Inter-word and Inter-letter spacing effects during reading revisited:

Interactions with word and font characteristics

Timothy J. Slattery[1]

Mark Yates[2]

Bernhard Angele[1]

Department of Psychology, Bournemouth University[1]

Department of Psychology, University of South Alabama[2]

Correspondence to:

Timothy J. Slattery

Department of Psychology

Bournemouth University

Email: tslattery@bournemouth.ac.uk

Running Head: Text spacing effects revisited

Abstract

Despite the large number of eye movement studies conducted over the past 30+ years, relatively few have examined the influence that font characteristics have on reading. However, there has been renewed interest in one particular font characteristic, letter spacing, which has both theoretical (visual word recognition) and applied (font design) importance. Recently published results that letter spacing has a bigger impact on the reading performance of dyslexic children have perhaps garnered the most attention (Zorzi et al. 2012). Unfortunately, the effects of increased inter-letter spacing have been mixed with some authors reporting facilitation and others reporting inhibition (van den Boer & Hakvoort, 2015). We present findings from three experiments designed to resolve the seemingly inconsistent letter-spacing effects and provide clarity to researchers and font designers and researchers. The results indicate that the direction of spacing effects depend on the size of the 'default' spacing chosen by font developers. Experiment 3, found that inter-letter spacing interacts with inter-word spacing, as the required space between words depends on the amount of space used between letters. Inter-word spacing also interacted with word type as the inhibition seen with smaller inter-word spacing was evident with nouns and verbs but not with function words.

American's infatuation with the automobile is undeniable. However, two score and three years before the first Model A rolled off the Ford assembly line, President Abraham Lincoln stated that, "Writing, the art of communicating thoughts to the mind through the eye, is the great invention of the world...enabling us to converse with the dead, the absent, and the unborn, at all distances of time and space." Cars have the ability to move us from place to place, but no automobile can compare with the power of the written word which can transport us to distant galaxies and even move us through time.

Just as we have designed different automobiles to suit different driving needs, so to have we created a variety of fonts to fit different reading situations. We have flashy "*Ferrari*" fonts to grab your attention for advertisements, familiar "Toyota" fonts for school papers, and boxy "Jeep" fonts for computer programming. Within this metaphoric framework even the often maligned comic sans has its place as the circus clown car.

In recent years, consumers have marveled at the technological advances now available from car manufacturers. We can now purchase cars that parallel park themselves, and even assist drivers in avoiding crashes. These advances improve the person machine system. That is, they were born from the knowledge that the drive is a cooperative experience between the car, the road, and the driver. Research into driver preferences, biases, and abilities was necessary to best identify how to design a harmonious system. The designing of fonts is no different. Reading involves a complex interplay between the text display, the language content, and the reader. Therefore, designing fonts to improve the reading experience requires a thorough understanding of reading itself (Slattery, 2016).

This paper examines the optimization of one aspect of font design, horizontal text spacing, to determine how spacing optimization is influenced by other font characteristics and language content. For instance, serif fonts may require more space between letters than san serif fonts. We examine the possibility of such interdependence in Experiments 2 and 3. Additionally, text quality can interact with a word's lexical frequency which is a key variable in the study of reading (Norris, 1984; Slattery & Rayner, 2010; Yap & Balota, 2007). We examine the possibility of that spacing interacts with and lexical variables in Experiments 1 and 3.

Recently, there has been increased interest in the effects of text spacing on normal reading. Interestingly, this increased interest spans a number of different subfields—vision research (Blackmore-Wright, Georgeson, & Anderson, 2013; Chung, 2002, 2004; McLeish, 2007; Song, Levi, & Pelli, 2014; Yu, Cheung, Legge, & Chung, 2007), psycholinguistics (Cohen, Dehaene, Vinckier, Jorbet, Montavont, 2008; Perea & Gomez, 2012a, 2012b; Perea, Moret-Tatay, & Gomez, 2011; Risko, Lanthier, & Besner, 2011; Slattery & Rayner, 2013), and typography (Arditi, 2004; Arditi & Cho 2005; Reynolds & Walker, 2004).

Despite the increase in research examining the influence of text spacing on reading and reading related tasks, a lack of consistent results across studies has prevented font designers from benefitting from this research. In a recent review of the literature on letter spacing effects, van den Boer and Hakvoort (2015) examined 20 studies across 18 publications and concluded that 'default' spacing was optimal for reading[1]. Their conclusion was based partly on the observation that, of the studies examining increases to letter spacing, some reported facilitation of word processing (10), some reported inhibition (6), and others reported null effects (6)[2]. They also conducted their own word naming experiment and found inhibition for decreased letter spacing but null effects of increased spacing. However, we believe that there are well-founded theoretical reasons for the apparent discrepancies in the literature on the effects of increased horizontal text spacing. We propose that three main factors contribute to these apparent discrepancies:

1. van den Boer and Hakvoort's direct comparison between large and small spacing increases relative to a "default".
2. The incorrect belief that there exists a single "default" level of letter spacing.
3. A failure of researchers to differentiate inter-letter spacing and inter-word spacing.

We will address these issues before moving on to our current empirical investigations, which we believe will shed considerable light on the effects of spacing, to the benefit of both theoretical reading researchers and typographers.

1. Large versus small spacing manipulations

The titles of the research articles themselves do not make it clear, but there exist two very different letter spacing manipulations: large and small. Where to draw the line between these is addressed in this section. A rough mapping suggests adding less than a full character space is "small" while adding a full character space or more is large.

1.1 Large Letter-spacing manipulations

Large increases in letter spacing can disrupt word identification (Cohen, et al., 2008; Risko, et al., 2011; Vinckier, Qiao, Pallier, Dehaene, & Cohen, 2011). However, just how much space is too much? Legge, Pelli, Rubin, and Schleske (1985) described the relation between letter size and reading rate. Their reading rate curve indicated a decline in reading rate for letter widths greater than 0.3° of visual angle. This is approximately equal to the size of a single character in most eye tracking studies of reading. However, it should be noted that Legge et al. (1985) manipulated character width and not spacing.

Paterson and Jordan (2010) examined the impact of letter spacing in an eye movement study of adult sentence reading, using the Courier New font. They found that reading times were significantly longer when an e x t r a s p a c e w a s a d d e d b e t w e e n e a c h l e t t e r a n d w o r d[3]. Additionally, they found that the time cost was larger for low frequency words than for high frequency words. They concluded that adding a full space in this way disrupted normal word processing.

As it turns out, the majority of the studies examined by van den Boer and Hakvoort (2015), which yielded inhibition to word processing from increased letter spacing, used letter spacing that was as large or larger than that of Paterson and Jordan (2010). While this might seem strange from the point of view of font development, optimization was NOT the focus of these studies. Instead, these studies were designed to test aspects of the human visual system and/or computational models word recognition. For instance, Cohen et al. (2008) used increased spacing (from 1 to 5 additional spaces) to examine processing differences between the dorsal and ventral visual pathways in the brain. They hypothesized that the ventral pathway is responsible for the fast efficient reading characterized by parallel letter processing. However, when reading degraded or unusually displayed text (e.g. vertically aligned), the letters within a word are processed serially with the dorsal pathway. They reported significant inhibition to word recognition for these large spacing additions. Moreover, they reported significant word length effects for conditions with 2 or more additional spaces between letters. Cohen et al. (2008) argued that the inhibition seen for increased inter-letter spacing with the Courier font was attributable to a shift from parallel to serial letter processing. They reasoned that dual process models such as the Dual Route Cascade Model (DRC; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) and Connectionist Dual Process Model (CDP+; Perry, Ziegler, & Zorzi, 2007) could simulate the spacing effect by increasing reliance on the serial sublexical pathway when spacing is increased. This idea was tested by Risko et al. (2011) who showed that large increases to spacing result in slower, serial letter processing. However, the effects could not be attributed to the sublexical pathway assumed by the models. Instead, model simulations revealed a spacing effect on words, but not on nonwords. Risko et al. failed to find this interaction with participants. They suggest that the serial processing induced by large spacing results from an earlier stage of processing that is common to both words and nonwords (e.g., letter processing). The inability of Risko et al. to simulate the inhibitory effects of large spacing with traditional computational models of word recognition highlights a deficiency in their front end—they are too simplistic. To address this, more recent modeling work (SOLAR, Davis, 2010; SERIOL, Whitney, 2001) has sought to explain the initial orthographic processing of words. A thorough treatment of these models is beyond the scope of the current paper. However, the fact that spacing influences word recognition latencies provides an important constraint on the future development of these models.

It bears repeating that much of the research utilizing large additions to spacing were NOT attempting to optimize fonts in any way. Instead, these researchers were using large spacing increases to examine word recognition when letters are processed serially. The take home message for the current work is that increases larger than a single character will almost certainly be larger than optimal for typical adult readers.

1.2 Small Letter-spacing manipulations

Research exploring the optimization of inter-letter spacing have generally altered spacing by values less than a full character. Perea, et al. (2011), examined the impact that a slight increase of inter-letter spacing had on Spanish word recognition in the lexical decision task (LDT). They used the Times New Roman font with either default spacing or a small amount additional spacing (+1.2 pts). Across two LDT experiments, they found reaction time (RT) benefits with the small additional space. Moreover, the benefits they found for increased spacing did not interact with word frequency (Experiment 1) or word length (Experiment 2). The lack of such interactions suggests that spacing benefits occur at early encoding stages of word recognition.

The benefits for small increases to inter-letter spacing with Spanish words have since been replicated using a parametric experimental design. Perea & Gomez, (2012a) varied inter-letter spacing from -0.5 points to +1.5 points and found a linear decrease in RT. Further, using the diffusion model (Ratcliff, 1978; Ratcliff, Gomez, & McKoon, 2004), they were able to simulate the effect of spacing as an early encoding process.

Individual reading ability may also influence the optimal amount of inter-letter spacing. Perea, Panadero, Moret-Tatay, and Gomez (2012) compared the spacing benefits for adults, and children with and without dyslexia. They used both the LDT and eye movement recording of natural sentence reading. While all groups benefitted from slightly increased inter-letter spacing, the dyslexic group exhibited significantly larger benefits (see also Spinelli, de Luca, Judica, & Zoccolotti, 2002 for similar effects in Italian).

2. Default Spacing

While studies exploring the effects of spacing usually use a “default” spaced control condition, fonts differ widely in their amount of “default” inter-letter space (see Table 1). The decision on how much space to use as a default is at the typographer’s discretion. However, some word processing software (e.g. Microsoft Word) allows users to adjust the spacing of fonts. This suggests that some in the typographical community believe optimal inter-letter spacing may vary across readers and fonts. For instance, the default spacing values for serif fonts tend to be slightly wider than those for sans-serif fonts, but for a given reader this slightly wider default may still not be enough space while for others it may be too much. In fact, as Perea et al. (2011) noted, the results of studies that use subtle manipulations of inter-letter spacing are somewhat inconsistent, which may have been due in part to the use of different fonts across the studies.

Table 1. Examples of “default” spacing from the two most used fonts in eye movement research.

Font	Sentence
Times New Roman	The quick brown fox jumped over the lazy dog.
Courier New	The quick brown fox jumped over the lazy dog.

