\eta_p^2 = .110$. Reading fluency for the other LS conditions did not differ significantly from that for the default spacing. The interactions between word length and grade, $F(1, 193) = 0.133$, $p = .715$, $\eta_p^2 = .001$, LS and grade, $F(5, 189) = 1.014$, $p = .410$, $\eta_p^2 = .026$, and LS and word length, $F(5, 189) = 0.428$, $p = .828$, $\eta_p^2 = .011$, were not significant.

For the poor readers, the effects of LS on reading fluency were the same. The main effect of LS was significant, $F(5, 24) = 2.718$, $p = .044$, $\eta_p^2 = .362$. Planned contrasts, however, again indicated that only LS −0.5 differed significantly from the default spacing, $F(1, 28) = 5.500$, $p = .026$, $\eta_p^2 = .164$. Reading fluency for the other LS conditions did not differ significantly from the default spacing.

Discussion

Recently, LS has become a popular topic of research due to it being a seemingly simple means of increasing reading performance, especially in the poorest readers. Findings on LS, however, have varied. On the one hand, increased LS has been shown to foster reading performance as a result of decreased crowding effects, but on the other hand, increased LS was shown to hamper reading performance because it induces serial rather than parallel letter processing (see Table 1 for an overview of the studies).

In the current study, we did not find any beneficial effects of increased LS. Accuracy rates did not increase with increased letter spacing. Accuracy rates were, however, found to decrease in the smaller than default LS condition. The −0.5 LS condition also resulted in slower reading rates, although this was not found in the poor readers. Beginning and more advanced readers, as well as the selected group of poor readers, read words in the +0.5 LS condition slightly faster than words in the default condition. However, this difference was not significant in the analyses of fluency scores, indicating that the increase in reading rates goes hand in hand with a small decrease in accuracy, which together do not result in increased reading fluency. These findings are comparable to those of eye tracking studies that have shown that increased LS results in shorter fixation durations, but at the same time in an increase in the number of fixations (Perea & Gomez, 2012; Tai et al., 2006). Taken together, these findings indicate that improvements as a result of LS on one specific measure (i.e., reaction times or fixation durations) do not necessarily mean that increased LS fosters overall reading performance. Word naming appeared to be invariant across the increased LS conditions studied (from default up to +2). Results were the same for advanced readers in Grade 4 and beginning readers in Grade 2, indicating that the effects of LS are not different for children of different reading levels. Separate analyses for the poor readers in the sample indicated that the effects of LS were also not different for children with specific difficulties in learning to read.

Interestingly, reading performance was shown to deteriorate in a smaller than default LS condition (−0.5). This result indicates that the invariance across the increased LS conditions should not be ascribed to lack of power. The poorer reading performance at the decreased LS may point toward an effect of crowding, where reading performance is hampered by increased crowding. However, this effect was not found to be larger in beginning or poor readers than in more advanced readers. All children appeared to be affected by crowding in the decreased LS condition, but not in the default LS condition. This is in contrast with previous studies on word naming (Spinelli et al., 2002) and letter identification (Callens, Whitney, Tops, & Brysbaert, 2013; Martelli et al., 2009). Results from a letter identification study with students by Callens et al. (2013), for example, showed a decrease in letter identification within letter triplets with eccentricity in poor and fluent readers. Poor readers, however, were even less accurate at identifying the centre target letter if the stimulus was further away from the fixation location than controls, which the authors ascribed to enhanced crowding effects. Note, however, that although in many studies the effects of increased LS have been explained in terms of crowding, very few studies have explicitly measured crowding. Future studies are needed to understand to what extent findings, such as in the current study, but also in previous studies, should be ascribed to increased or decreased effects of crowding.

Our findings on the increased LS conditions could be in line with the bigram coding hypothesis, which states that bigram detectors sensitive to letter combinations allow small increases in letter spacing, but are disrupted when LS is too large (Dehaene et al., 2005). The LS conditions of the current study should then be assumed to fall within the scope of variations that are allowed by the reading system without hampering performance. Previous studies showed that for adults, reading performance is disrupted in LS conditions of +2 or larger (e.g., Cohen et al., 2008; Risko et al., 2011; Vinckier et al., 2011). Future studies could focus on the turning point at which children's word reading fluency is affected by increased LS. Based on the current findings, however, we conclude that default LS is the optimal LS for children's word reading fluency, since a decrease in LS hampers reading performance, and increases in LS (up to +2) do not foster reading fluency.