Therefore, the effect of adjusted spacing (smaller or larger), relative to a “default” should depend on how wide the “default” is. For wide spaced fonts like Courier New, increased spacing should yield longer reaction times (i.e. lower reading rates). However, for narrowly spaced fonts like Times New Roman increased spacing should yield shorter reaction time (i.e. higher reading rates). That is, whether an increase in inter-letter spacing results in faster or slower reading depends on where the font’s default spacing lies relative to an optimal spacing value.

3. Inter-letter Spacing vs. Inter-word Spacing

Being able to recognize individual words is of course vital to reading and lexical (word) processing is often viewed as the engine that drives the eyes while we read (Morrison & Inhoff, 1981; Rayner & Pollatsek, 1989; Rayner, Pollatsek, Ashby, & Clifton, 2012). However, we rarely read words in isolation, and when we are forced to do so via RSVP (rapid serial visual presentation[4]) there are sizable decrements to comprehension (Schotter, Tran, & Rayner, 2014). Thus, while single word reading tasks offer researchers a simplified theater in which to examine many of the processes related to reading, other important processes are not allowed to play a role. For instance, letter strings in an LDT experiment do not need to be segmented into individual word units nor do they need to be parsed into a meaningful syntactic phrase.

When reading sentences, a necessary starting point for word recognition is to determine the beginning and ending of words. In English and many other alphabetic languages, word boundaries are indicated by additional space between letters belonging to different words (inter-word spaces). The importance of inter-word spaces reading English is evident from studies that remove these spaces. There is a substantial reduction in reading rate for English text when inter-word spaces are removed (McGowan, White, & Paterson, 2015; Morris, Rayner, & Pollatsek, 1990; Perea & Acha, 2009; Pollatsek & Rayner, 1982; Rayner, Fischer, & Pollatsek, 1998; Rayner, Yang, Schuett, & Slattery, 2013; Sheridan, Rayner, & Reingold, 2013; Sheridan, Reichle, & Reingold, 2016). Additionally, this reduction in reading rate associated with the absence of inter-word spaces interacts with word frequency, being larger for low frequency words than high frequency words, suggesting that the inhibition occurs during word recognition. As with inter-letter space, there is variability in the amount of default inter-word space used across different fonts (see again Table 1). Therefore, we may anticipate that the relative ease of word segmentation processes will differ from font to font based on the amount of inter-word space they use. Fonts with larger inter-word space relative to inter-letter space should facilitate word segmentation compared to fonts with small inter-word space relative to inter-letter space.

A growing number of studies have explored the influence of the horizontal text spacing using sentence or passage reading (Blackmore-Wright et al., 2013; McLeish, 2007; Paterson &

Jordan, 2010; Perea & Gomez, 2012b; Perea et al., 2012; Reynolds & Walker, 2004; Slattery & Rayner, 2013; Zorzi et al., 2012). The majority of these studies report that increased spacing benefits reading. Perea and Gomez (2012b) and Perea et al. (2012) found faster reading rates with subtle increases in inter-letter spacing using the proportional width font Times New Roman. Additionally, Perea et al. (2012) found the benefits of increased inter-letter spacing were greater for readers with developmental dyslexia than for those without developmental dyslexia (see also Zorzi et al., 2012). However, it is important to note that, although these studies report inter-letter spacing effects, they manipulated inter-character spacing. That is, these studies examined conditions where additional space was either added or removed between all characters of text—including the inter-word space character. Consequently, inter-letter spacing was confounded with inter-word spacing. Therefore, they have no ability to separate the influence of one type of spacing from the other.

In contrast to the benefits seen with increases to inter-letter spacing reported by the above studies, Paterson and Jordan (2010) found a detrimental effect of increased spacing on eye movements when using the fixed width font Courier. However, in their experiment the smallest addition to inter-letter spacing added an extra space b e t w e e n e a c h l e t t e r and this most likely disrupted the overall integrity of the words in the sentences resulting in more serial letter processing. In fact, Paterson and Jordan found that the effect of word frequency was larger for increased spacing conditions relative to the default spacing control condition. From this, they argue that the increased spacing interfered with normal word processing. Paterson and Jordan also manipulated both inter-letter and inter-word spacing across their conditions but not in a fully factorial experimental design and therefore could not assess the individual contributions of each type of space nor test for interactions between inter-word and inter-letter spacing.

Slattery and Rayner (2013) also manipulated inter-letter and inter-word spacing in two eye movement studies of sentence reading. Their first experiment was similar to those of Perea and colleagues as they manipulated spacing between all characters. However, Slattery and Rayner compared two proportional width fonts: Cambria and Times New Roman. They found that adding or removing space between the characters of these fonts increased total reading times. In their second experiment, they used a novel manipulation which reduced the inter-letter spacing of words and added this space to the end of the word thereby increasing inter-word spacing. This condition was then compared to the default spaced condition. In this second experiment, they also compared these spacing effects for the proportional width Georgia font and the fixed width Consolas font. They found that the adjusted spacing condition yielded shorter gaze durations and that this benefit of increased inter-word/decreased inter-letter spacing was largely limited to the Georgia font. Blackmore-Wright et al. (2013) also report benefits to reading rate with increased inter-word spaces for readers with macular disease.

Current Studies

The current studies were designed to accomplish a number of goals. First, we wanted to assess the replicability of inter-letter spacing benefits in single word tasks which is currently complicated by the use of different fonts across published studies. If the reported spacing effects represent Type I errors, then we should be unlikely to replicate them under similar conditions. However, if the effect of altering inter-letter spacing depends on the size of the 'default' spacing of the font used, then these effects may appear more stable and reproducible when taking font into consideration. We address this reproducibility issue in Experiment 1 using the same font (Times New Roman) and a highly similar manipulation as Perea and colleagues.

Second, we wanted to explore how changes in spacing may interact with font characteristics (e.g. the presence or absence of serifs). Such interactions might help explain why spacing studies have yielded inconsistent results. Related to this goal we wanted to develop a measure of spacing capable of capturing the differences that exist between the 'default' spacing values of various fonts.

Third, we wanted to investigate inter-word and inter-letter spacing within a fully factorial experimental design so that the independent contributions of both types of spacing and their potential interaction could be assessed. We addressed this issue in Experiment 3 with an eye movement study of sentence reading.

Finally, we wanted to explore the possibility that the inter-letter spacing effect might be different for different categories of words. There is evidence that verbs are more difficult to process than nouns. Nouns are recalled better than verbs (Engelkamp, Zimmer, & Mohr, 1990; Helstrup, 1989; Reynolds & Flagg, 1976) and are fixated for less time than verbs even when controlling for length and word frequency (Rayner, 1977). If spacing facilitates word recognition, we might expect that the facilitation would be greater for the more difficult words (i.e. verbs). Also, function words (e.g., for, and, from), which tend to be shorter and much more frequent than content words (nouns and verbs), receive fewer and shorter fixations during reading than content words (Rayner, 2009). Therefore, function words might not benefit from increased spacing to the same extent as nouns and verbs. We explore word type (verb, noun, function) effects in Experiment 1 and Experiment 3.

Experiment 1

Experiment 1 was a replication and extension of the LDT experiments of Perea et al. (2012). Similar to Perea et al., we used the Times New Roman font and varied the inter-letter spacing of words and non-words. However, we used English rather than Spanish stimuli. English is considered a deep orthography as the same letter units can map onto multiple phonological units. Spanish on the other hand has a shallow orthography with largely consistent mappings to phonology. It is possible that letter spacing will matter less in an environment where a letters pronunciation is derived in large part from the context provided by nearby letters—as in English. Additionally, we included five levels of spacing rather than just two. If there exists an optimal level of inter-letter spacing, it may appear as a higher order (e.g. quadratic) trend. With only two levels of spacing, it is impossible to assess such higher order effects. Finally, we extended upon Perea et al. (2012) by examining how spacing influences different word types (nouns, verbs, and function words).

Methods

Participants: Twenty-four members of the Psychology Department subject pool at the University of South Alabama participated in the study. All participants self-identified as native speakers of American English, were naïve to the purpose of the experiment, and had normal or corrected to normal vision in at least one eye.

Apparatus: Eye gaze position was sampled every millisecond using an SR Research Eyelink 1000 eye-tracker. Eye movement data were only collected from the right eye, though viewing was binocular. Stimuli were displayed on a 24 inch BenQ gaming LCD monitor with a 120 Hz refresh rate and 1 ms response time. Participants were seated 60 cm from the monitor. Responses were collected with a VPixx brand five button response box.

Materials: Words for Experiment 1 were selected from the English Lexicon Project's full database. We selected 180 verbs, 180 nouns, and 90 functions words for a total of 450 words. The characteristics of these words can be found in Table 2. For each word type, we selected from a wide range of word lengths and word frequencies. For the nouns and verbs, 30 words were selected from each word length between 3 and 8 letters. Within each word length words were chosen to represent a range from low to high frequency. A similar approach was taken for the function words. However, due to the scarcity of such words relative to verbs and nouns, there were two differences in the selection of function words. First, function words were not selected equally over different word lengths. Though the range of length for function words still extended from 3 to 8 letters, the majority of function words are short (3-4 letters). This difference in length between function words and the nouns and verbs was statistically significant, $t(268) = 4.03$, $p < .001$. Second, function words tended to range from high to very high in frequency. This lead the function words to be higher in frequency than the nouns, $t(268) = 19.7$, $p < 0.001$, and the verbs, $t(268) = 15.3$, $p < 0.001$. However, there was no significant difference in word frequency between nouns and verbs, $t(368) = 1.4$, $p > 0.15$.

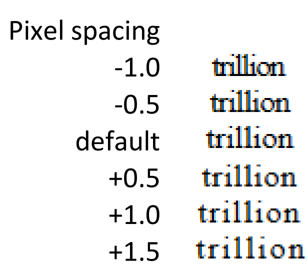
Table 2. Average Word Properties

Type	Length in Letters	Log Frequency
Noun	5.5 (1.7)	1.7 (1.8)
Verb	5.5 (1.7)	2.0 (2.3)
Function	4.7 (1.4)	6.2 (2.1)

Note: standard deviations appear in parentheses.

Non-word stimuli were chosen by selecting a word with similar characteristics to the word stimuli and changing one letter to yield a pronounceable non-word. The spacing of the word and non-word stimuli was manipulated using Microsoft's ClearType sub-pixel rendering (for details see Slattery & Rayner, 2013). Six levels of inter-letter spacing were utilized: 1 pixel removed, 0.5 pixels removed, default, 0.5 pixels added, 1 pixel added, 1.5 pixels added (see figure 1 below). For reference, the normal spacing and 1 pixel added conditions are closest to those used by Perea et al. (2012). Latin square counterbalancing was used to create six lists that crossed words and spacing conditions.

Figure 1. Example Stimulus



Procedure: Participants were familiarized with the LDT procedure in 10 practice trials. Similar to the standard LDT, target letter strings were presented in the center of a computer monitor and participants were required to decide if the string was a word or not via a button press.

However, the LDT was modified in that eye movements were monitored during the task. Participants were calibrated using a full screen 9-point procedure. Prior to the presentation of the text sting, a fixation cross appeared randomly in one of 6 locations around and equidistant (1° of visual angle) from the center of the computer monitor. Participants were to first fixate this cross. Upon doing so, the cross would disappear and the string would appear in the center of the monitor. Participants would then fixate the string and make their decision via button press. The procedure took 1 hour and 45 minutes to complete including breaks, which were given frequently due to the repetitive nature of the task.

Results

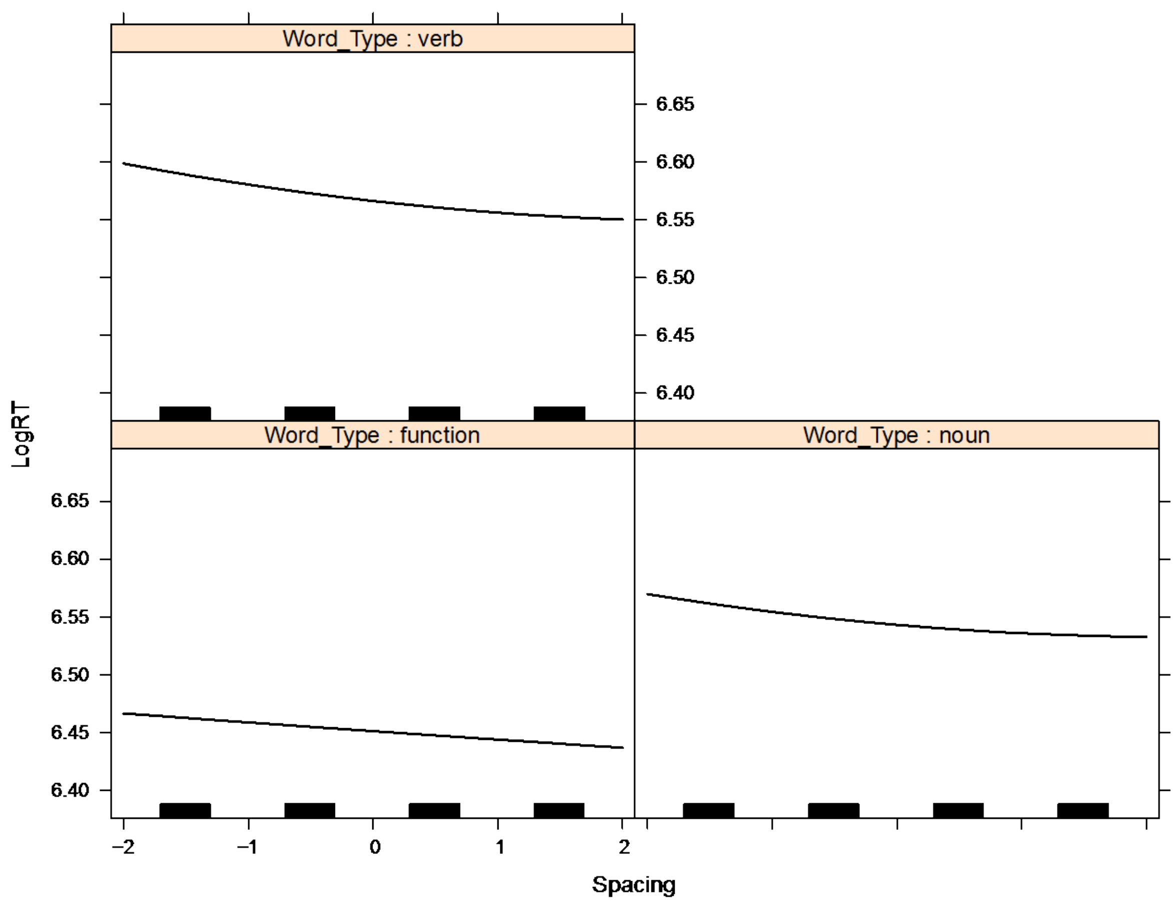
Of the 450 words, 13 were judged incorrectly by more than 50% of participants. These words were excluded from analysis. An additional 5.2% of the word trials were excluded from analysis due to: eye blinks, trials where the initial saccade from the fixation cross to the target took longer than 350 ms, problems with stimulus presentation, or anticipatory button presses. Finally, trials with RTs greater than 3000 ms (<.1%) or less than 250 ms (<.1%) were excluded from analysis. Participants were accurate on 93.5% of word decisions. For analyses, we used the *lmer* function from the *lme4* package (version 1.1-11; Bates, Maechler, Bolker, & Walker, 2015) within the R Environment for Statistical Computing (R Core Team, 2015) to fit linear mixed models (LMMs) of reaction time (RT) and generalized linear mixed models (GLMMs) using a logit link for accuracy data. For all statistical models, we present effect coefficients (b), standard errors (SE), and t-values (t) or z-values (for GLMMs). Determining the degrees of freedom for t-statistics estimated by LMMs is unclear which makes estimating exact p-values difficult

(Baayen, Davidson, & Bates, 2008). However, with a large number of subjects and items, and relatively few fixed and random effects to estimate (as in the current studies), the distribution of the t-values estimated by the LMMs approximates the normal distribution. Therefore, we used the two-tailed criterion $|t| \geq 1.96$ corresponding to a significance test at the .05 α-level. The z-values from the GLMMs can be interpreted similarly.

The GLMM for accuracy predicted participant decisions to word stimuli from the spacing condition (as a centered continuous predictor) and the word type (as a discrete factor, using successive difference contrast coding, Venables & Ripley, 2002). There was a significant effect of word type as function words were responded to more accurately than nouns and verbs, b = 1.369, SE = .286, z = 4.792, but responses to nouns were not significantly more accurate than those to verbs, b = .216, SE = .196, z = -1.10. There was no significant effect of the spacing manipulation on response accuracy, b = -.001, SE = .034, z = -0.215, nor was there a significant interaction between spacing and word type, b = .121, SE = .095, z = 1.282; b = -0.025, SE = .055, z = -0.456.

Accurate word responses were analyzed with an LMM that predicted log-transformed reaction time (RT) from the inter-letter spacing condition (both linear and quadratic) and word type. All predictors were centered around their mean. The LMM included fixed effects for all main effects and interactions. The model also included random intercepts for subjects and items as well as random slopes for the main effects. Inter-letter spacing significantly predicted log RTs but the effect was strictly linear, b = -1.602, SE = .326, t = -4.91: log RT decreased as inter-letter spacing increased (see Figure 2). The quadratic component for inter-letter spacing did not approach significance, b = .341, SE = .259, t = 1.32. Function words were also responded to faster than nouns and verbs, b = .098, SE =.028, t = 3.43. However, noun responses were not significantly faster than verbs, b = .023, SE = .019, t = 1.24 As with the accuracy data there were no significant interactions, all ts < 1.

Figure 2. Plot of the partial effects of inter-letter spacing and word type on log RT as predicted by the LMM.



Having replicated the inter-letter spacing effects found by Perea et al., and establishing that there is no indication that this inter-letter spacing effect varies by word type (at least in an

LDT), we further investigated the possibility that spacing interacted with the lexical characteristics which varied over the stimuli in our word types. The LMM predicted the log RTs from the log word frequency, length in letters, and inter-letter spacing. All predictors were centered about their mean. The LMM included fixed effects for all main effects and interactions. The model also included random intercepts for subjects and items as well as random slopes for the main effects. The results confirmed a main effect of inter-letter spacing on log RTs with larger spacing yielding shorter times, $b = .010$, $SE = .002$, $t = -4.98$. There was also a main effect of word frequency, $b = -.037$, $SE = .004$, $t = -8.40$, as higher frequency words were responded to faster than lower frequency words. The effect of the number of letters in the word failed to reach significance, $b = .008$, $SE = .004$, $t = 1.94$, however, there was an interaction between word frequency and word length, $b = -.006$, $SE = .001$, $t = -4.59$, as the effect of word frequency was stronger for longer words. Crucially though, there were no interactions with inter-letter spacing, $|ts| < 1$.

Discussion Experiment 1

In Experiment 1, we were able to successfully replicate the facilitative effects of increased inter-letter spacing reported by Perea et al. (2012) using the same proportional width serif font, Times New Roman, which has relatively small default inter-letter spacing. Additionally, function words were responded to faster than the nouns and verbs. This was not surprising given they were higher in frequency and shorter in length. This indicates that the function words in our stimulus set can indeed be assumed to be very easy to process and therefore useful for testing for potential interactions between lexical difficulty and spacing.

However, despite strong main effects of word type and spacing, we found no evidence of an interaction between these variables. The spacing effect was statistically similar for nouns, verbs, and function words and was not influenced by word length or word frequency.

Having shown that the influence of inter-letter spacing on LDT latencies does not depend on the lexical properties of words, we next wanted to explore the possibility that inter-letter spacing effects are interrelated to font characteristics. Fonts differ in a number of ways including default inter-letter spacing. Prior research has shown that increasing the spacing between letters within a word beyond a tipping point results in less efficient word processing (Cohen et al., 2008; Legge et al., 1985; Risko et al., 2011). Some font's default spacing may be closer to this tipping point than others (or even on the other side of the tipping point), and the tipping point may depend on other font variables such as the presence or absence of serifs and whether the font uses fixed or proportional width letters.

Experiment 2

Experiment 2 explored the possibility that intra-letter spacing effects may differ across fonts. We used the same words and procedures as in Experiment 1 but with six fonts that differed in a variety of ways, which allowed us to examine two specific font characteristics. These characteristics were: (a) whether the font had serifs or not, and (b) whether the font used fixed or proportional width letters. Three of the fonts were sans serif fonts (Calibri, Verdana, and Consolas) while three were serif fonts (Cambria, Georgia, Courier New). Additionally, two fonts (one serif and one sans-serif) were fixed width while the remaining were proportional width (see table 3). The majority of text that readers encounter in their daily lives will be in a proportional width font where each letter can be of a different width. However, letters in fixed width fonts, like Courier New are popular for computer programming and are commonly used in psycholinguistic research (Slattery, 2016) due to their uniform width. Fixed width fonts have advantages in psycholinguistic research, as words with the same number of letters will necessarily be the same size (i.e. occupy the same horizontal extent). However, fixed width fonts will often appear to have larger and less uniform inter-letter spacing. Compare the Cambria version of the name "William" with the Courier New version "William". The letters Cambria all appear to unite into a cohesive whole. However, there is more space between letters in Courier New and the 'i' appears more separated from the other letters while the 'a' and 'm' appear quite close.

Method

Participants: The participants were drawn from the same pool as in Experiment 1. However, due to the addition of the additional independent variable (font) we recruited 66 participants for Experiment 2. All participants had normal or corrected to normal vision in at least one eye.

Apparatus & Procedure: All the details were the same as in Experiment 1.

Materials: The words were the same as those used in Experiment 1. The characteristics of the 6 fonts we chose appear in table 3 below. With the addition of the font variable, we reduced the number of spacing levels tested to three (remove a pixel, normal, add a pixel). As with Experiment 1, all independent variables were fully counterbalanced within participants.