Our results do not support previous findings that increased LS is especially beneficial for poor readers. Differences in findings could be ascribed to the tasks and materials used. First, our results differ from the positive effects of increased LS found in lexical decision tasks (e.g., Perea et al., 2012). If increased LS indeed fosters letter identification, effects could be expected on lexical decision, which is often seen as a search of the orthographic lexicon (e.g., Coltheart et al., 2001). Word naming, in contrast, requires letter identification, but most importantly activation of a phonological code. The different effects of LS could indicate that word processing in a lexical decision task differs from word processing in a naming task, in line with previous findings that lexical decision and naming differ, for example in the effects of lexical variables (i.e., length, frequency, neighbourhood size) on identification latencies (e.g., Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004).

Second, we did not find an effect on word reading fluency, but fluency at the sentence or text level has been shown to improve with increased LS, especially for poor readers (e.g., Perea et al., 2012; Zorzi et al., 2012). It remains to be studied whether the results at the text level could be ascribed to increases in attention, given that the increased LS text is of an unusual size and format. Another possibility, however, could be that crowding, to which positive effects of increased LS are most often ascribed (e.g., Perea et al., 2012; Spinelli et al., 2002; Zorzi et al., 2012), affects performance at the sentence more than at the word level. Indeed, Spinelli et al. (2002) showed that children with dyslexia were more affected than controls by surrounding words when identifying target words. Furthermore, Moll and Jones (2013) recently showed that crowding operates in a larger visual span in dyslexic than in control readers. Similarly, poor readers have been shown to benefit from text presented in shorter lines, assumed to eliminate crowding of surrounding text (Schneps et al., 2013).

A final important issue in studies on the effects of LS is the font that materials are presented in. Some authors have used fonts of fixed width (e.g., Chung, 2002; Paterson & Jordan, 2010; Yu et al., 2007), but others have used fonts with a proportional width (e.g., Perea et al., 2012). In fixed-width fonts the total width of each letter is equal, but as a consequence, the spacing between letters varies. In proportional fonts, in contrast, the width of each letter varies, but the spacing between the letters is equal. For this reason, we chose to use a proportional-width font because our focus was on variations in LS rather than in the total width of words. Furthermore, the font used in the current study has been used previously (see Table 1) and is widely used in child literature. Findings of previous studies, however, do not seem to be specific to either fixed- or proportional-width fonts, so we expect the influence of our choice of font on the outcome to be limited.

In sum, results from this study do not support the idea that a larger LS facilitates children's reading fluency. Our findings suggest that a larger LS does not have an effect on naming fluency, but that a smaller spacing reduces naming fluency. The default spacing thus seems to be optimal for both beginning and more advanced readers. With respect to poor readers, no indications were found that increased LS fosters word reading fluency, which is currently taken to be the core deficit in dyslexia (e.g., Vellutino et al., 2004). However, it has been shown that dyslexics as a group (Moll & Jones, 2013), and a subgroup of dyslexics in particular (Spinelli et al., 2002), show increased crowding effects when identifying a target among other items. Accordingly, spacing might be more effective at the sentence or text, rather than the word, level.

Original manuscript received 4 September 2014
Accepted revision received 6 September 2014
First published online 7 November 2014

REFERENCES

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database [CD-ROM]*. Philadelphia, PA: University of Pennsylvania, Linguistic Data Consortium.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of*

Experimental Psychology: General, *133*, 283–316. doi:10.1037/0096-3445.133.2.283

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*, 177–178. doi:10.1038/226177a0

Brus, B., & Voeten, B. (1995). *Eén minuut test vorm A en B. Verantwoording en handleiding* [one-minute-test manual]. Lisse, The Netherlands: Swets & Zietlinger.

Callens, M., Whitney, C., Tops, W., & Brysbaert, M. (2013). No deficiency in left-to-right processing of words in dyslexia but evidence for enhanced visual crowding. *The Quarterly Journal of Experimental Psychology*, *66*, 1803–1817. doi:10.1080/17470218.2013.766898

Chung, S. T. L. (2002). The effect of letter spacing on reading speed in central and peripheral vision. *Investigative Ophthalmology & Visual Science*, *43*, 1270–1276.