The use of different fonts in Experiment 2 necessitated a font independent measure of default letter spacing for comparisons. Characters in text are composed of the visible letter pixels as well as space to either side. When two characters are printed next to each other in a word their inter-letter space consists of the space to the right of the left letter plus the space to the left of the right letter. It has been argued that visual crowding depends on center-to-center letter distance (character width) rather than letter size or spacing (Arditi, Knoblauch, & Grunwald, 1990; Bricolo, Salvi, Martelli, Arduino, & Daini, 2015; Strasburger, Harvey, & Rentschler, 1991). So, wider letters with less spacing should be just as crowded as thinner letters with more spacing so long as the distance from the center of one letter to the next is the same in both cases. We calculated this measure in visual angle for 10 randomly chosen words from our stimulus list (see character width in Table 3). In addition to character width, we also estimated the letter width and spacing from these 10 randomly chosen words with a new measure of inter-letter spacing; e-bar space (see Figure 3). We calculated e-bar space by counting the space pixels between letters at the height of the e-bar (the horizontal line of the lowercase ‘e’) for each font, then converted this into a visual angle. For Calibri (Figure 3—left) there were three space pixels between the ‘t’ and ‘a’, one between the ‘a’ and ‘k’ and three between the ‘k’ and ‘e’. For Consolas, (Figure 3—right) there were five space pixels between the ‘t’ and ‘a’, two between the ‘a’ and ‘k’ and four between the ‘k’ and ‘e’. Letter width was defined as character width minus spacing. Finally, in order to represent font spacing in a single measure for comparison between fonts, we calculated space as a percentage of letter width. As can be seen in Table 3, the fixed width fonts are composed of a larger percentage of “default” inter-letter space than the proportional width fonts, t(9) = 3.397, $p < .05$. Additionally, the serif fonts are composed of more "default" space than the san serif fonts, t(9) = 4.097, $p < .05$.

Figure 3. Calibri and Consolas example for the word "take" with contrast adjusted to make the spacing pixels more apparent and shown with pixel grid overlay.

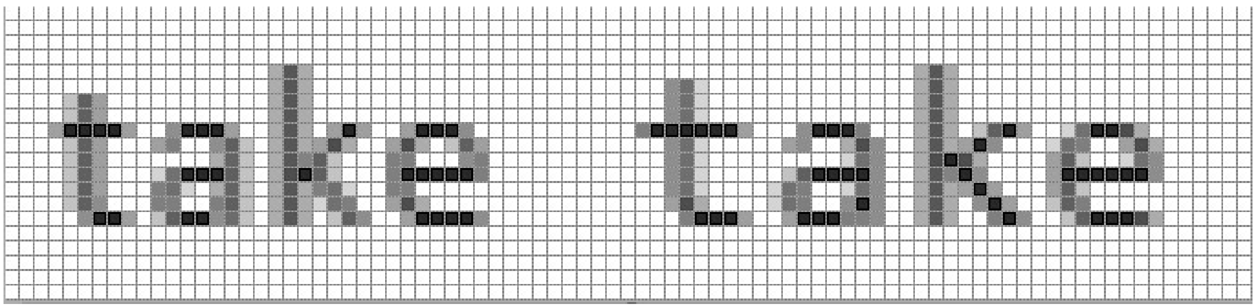


Table 3. Font Characteristics for Experiment 2.

Font	Fixed Width	Serif	Character Width	Letter Width	Space Width	Percent Space
Calibri	No	No	.201	.166	.035	22.63
Cambria	No	Yes	.215	.171	.044	30.04
Consolas	Yes	No*	.240	.175	.066	40.70
Courier New	Yes	Yes	.271	.186	.085	50.00
Georgia	No	Yes	.228	.181	.047	28.42
Verdana	No	No	.246	.202	.044	23.19
Times NR	No	Yes	.203	.163	.040	26.52

Note: Widths in visual angle. Percent space is calculated as (space width / letter width) x 100). Times New Roman shown for comparison with Exp 1. * Only the "i" and lower case "l" have serifs in Consolas.

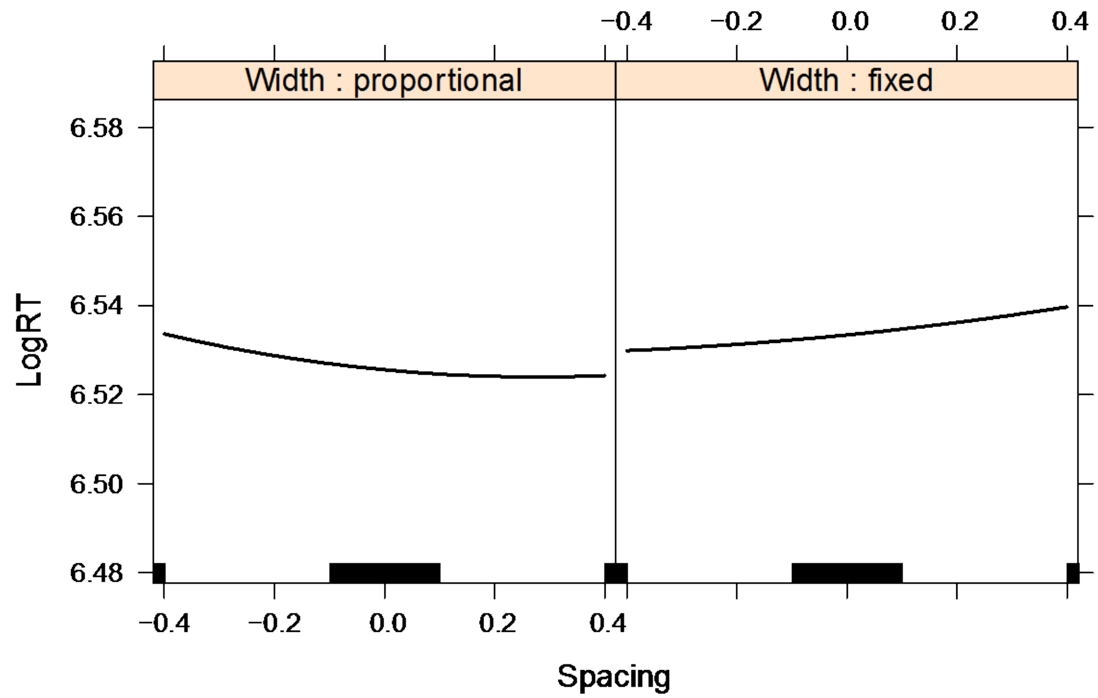
Results

As with Experiment 1, there were thirteen words out of 450 that were answered incorrectly by more than 50% of participants (ten of the words were the same as in Experiment 1). These words were excluded from further analysis. We used the same trial exclusion criteria from Experiment 1, which resulted in the removal of 4.75% of the word responses. After exclusions, participants were correct on 93.8% of word trials. We assessed the impact of our experimental manipulations on accuracy with a GLMM that predicted participant response from the fixed effects of spacing, serifs, font width, and word type, as well as the interactions between spacing and each of the font characteristic variables. The model included random effects for participants only as models including random effects for items or random slopes failed to converge (likely due to the small number of errors). Accuracy was again higher for the function words (97%) than for the nouns (94%), or verbs (92%), $b = -.915$, $SE = .092$, $z = -9.982$. Additionally, in Experiment 2, the accuracy for nouns was significantly higher than for verbs, $b = -.197$, $.054$, $z = -3.659$. However, accuracy was not significantly impacted by the spacing, serif, or font width variables, nor were there any significant interactions, $|zs| < 1$.

To test our two font characteristic hypotheses, we fit an LMM to log reaction time (RT). The model included the fixed effects predictors of inter-letter spacing (linear and quadratic trends) and the two font characteristic variables (serifs: presence or absence, and width type: proportional or fixed) as well as the interaction term between inter-letter spacing and each of these font characteristics. We included log word frequency, word length, and their interaction which had been significant in Experiment 1. The model included random intercepts for subjects and items as well as random slopes for inter-letter spacing and font width. Models including random slopes for the serif variable failed to converge. All predictors were centered about their mean.

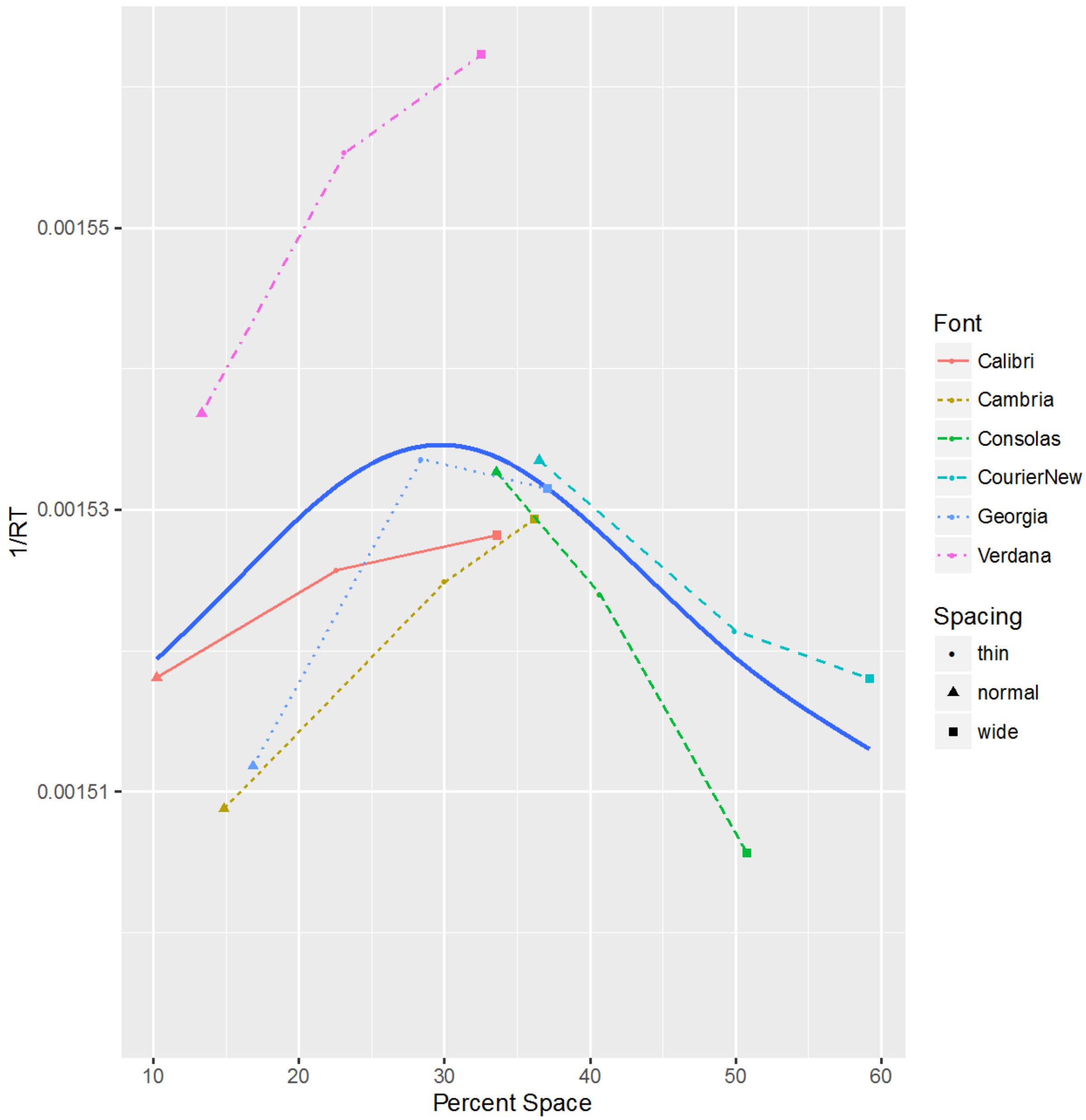
As in Experiment 1, there was a main effect of word frequency, b = -.037, SE = .002, t = -18.51, and word length, b = .010, SE = .003, t = 3.83, as RTs were shorter for higher frequency words and shorter words. Additionally, there was as an interaction between word length and word frequency, b = -.005, SE = .001, t = -4.33. This interaction, which indicates that the effect of word frequency is smaller for short than for long words, was nearly identical in Experiment 1. Unlike Experiment 1, there was no main effect of the spacing manipulation in Experiment 2: linear b = .032, SE = .252, t = .13; quadratic b = .279, SE = .245, t = 1.14. Neither the main effect of serifs nor the main effect of width type reached significance: b = .004, SE = .003, t = .129; b = .006, SE = .003, t = 1.83 respectively. However, there was a significant interaction between the linear effect of spacing and whether the font was proportional or fixed width, b = 1.543, SE = .4907, t = 3.14. Increasing the spacing between letters decreased RTs for proportional width fonts but increased RTs for fixed width fonts (see Figure 4). No other interactions approached significance, all |ts| < 1.

Figure 4. Plot of the partial effects of inter-letter spacing and font width type on log RT as predicted by the LMM.



The interaction between inter-letter spacing and the font characteristic of width (proportional vs. fixed) provides strong support for our first hypothesis. Specifically, that the large and uneven spacing of fixed width fonts would place these fonts on the right side of the reading rate peak while the small uniform space of proportional width fonts would place them on the left side of the peak. In order to examine this directly, we transformed reaction times into their reciprocal so that so that faster reading is indicated by larger values (to match the standard reading rate curve). We then calculated for each font and spacing level the amount of inter-letter space as a percentage of the width of the font's letters. Figure 5 plots reciprocal RT by the percentage of space showing both the average effect (with a smoother) as well as the data points from each font. There is a clear peak in the average effect for a spacing percentage of ~30%. Additionally, the proportional width fonts (Calibri, Cambria, Georgia, and Verdana) all lie on the left side of this peak while the fixed width fonts (Consolas, and Courier New) both lie to the right of it. To confirm that the effects shown in Figure 4 are statistically reliable, we fit an LMM to the reciprocal RT data. This model included fixed effects for inter-letter spacing percentage (linear and quadratic components), letter width in visual angle, and their interaction. We also included word frequency, word length, and their interaction, which had been significant in all earlier models. Random effects of subjects and items were included as well as random slopes for inter-letter space percentage and letter width. The results confirm the significant quadratic component of inter-letter space percentage, $b = -75.641$, $SE = 22.574$, $t = -3.35$. Additionally, there was a significant effect of letter width, $b = .484$, $SE = .127$, $t = 3.82$, as wider letters yielded faster reading. However, there was no interaction between inter-letter spacing percentage and letter width, $|ts| < 1$.

Figure 5. Reciprocal RT as a function of inter-letter spacing represented as a percentage of letter width. The solid curved line represents the average across the fonts.



Discussion Experiment 2

Experiment 2 tested two hypotheses regarding the font characteristics and inter-letter spacing. We found no evidence to support the serif hypothesis. That is, inter-letter spacing affected RT similarly for serif and san serif fonts. However, we did find evidence to support the font width (proportional vs. fixed) hypothesis in a significant interaction between inter-letter spacing and font width (proportional vs. fixed). If we take the LDT latencies as an indicator of font readability, then Experiment 2 demonstrated that inter-letter spacing can influence font readability in different ways. Bigger isn't always better. Fixed width fonts with large and less uniform default inter-letter spacing become less readable with additional inter-letter spacing. However, proportional width fonts with uniformly small spacing become more readable with added space. This also helps to explain the different inter-letter spacing effects reported in the literature. Perea and colleagues (2011, 2012) found facilitation from increased spacing using the proportional width font Times New Roman. However, Paterson and Jordan (2010), and Cohen et al. (2008) both used a fixed width font (Courier and Courier New respectively) and both reported increased processing time with increased spacing.

When the data from Experiment 2 are plotted in words per minute (WPM) as a function of the inter-letter spacing percentage for fonts and spacing conditions a peak emerges at ~30%. Plotting the data points for the individual fonts over the average plot line was illuminating. Statistical analyses confirmed two important effects. Reading rate improved as letter width increased, and as inter-letter spacing (taken as a percentage of letter width) approached 30%. However, there was no hint of an interaction between these variables. That is, the optimal inter-letter spacing percentage (~30%) was roughly the same for all fonts tested regardless of the width of their letters. Thus, Verdana with its relatively wide letters (.243° per letter) and Consolas with more narrow letters (.210° per letter) both improved as their inter-letter spacing percentage approached 30%. However, these two fonts approached the peak from different directions because of their default inter-letter spacing percentage values. Note that these two fonts have almost identical center-to-center letter distances (character widths = .295°and .289° respectively). These data argue against the use of center-to-center letter distance as a metric for determining optimal inter-letter spacing. Instead, optimal spacing in the LDT is better represented as a percentage of letter width.

Experiment 3

In the single word reading LDT, added inter-letter spacing has been shown to improve readability but only for proportional width fonts. Readability of fixed width fonts, like those typically used in psycholinguistic research, suffers with the addition of extra inter-letter space.

Single word reading tasks are simplified and fail to capture many important aspects of normal reading. One such aspect is the need to parse text into individual word units. Up until now we have only had to describe one type of spacing— inter-letter. However, in natural reading, inter-word spacing (the space between words) may be more important for readability than inter-letter spacing (Slattery & Rayner, 2013). Prior studies have shown that word segmentation is an important early process in reading (see Perea & Acha, 2009). Reading studies of alphabetic languages have shown that reading becomes much more difficult when inter-word spaces are removed from text (Morris et al., 1990; Perea & Acha, 2009; Pollatsek & Rayner, 1982; Rayner et al., 1998; Rayner et al., 2013; Sheridan et al., 2013; Sheridan et al., 2016).

While Slattery and Rayner (2013) examined the influence of both inter-letter and inter-word spacing, they did not use a fully factorial design in their study. Instead of crossing levels of inter-letter spacing with levels of inter-word spacing, they compared a condition with normal text to one with increased inter-word and decreased inter-letter spacing. They accomplished this by placing the inter-letter space they saved at the end of the word thus increasing inter-word space. Their unique manipulation allowed them to control for the visual angle of the words within the experimental texts. However, a by-product of this manipulation was that inter-word spaces were variable, being larger after longer words. Nevertheless, they found benefits for the modified spacing condition relative to normal text. This finding runs counter to the evidence from single word studies which show that increased inter-letter spacing yields faster reaction times. Slattery and Rayner also examined these spacing effects for two fonts; Georgia (proportional width) and Consolas (fixed width). They reported larger benefits of their spacing adjustment for Georgia compared to Consolas—consistent with the results of Experiment 2 and foreshadowing the results of our next experiment.

Experiment 3 manipulated inter-word and inter-letter spacing in a fully factorial design to obtain a better understanding of how these variables influence a font's readability. We used the proportional width font Calibri and the fixed width font Consolas, which were both used in Experiment 2. Given the results of Experiment 2, we should predict that, as inter-letter spacing increases, measures of reading performance should improve for Calibri but should decline for Consolas. When inter-word spaces are completely eliminated, reading rate drops significantly. Therefore, we predict that reading rate should decrease as inter-word spaces become smaller. However, we also anticipate a potential interaction between inter-letter and inter-word spacing. When inter-letter spacing is increased and inter-word spacing is reduced, the ratio of inter-word to inter-letter spacing is reduced thereby making inter-word spaces less apparent.

Method

Participants: Ninety participants were drawn from the same pool as in Experiment 1 and 2, and had normal or corrected to normal vision in at least one eye.

Apparatus: Same as Experiment 1 and 2.

Materials: The target words were the same as those used in Experiment 1 and 2. However, in Experiment 3 these words were embedded into 180 unique sentences. To do this, nouns, verbs and function words were pseudo-randomly[5] grouped together (one of each per sentence). In order to accomplish this, each function word had to be grouped with two different noun/verb pairings.

The inter-letter and inter-word spacing manipulation was accomplished with the same software used for the words in Experiment 1 and 2. Three equidistant levels of inter-letter and inter-word spacing were chosen. The middle level was equal to the default spacing for the font. Additionally, the levels were chosen such that on average across all the items, the added or removed inter-letter space would equal the added or removed inter-word space (see Figure 5). This was done to allow for specific tests of spacing under conditions in which line length was equated. So, the increased inter-letter / default inter-word spacing sentences occupied the same visual angle as the default inter-letter / increased inter-word spacing sentences. We chose one fixed width (Consolas) and one proportional width (Calibri) font for Experiment 3 (both were san serif). Each of the 180 unique sentences could be seen in one of 18 conditions created by a factorial crossing of independent variables 2 (fonts) X 3 (inter-letter spacing) X 3 (inter-word spacing). All independent variables were fully manipulated within participants. Latin square counterbalancing insured each participant saw an equal number of sentences in every condition and no sentence in more than one condition. Additionally, over all participants, every item was seen an equal number of times in each condition. An example sentence is shown in all of its conditions in Figure 6. Table 4 displays the default inter-word and inter-letter spacing in visual angle and the proportion of inter-word to inter-letter spacing for each font.

Table 4. Font Characteristics Experiment 3.

Font	Inter-letter Space	Inter-word Space	Proportion Inter-word / Inter-letter Space
Calibri	.035	.138	3.94
Consolas	.066	.305	4.62

Note: Spaces shown as degrees of visual angle.

Procedure: Sentences were presented centered vertically on the LCD monitor in a random order. Participants were calibrated using a full screen 9 point procedure. Validation errors greater than 0.3° of visual angle resulted in a repetition of the calibration procedure. Participants read the sentences silently to themselves for comprehension while their eye movements were recorded. Participants were instructed to press a button on the response box to indicate that they were finished reading the sentence. Following one third of the sentences, a two alternative comprehension question was displayed. Participants responded via button press. The procedure took approximately 1 hour and 30 minutes to complete including frequent breaks.

Figure 6. An example sentence in all 18 conditions. Calibri sentences appear first. The conditions for each font are grouped by inter-word spacing then by inter-letter spacing.

Inter-word spacing	Inter-letter spacing	
-1.8	-0.5	The hungry lions devour over a hundred pounds of meat every single day.
-1.8	default	The hungry lions devour over a hundred pounds of meat every single day.
-1.8	+0.5	The hungry lions devour over a hundred pounds of meat every single day.
default	-0.5	The hungry lions devour over a hundred pounds of meat every single day.
default	default	The hungry lions devour over a hundred pounds of meat every single day.
default	+0.5	The hungry lions devour over a hundred pounds of meat every single day.
+1.8	-0.5	The hungry lions devour over a hundred pounds of meat every single day.
+1.8	default	The hungry lions devour over a hundred pounds of meat every single day.
+1.8	+0.5	The hungry lions devour over a hundred pounds of meat every single day.
-1.8	-0.5	The hungry lions devour over a hundred pounds of meat every single day.
-1.8	default	The hungry lions devour over a hundred pounds of meat every single day.
-1.8	+0.5	The hungry lions devour over a hundred pounds of meat every single day.
default	-0.5	The hungry lions devour over a hundred pounds of meat every single day.
default	default	The hungry lions devour over a hundred pounds of meat every single day.
default	+0.5	The hungry lions devour over a hundred pounds of meat every single day.
+1.8	-0.5	The hungry lions devour over a hundred pounds of meat every single day.
+1.8	default	The hungry lions devour over a hundred pounds of meat every single day.
+1.8	+0.5	The hungry lions devour over a hundred pounds of meat every single day.