Cohen, L., Dehaene, S., Vinckier, F., Jobert, A., & Montavont, A. (2008). Reading normal and degraded words: Contribution of the dorsal and ventral visual pathways. *NeuroImage*, *40*, 353–366. doi:10.1016/j.neuroimage.2007.11.036

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204–256. doi:10.1037/0033-295X.108.1.204

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *TRENDS in Cognitive Sciences*, *9*, 335–341. doi:10.1016/j.tics.2005.05.004

Legge, G. E., & Bigelow, C. A. (2011). Does print size matter for reading? A review of findings from vision science and typography. *Journal of Vision*, *11*, 1–22. doi:10.1167/11.5.8

Martelli, M., Di Filippo, G., Spinelli, D., & Zoccolotti, P. (2009). Crowding, reading, and developmental dyslexia. *Journal of Vision*, *9*(4), 14, 1–18. doi:10.1167/9.4.14

Marzouki, Y., & Grainger, J. (2014). Effects of stimulus duration and inter-letter spacing on letter-in-string identification. *Acta Psychologica*, *148*, 49–55. doi:10.1016/j.actpsy.2013.12.011

Moll, K., & Jones, M. (2013). Naming fluency in dyslexic and nondyslexic readers: Differential effects of visual crowding in foveal, parafoveal, and peripheral vision. *The Quarterly Journal of Experimental Psychology*, *66*, 2085–2091. doi:10.1080/17470218.2013.840852

Paterson, K. B., & Jordan, T. R. (2010). Effects of increased letter spacing on word identification and eye guidance during reading. *Memory & Cognition*, *38*, 502–512. doi:10.3758/MC.38.4.502

Perea, M., & Gomez, P. (2012). Subtle increases in interletter spacing facilitate the encoding of words during normal reading. *PLOSone*, *7*, e47568. doi:10.1371/journal.pone.0047568

Perea, M., Moret-Tatay, C., & Gómez, P. (2011). The effect of interletter spacing in visual-word recognition. *Acta Psychologica*, *137*, 345–351. doi:10.1016/j.actpsy.2011.04.003

Perea, M., Panadero, V., Moret-Tatay, C., & Gómez, P. (2012). The effects of inter-letter spacing in visual-word recognition: Evidence with young normal readers and developmental dyslexics. *Learning and Instruction*, *22*, 420–430. doi:10.1016/j.learninstruc.2012.04.001

Quené, H., & van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: A tutorial. *Speech Communication*, *43*, 103–121. doi:10.1016/j.specom.2004.02.004

Rasbash, J., Steele, F., Browne, W. J., & Goldstein, H. (2008). *A User's guide to MLwiN version 2.10*. Bristol: Centre for Multilevel Modelling, University of Bristol.

Reynolds, L., & Walker, S. (2004). "You can't see what the words say": Word spacing and letter spacing in children's reading books. *Journal of Research in Reading*, *27*, 87–98. doi:10.1111/j.1467-9817.2004.00216.x

Risko, E. F., Lanthier, S. N., & Besner, D. (2011). Basic processes in reading: The effect of interletter spacing. *Journal of Experimental Pscyhology: Learning, Memory, and Cognition*, *37*, 1449–1457. doi:10.1037/a0024332

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime: User's guide*. Pittsburgh, PA: Psychology Software Inc.

Schneps, M. H., Thomson, J. M., Sonnert, G., Pomplun, M., Chen, C., & Heffner-Wong, A. (2013). Shorter lines facilitate reading in those who struggle. *PLoS One*, *8*, e71161. doi:10.1371/journal.pone.0071161

Slattery, T. J., & Rayner, K. (2013). Effects of intraword and interword spacing on eye movements during reading: Exploring the optimal use of space in a line of text. *Attention, Perception, & Psychophysics*, *75*, 1275–1292. doi:10.3758/s13414-013-0463-8

Spinelli, D., de Luca, M., Judica, A., & Zoccolotti, P. (2002). Crowding effects on word identification in developmental dyslexia. *Cortex*, *38*, 179–200. doi:10.1016/S0010-9452(08)70649-X

Tai, Y. C., Sheedy, J. E., & Hayes, J. (2006). Effect of letter spacing on eye movements and reading performance. Proceedings of the 2006 symposium on Eye tracking research & applications, San Diego, California, 27–29 March 2006.

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades?. *Journal of Child Psychology and Psychiatry*, *45*, 2–40. doi:10.1046/j.0021-9630.2003.00305.x

Vinckier, F., Qiao, E., Pallier, C., Dehaene, S., & Cohen, L. (2011). The impact of letter spacing on reading: A test of the bigram coding hypothesis. *Journal of Vision*, *11*, 1–21. doi:10.1167/11.6.8

Yu, D., Cheung, S. H., Legge, G. E., & Chung, S. T. L. (2007). Effect of letter spacing on visual span and reading speed. *Journal of Vision*, *7*, 1–10. doi:10.1167/7.2.2

Zorzi, M., Barbiero, C., Facoetti, A., Lonciari, L., Carrozzi, M., Montico, M., … , Ziegler, J. C. (2012). Extra-large letter spacing improves reading in dyslexia. *Proceeding of the National Academy of Sciences of the United States of America*, *109*, 11455–11459. doi:10.1073/pnas.1205566109