Results

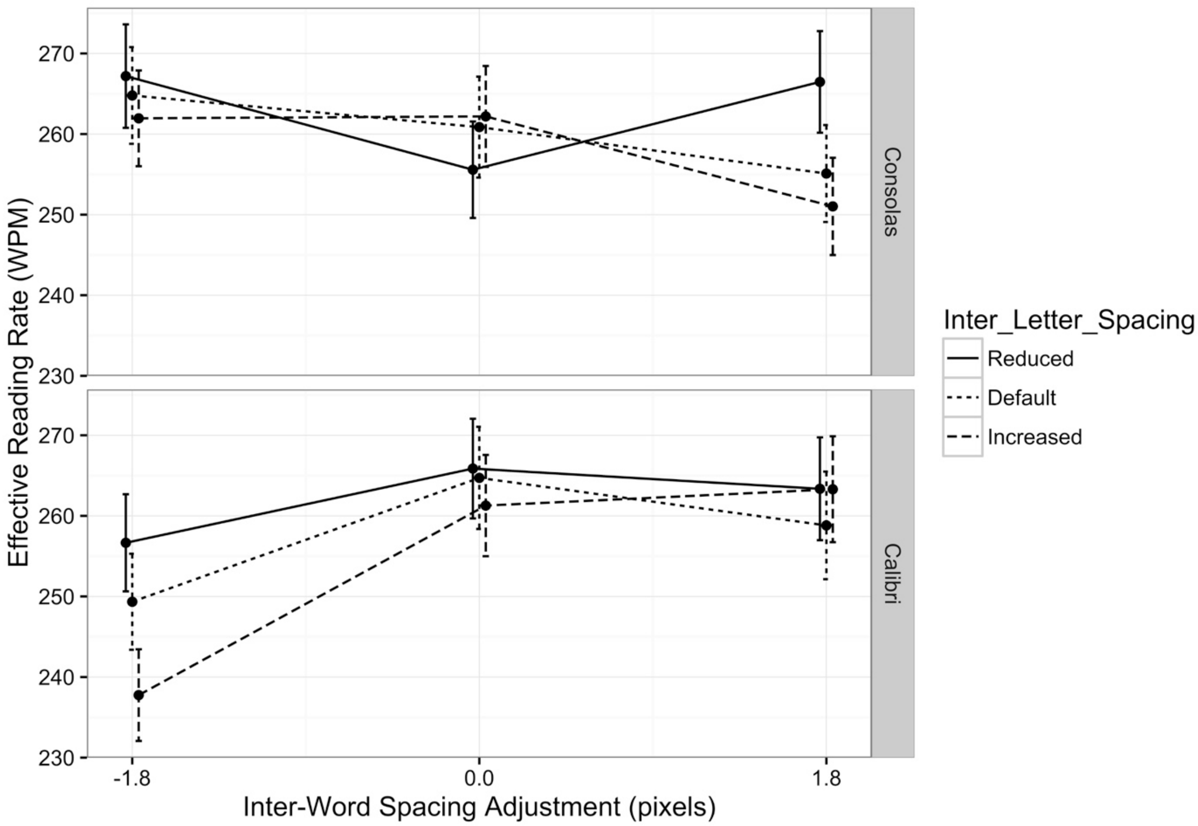
We present analysis of two dependent measures: effective reading rate, and target word gaze durations. Prior to analysis, sentence reading trials with more than 2 blinks (<1% of trials) or more than 50 fixations (<1% of trials) were excluded. Additionally, fixations below 80 milliseconds (1.4% of total fixations) were combined with a temporally adjacent fixation if the two fixations were within 10 pixels (about the size of one character) of each other (6.5% of the short fixations). Trial initial and final fixations were excluded from analysis as were fixations that were contaminated by blinks (1.0% of all fixations). Accuracy to comprehension questions was high (91.4%) and uninfluenced by the experimental variables (all Fs <1.)

Effective Reading Rate (Jackson & McClelland, 1979; Rayner, Abbott, & Plummer, 2015) compensates for speed accuracy trade-offs by multiplying a participant's WPM reading rate by their average accuracy to comprehension questions in that condition. So, if in a given condition, a participant's average reading rate was 300 WPM, and their average accuracy was 90%, their effective reading rate would be 270 WPM. Effective reading rate was analyzed with linear mixed models which included fixed effects predictors of font, inter-letter spacing, and inter-word spacing along with all their interactions. The model also included random intercepts for items and subjects as well as random subject slopes for the main effects. Both spacing variables were fit with polynomial contrasts to test for both linear and quadratic trends.

There was no main effect of font on effective reading rate, $|t| < 1$. There was also no main effect of inter-word spacing, linear trend: $b = 2.59$, $SE = 1.94$, $t = 1.34$; quadratic trend: $b = -3.05$, $SE = 2.12$, $t = -1.44$. However, there was a marginal main linear effect of inter-letter spacing, $b = -4.13$, $SE = 2.12$, $t = -1.92$, as larger inter-letter spacing was associated with slower effective reading rates.

While the main effects were largely non-existent, there were numerous important interactions (see Figure 7). Font interacted with both the linear and quadratic components of inter-word spacing, linear $b = -14.87$, $SE = 2.26$, $t = -6.58$; quadratic $b = 9.12$, $SE = 2.26$, $t = 4.04$. With the fixed width Consolas, there was little difference between the reduced and default inter-word spacing conditions but the increased inter-word spacing condition resulted in slower reading rates. The situation was largely the reverse for the proportional width Calibri. There was little difference between the default and increased inter-word spacing condition but the reduced inter-word condition resulted in slower reading rates.

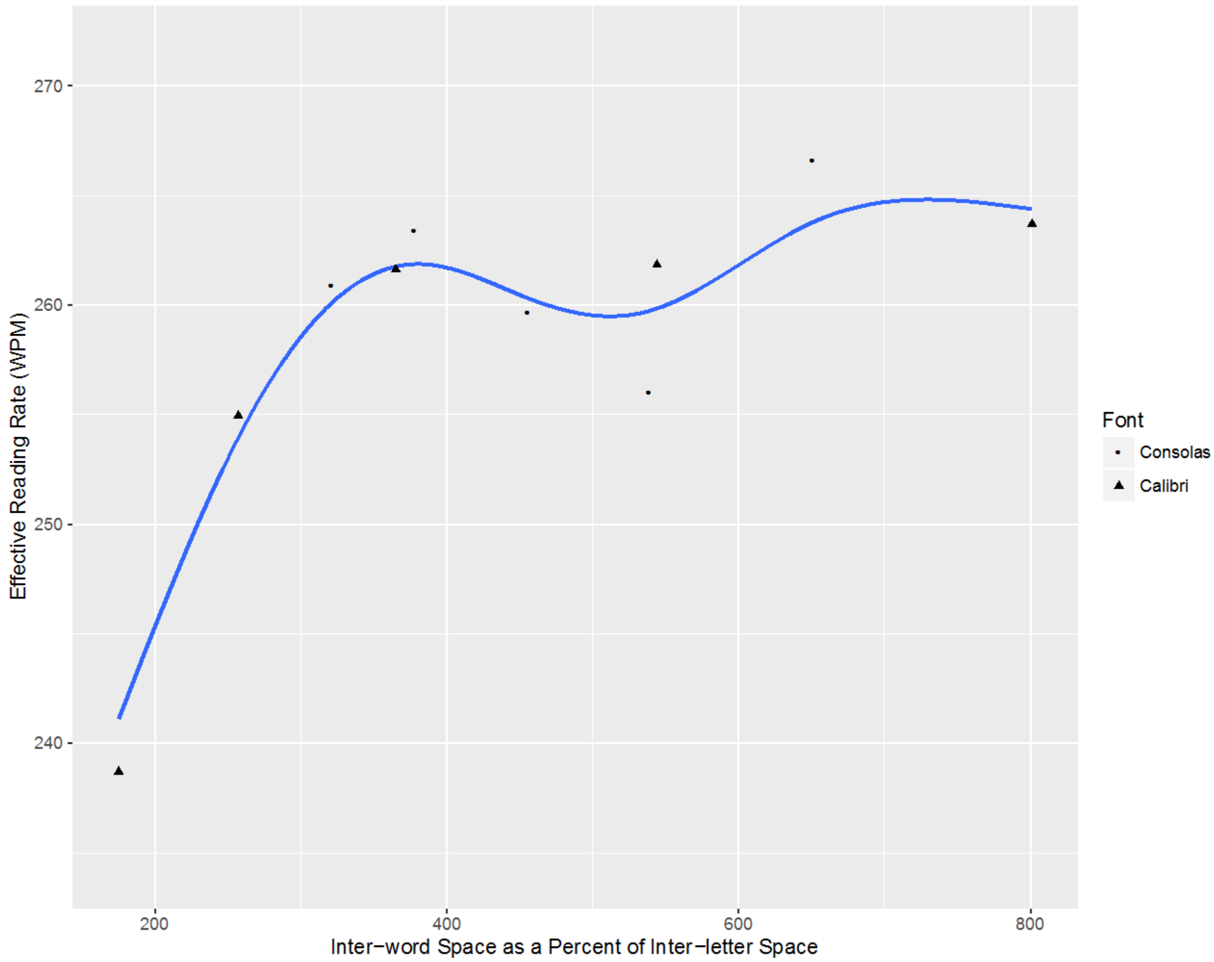
Figure 7. Effective reading rate. Error bars represent within subject confidence intervals (Morey, 2008).



Additionally, the quadratic effect of inter-word spacing interacted with the linear effect of inter-letter spacing, b = -6.05, SE = 1.96, t = -3.09 and this interaction, was marginally stronger for Consolas, b = -7.75, SE = 3.91, t = -1.98. These interactions were both largely driven by the data point for the default inter-word spacing and reduced inter-letter spacing condition for Consolas (see Figure 7). The data for this condition are not in accord with our prediction that reading rate should increase with decreasing spacing with Consolas. However, the data from the other conditions do follow the general predicted trend.

Finally, there was a significant three-way interaction between font, and the linear components of inter-word and inter-letter spacing, $b = -15.14$, $SE = 3.92$, $t = -3.86$. For Consolas, the effect of inter-letter spacing grew larger with increasing inter-word spacing. For Calibri, the effect of inter-letter spacing grew larger with decreasing inter-word spacing. This is consistent with the notion that reading becomes more difficult as inter-word spaces become less apparent. Figure 8 plots the effective reading rate as a function of inter-word space calculated as a percentage of inter-letter space[6]. The figure approximates the reading rate curve from Legge et al. (1985), especially for Calibri, and indicates that inter-word space needs to me more than three and a half times inter-letter space in order for reading to be efficient.

Figure 8. Effective Reading Rate by Percent Inter-word Space. A Loess (local regression) curve has been added to aid interpretation of the trend (shaded area denotes the confidence interval of the Loess regression).



To examine how these font and spacing manipulations influence lexical processing, we analyzed gaze durations on the target word. Gaze duration on a word is defined as the sum of reading fixations from initially fixating the word during first pass reading until leaving the word in either direction (Rayner, 1998, 2009), and can be taken as a measure of word recognition similar to the lexical decision RT in Experiments 1 and 2 (Schilling, Rayner, & Chumbley, 1998). The gaze duration LMM included the same fixed effect predictors used in the analyses of effective reading rate with the addition of a target word type variable (verb, noun, or function word). We coded contrasts for the word type variable to test two effects. The first contrast compared function words to the average of nouns and verbs (content words). The second contrast directly compared the nouns to the verbs. The model included all interactions between the main fixed effects predictors as well as random intercepts for items and subjects, and random slopes for target word type. The inclusion of random slopes for additional variables resulted in models that failed to converge.

Gaze durations (Figure 9) were shorter for target words presented in Consolas than for those presented in Calibri[7], b = -13.69, SE = 0.83, t = -16.54. There were also significant main effects of both target word type contrasts. Nouns had shorter gaze durations than verbs, b = 6.50, SE = 2.16, t = 3.02. Additionally, gaze durations were shorter on function than content words, b = 52.24, SE = 5.46, t = 9.57. There was a linear main effect of inter-word spacing, b = -11.04, SE = 1.43, t = -7.70, as gaze durations tended to decrease with increasing inter-word space. The linear effect of inter-letter spacing was also significant, b = 2.88, SE = 1.43, t = 2.01, as gaze durations tended to increase with increasing space.

Figure 9. Gaze duration as a function of font, target type, and inter-word and inter-letter spacing. Error bars represent within subject confidence intervals (Morey, 2008).

Each of the spacing variables significantly interacted with font as the spacing effects were clearly more pronounced with Calibri. The linear inter-word spacing effect was smaller for Consolas than Calibri, b = 6.66, SE = 1.43, t = 4.64. The quadratic inter-word spacing effect also interacted with font, b = -3.27, SE = 1.43, t = -2.28. Additionally, the linear inter-letter spacing effect interacted with font, b = -4.62, SE = 1.43, t = -3.23. The linear effect of inter-word spacing interacted with the linear effect of inter-letter spacing, b = -6.43, SE = 2.48, t = -2.59, as the increase in gaze durations associated with increased inter-letter spacing was more pronounced with reduced inter-word spacing.

There were also a number of interactions with target word type. The size of the inter-word spacing effect was significantly smaller for function words than for content words, b = -17.28, SE = 6.19, t = -2.79. There was a three-way interaction between the quadratic effect of inter-word space, the linear effect of inter-letter space, and the function vs. content word contrast, b = 23.20, SE = 10.69, t = 2.17. Additionally, this three-way interaction was only present for the Calibri font resulting in a four-way interaction, b = -24.07, SE = 10.68, t = -2.25. Content words (nouns and verbs) required longer gaze durations for reading when the inter-word spacing was reduced but function words were relatively immune to the inhibition associated with smaller inter-word spaces.

Finally, inter-word and inter-letter spacing influenced nouns and verbs differently as well. There was a three-way interaction between the noun vs. verb contrast, the linear effect of inter-word spacing, and the quadratic effect of inter-letter spacing and, b = 19.21, SE = 5.97, t = 3.22, as well as a three-way interaction between the noun vs. verb contrast, and the two quadratic effects of inter-word and inter-letter spacing, b = -14.18, SE = 5.98, t = -2.37. We interpret these interactions as being the result of the interplay between lexical complexity, and word segmentation processes. We interpreted the difficulty associated with increased inter-letter and decreased inter-word spacing as being due to the reduced appearance of inter-word spaces required for effective word segmentation. The current interactions suggest that the more complex verb word type suffered this difficulty not only in the condition that results in the least apparent inter-word spaces (reduced inter-word / increased inter-letter) but also the condition with the second least apparent inter-word spaces (reduced inter-word / default inter-letter). Nouns however only suffer in the condition with the least apparent inter-word space, and the least complex function words avoided difficulty with segmentation altogether.

Discussion Experiment 3

In Experiment 3, we were able to explore the influence of inter-letter spacing in conjunction with inter-word spacing for normal reading where word segmentation is crucial. While others have investigated text spacing during reading, most have confounded inter-word spacing with inter-letter spacing by manipulating both simultaneously. The current study is the first to manipulate inter-word and inter-letter spacing independently in a factorial experiment allowing us to examine the possibility of interactions between these two types of spacing. Indeed, our results highlight the importance of such interactions. As with the results of Experiment 2, spacing effects differed between fixed and proportional width fonts. The inhibition associated with increased inter-letter spacing and the facilitation associated with increased inter-word spacing both tended to be larger with the proportional width font Calibri than with the fixed width font Consolas. These interactions are likely the result of the small default inter-word spacing for Calibri relative to Consolas. Calibri's small default inter-word spacing means that increases in inter-letter spacing will obscure word boundaries more than for Consolas. Similar interactions between fonts and spacing were reported in Slattery and Rayner (2013). These interactions indicate that when assessing the influence of spacing manipulations reported in the literature one must pay special attention to the characteristics of the fonts used. This finding may also help explain why researchers using proportional width fonts have found facilitation for increased text spacing (Perea et al. 2011, 2012) and those using fixed width fonts have found inhibition (Paterson and Jordan, 2010; Cohen et al. 2008).

For typographers, these results point to additional ways in which fonts can be manipulated for optimization. Specifically, average inter-letter spacing greater than 35% of average letter width reduces the efficiency of word recognition. More importantly, the optimal amount of inter-word space depends on the amount of inter-letter space. There seemed to be a wide range of tolerable widths so long as the inter-word space was at least 3 and a half times the inter-letter space. Additionally, while some word processing software, such as Microsoft Word, allow users to make changes to font spacing, the option only allows spacing to be adjusted in a manner that maintains the proportion of inter-word to inter-letter spacing. We know of no device or software package that currently allows a user to make independent adjustments to inter-word and inter-letter spacing. Such options appear warranted given the current findings.

The local target word processing analyses not only confirmed the effects seen in the global analyses but also found important effects yet to be reported in the literature. The interactive effects of inter-word and inter-letter spacing resulting from less apparent word boundaries interacted with the type of target word. We have speculated that these interactions are due to word segmentation processes being sensitive to lexical complexity. Therefore, we would predict that as words become more difficult to integrate into the developing sentence context, the importance of inter-word space increases. Slattery and Rayner (2013) hinted at this possibility in their general discussion and finding evidence of this effect has implications for theories of written language parsing and lexical access. This effect also suggests that optimal word spacing on a line of text may involve variable inter-word spaces.

General Discussion

In the introduction, we proposed three ways in which the literature might make the effects of text spacing seem contradictory.

1. Comparison between large and small spacing increases relative to a “default”.
2. The incorrect belief that there exists a single “default” level of letter spacing.
3. A failure of researchers to differentiate inter-letter spacing and inter-word spacing.

We developed four main goals from these points, which we used to guide our research. First, we wanted to assess the replicability of inter-letter spacing benefits in single word tasks. If the conflicting reports in the literature were due to type 1 errors, there would be no benefit in exploring them further. However, we were highly successful in replicating the findings of Perea and colleagues wherein increased inter-letter spacing resulted in faster lexical decision times. We therefore dismiss the possibility that the seemingly conflicting results of inter-letter spacing were due to type 1 errors.

Second, given the differences that exist in default spacing between fonts, we wanted to explore how changes in spacing may interact with font characteristics. We tested two hypotheses connected to font characteristics: width hypothesis, and serif hypothesis. We found strong support for the width hypothesis in the form of a crossover interaction between inter-letter spacing and font width (proportional vs. fixed). As predicted, proportional width fonts, such as Times New Roman, benefitted from a small *addition* to inter-letter space, while fixed width fonts such as Courier New benefitted from a small *reduction* to inter-letter space. We found no evidence to support the serif hypothesis. That is, spacing influenced serif and san serif font in a similar manner. Additionally, we developed a new measure of inter-letter spacing, e-bar space. We used this new measure to demonstrate the differences in “default” spacing that exist between fonts. This measure suggested an optimal value of inter-letter space as ~ 35% of average letter width—at least for isolated word recognition. This ~35% inter-letter space rule held for all 6 of the fonts we tested despite fairly substantial differences between each font.

Third, we wanted to investigate inter-word and inter-letter spacing within a fully factorial experimental design so that the independent contributions of both types of spacing and their potential interaction could be assessed. This distinction has often been ignored in the literature and the majority of studies investigating text spacing confound these two variables. Ours is the first study to report a fully factorial manipulation of inter-letter and inter-word spacing. We report additional evidence of the importance of two distinct but interrelated forms of horizontal text spacing: inter-letter, and inter-word space. Experiment 3 accomplished this and found numerous interactions between inter-letter and inter-word spacing. These interactions highlight the important relationship between inter-letter and inter-word space as they indicate that inter-word space needs to be at least 3.5 times the inter-letter spacing for reading to proceed efficiently.

While Experiment 3 provided clear guidance for typographic designers regarding the use of horizontal space along a single line of text, more research is required to examine the influence of inter-word and inter-letter spacing across multiple lines of text. One side effect of the recommendations of increased inter-word relative to inter-letter spacing may be an increased prevalence of vertical “rivers” of white space through the text. These rivers occur when inter-word spaces from adjacent lines happen to overlap to a great enough extent that it gives the impression of a river of white space running vertically through a paragraph (see Dowding, 1995). Many font designers view these rivers as something to avoid. However, studies of eye movements during reading suggest that readers do not obtain useful preview benefit from the text of the line beneath the one currently being read (Pollatsek, Raney, Lagasse, & Rayner, 1993). Therefore, vertical rivers may be an aesthetic property of a text that has little influence on reading performance when readers are fully engaging their attention to the task of understanding the meaning of the text.

Finally, we wanted to explore the possibility that the inter-letter spacing effect might be different for different categories of words. If this were the case, it would open the door to new explorations in text layout. In Experiment 1, we found no evidence that inter-letter spacing differentially impacted nouns, verbs, or function words in a task involving lexical decisions of words presented in isolation. However, with the sentence reading task of Experiment 3 we found evidence that the spacing manipulations differentially influenced the types of words used in this study. These results suggest that the optimal spacing between words may depend on the characteristics (length, frequency, predictability) of the words themselves. This is not the first study to suggest that reading may benefit from the use of more variable inter-word spaces. Jandreau and Bever (1992) found that using larger inter-word spaces at phrase boundaries increased reader’s comprehension of texts. However, more research is needed to determine how sentence parsing and word segmentation processes influence each other during reading. This information will be vital to assessing the potential for using variable size inter-word spaces in text layout.

The current studies also highlight the importance of examining the reading process in settings that allow for the assessment of all the relevant cognitive processes involved. While pseudo-reading tasks involving words in isolation can be useful, they fail to capture numerous reading processes such as word segmentation (Sheridan et al., 2016), parafoveal preview (Schotter, Angele, & Rayner, 2012), syntactic parsing (Ferreira, Bailey & Ferraro, 2002; Frazier & Rayner, 1982), and predictive inference (Staub, 2015), to name just a few. Comparisons of the inter-letter spacing effect across experiments clearly demonstrate how a variable can have different effects between isolated single word tasks and natural reading. For instance, the effect of increased inter-letter spacing which had been facilitative in the isolated word task of Experiments 1 and 2 was largely inhibitory with natural sentence reading in Experiment 3. Additionally, potentially important interactions between spacing and word type that had been absent in the lexical decision experiments manifested under the more natural sentence reading task.

Fonts are the vehicle in which the eyes drive along the road of written language. Automobiles are designed with driver tendencies and terrain limitations in mind. The current research indicates that the design of fonts should consider how both reader ability and reading content are influenced by design choices. These findings also open the door to exciting new research possibilities in typographic optimization. Perhaps the amount of space placed between words in text should be a function of the written language being read, just as the best type of tires for a car depends on the seasonal road conditions. However, these font optimization effects are likely to be modest in size for readers with normal vision. Additionally, while some have reported larger benefits of increased spacing for dyslexic readers compared to typical readers, font optimization should not be seen as a cure for dyslexia (Henderson, 2014). While reading ability interacted with font spacing in Experiment 3, there was no level of spacing in the current studies that resulted in low reading ability participants displaying the fast, efficient reading performance of high ability readers. Similarly, if you gave me a Formula 1 race car, I would still be no match in a race against a professional driver—even if they were driving a far inferior vehicle. Nonetheless, it is reasonable to assume anything which makes reading easier may lead to increased engagement with reading. Therefore, even if improvements to typography result in only modest gains in effective reading rate, but bring with them a greater willingness to read, then such improvements should be pursued.

Footnotes:

1. Complicating matters, the table of studies reported by van den Boer and Hakvoort (2015) fails to indicate the precise nature of the spacing manipulations used. For instance, both Cohen et al. (2008) and Perea and Gomez (2012) are shown to use a "spacing +1.5" manipulation. However, Cohen et al. (2008) used 1.5 extra character spaces with Courier New (w o r d) while Perea and Gomez (2012) used only 1.5 extra points of space with Times New Roman (word).
2. Some studies with multiple experiments reported a combination of facilitation, inhibition, and/or null effects over the increased spacing conditions leading the total number of reported effects to be larger than the number of studies.
3. Paterson and Jordan (2010) used 3 different increased spacing conditions and found inhibition in target word reading for all three. For brevity we focused on only one of their conditions.
4. RSVP is a technique that presents words one at a time in the center of a computer monitor for a predetermined amount of time.
5. After an initial randomization, there were 23 word groupings that proved too difficult to write sentences for. These remaining words were re-randomized into new groups until all words had been written into sentences.
6. The graph has only 5 points for each font instead of 9. This is due to the fact that the design of the study resulted in some conditions having the same proportion of inter-word to inter-letter space. For instance, the default inter-word / default inter-letter condition had nearly identical proportions as the increased inter-word / increased inter-

letter condition, and the reduced inter-word / reduced inter-letter condition. Such nearly identical conditions have been averaged together for this plot.

7. While gaze durations were shorter with Consolas than with Calibri, words were skipped more often in Calibri than in Consolas. See the supplemental materials for an analysis of skipping rates and their interpretation.

References

Arditi, A. (2004). Adjustable typography: An approach to enhancing low vision text accessibility. *Ergonomics*, 47(5), 469-482. doi:10.1080/0014013031000085680

Arditi, A., & Cho, J. (2005). Serifs and font legibility. *Vision Research*, 45(23), 2926-2933. doi:10.1016/j.visres.2005.06.013

Arditi, A., Knoblauch, K., & Grunwald, I. (1990). Reading with fixed and variable character pitch. *JOSA A*, *7*(10), 2011-2015.

Bates, D., Maechler, M., Bolker, & Walker (2015). Lme4: Linear mixed-effects models using S4 classes. R Package Version 1.1-11.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, 59(4), 390-412.

Blackmore-Wright, S., Georgeson, M.A., & Anderson, S.J. (2013). Enhanced text spacing improves reading performance in individuals with macular disease. *PLoS ONE* 8(11): e80325. doi:10.1371/journal.pone.0080325

Bricolo, E., Salvi, C., Martelli, M., Arduino, L. S., & Daini, R. (2015). The effects of crowding on eye movement patterns in reading. *Acta psychologica*,*160*, 23-34.

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.

Chung, S. T. L. (2004). Reading speed benefits from increased vertical word spacing in normal peripheral vision. *Optometry and Vision Science*, 81: 525-535.

Cohen, L., Dehaene, S., Vinckier, F., Jobert, A., &Montavont, A. (2008). Reading normal and degraded words: Contributions of the dorsal and ventral visual pathways. *NeuroImage*, 40: 353-366.

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.

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

Dowding, G. (1995). *Finer points in the spacing and arrangement of type*. Hartley & Marks Publishers.

Engelkamp, J., Zimmer, H. D., & Mohr, G. (1990). Differential memory effects of concrete nouns and action verbs. *Zeitschrift für Psychologie mit Zeitschrift für angewandte Psychologie*.

Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, 11, 11–15.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14, 178–210.

Helstrup, T. (1989). Memory for performed and imaged noun pairs and verb pairs. *Psychological Research*, *50*(4), 237-240.

Henderson, J.M. (2014, November 17th). Looking at dyslexia backwards [Web log post]. Retrieved from http://jhenderson.org/vclab/Blog/Entries/2014/11/17_Looking_at_Dyslexia_Backwards.html

Jackson, M. D., & McClelland, J. L. (1979). Processing determinants of reading speed. Journal of Experimental Psychology: General, 108(2), 151.

Jandreau, S., & Bever, T. G. (1992). Phrase-spaced formats improve comprehension in average readers. *Journal of Applied Psychology*, 77(2), 143-146. doi:10.1037/0021-9010.77.2.143

Legge, G. E., Pelli, D. G., Rubin, G. S., & Schleske, M. M. (1985). Psychophysics of reading—I. Normal vision. *Vision research*, *25*(2), 239-252.

McGowan, V. A., White, S. J., & Paterson, K. B. (2015). The effects of interword spacing on the eye movements of young and older readers. *Journal of Cognitive Psychology*, 27(5), 609-621.

McLeish, E. (2007). A study of the effect of letter spacing on the reading speed of young readers with low vision. *British Journal of Visual Impairment*, 25(2), 133-143. doi:10.1177/0264619607075995

Morey, R.D. (2008). Confidence intervals for Normalized Data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology*, Vol.4(2), 61-64.

Morris, R. K., Rayner, K., & Pollatsek, A. (1990). Eye movement guidance in reading: The role of parafoveal letter and space information. *Journal of Experimental Psychology: Human Perception And Performance*, 16(2), 268-281. doi:10.1037/0096-1523.16.2.268

Morrison, R. E., & Inhoff, A. (1981). Visual factors and eye movements in reading. *Visible Language*, 15(2), 129-146.

Norris, D. (1984). The effects of frequency, repetition and stimulus quality in visual word recognition. *The Quarterly Journal of Experimental Psychology*,*36*(3), 507-518.

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.

Perea, M., & Acha, J. (2009). Space information is important for reading. *Vision Research*, 49, 1994-2000.

Perea, M., & Gomez, P., (2012a). Increasing interletter spacing facilitates encoding of words. *Psychonomic Bulletin & Review*, online DOI 10.3758/s13423-011-0214-6

Perea, M., & Gomez, P., (2012b). Subtle increases in interletter spacing facilitate the encoding of words during normal reading. PLos ONE, 7(10): e47568. Doi:10.1371/journal.pone.0047568.

Perea, M., Moret-Tatay, C., & Gomez, P. (2011). The effects of interletter spacing in visual-word recognition. *Acta Psychologica*, 137: 345-351.

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(6), 420-430. doi:10.1016/j.learninstruc.2012.04.001

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review*, 114, 273-315.

Pollatsek, A., Raney, G. E., Lagasse, L., & Rayner, K. (1993). The use of information below fixation in reading and in visual search. *Canadian Journal of Experimental*

Psychology/Revue Canadienne De Psychologie Expérimentale, 47(2), 179-200. doi:10.1037/h0078824

Pollatsek, A., & Rayner, K. (1982). Eye movement control in reading: The role of word boundaries. *Journal of Experimental Psychology: Human Perception and Performance*, 8(6), 817-833. doi:10.1037/0096-1523.8.6.817

R Development Core Team. (2015). R: A language and environment for statistical computing. Retrieved from http://www.R-project.org.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological review*, *85*(2), 59.

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological review*, *111*(1), 159.

Rayner, K. (1977). Visual attention in reading: Eye movements reflect cognitive processes. Memory & Cognition, 5(4), 443-448.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124: 372-422.

Rayner, K. (2009). The Thirty-fifth Sir Frederick Bartlett lecture: Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, 68: 1457-1506.

Rayner, K., Abbott, M. J., & Plummer, P. (2015). Individual Differences in Perceptual Processing and Eye Movements in Reading. Handbook of Individual Differences in Reading: Reader, Text, and Context, 348.

Rayner, K., Fischer, M.H.., & Pollatsek, A. (1998). Unspaced text interferes with both word identification and eye movement control. *Vision Research*, 38: 1129-1144.

Rayner, K. & Pollatsek, A. (1989). *The Psychology of Reading*. Englewood Cliffs, NJ, US: Prentice-Hall, Inc.

Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C. (2012). *The Psychology of Reading*. New York: Psychology Press.

Rayner, K., Yang, J., Schuett, S., & Slattery, T. J. (2014). The effect of foveal and parafoveal masks on the eye movements of older and younger readers. *Psychology and Aging*, 29(2), 205-212. doi:10.1037/a0036015

Reynolds, A. G., & Flagg, P. W. (1976). Recognition memory for elements of sentences. *Memory & Cognition*, *4*(4), 422-432.

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(1), 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 Psychology: Learning, Memory, and Cognition,* 37, 1440-1457.

Schilling, H. E., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition*, *26*(6), 1270-1281.

Schotter, E.R., Tran, R., & Rayner, K. (2014). Don't believe what you read (only once): Comprehension is supported by regressions during reading. *Psychological Science*, 25, 1218-1226.

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, & Psychophysics*, 74(1), 5-35.

Sheridan, H., Rayner, K., & Reingold, E. M. (2013). Unsegmented text delays word identification: evidence from a survival analysis of fixation durations. *Visual Cognition*, 21(1), 38-60.

Sheridan, H., Reichle, E. D., & Reingold, E. M. (2016). Why does removing inter-word spaces produce reading deficits? The role of parafoveal processing. *Psychonomic Bulletin & Review*, 1-10.

Slattery, T.J. (2016). Eye movements: From psycholinguistics to font design. In M. Dyson (Ed). *Digital fonts and reading*. (pp. 54-78) World Scientific.

Slattery, T. J., & Rayner, K. (2010). Eye movements and text legibility. *Applied Cognitive Psychology*, 24: 1129-1148.

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(6), 1275-1292. doi:10.3758/s13414-013-0463-8

Song, S., Levi, D.M., & Pelli, D.G. (2014). A double dissociation of the acuity and crowding limits to letter identification, and the promise of improved visual screening. *Journal of Vision* 14(5):3, 1-37. doi: 10.1167/14.5.3

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

Staub, A. (2015). The effect of lexical predictability on eye movements in reading: Critical review and theoretical interpretation. *Language and Linguistics Compass*, 9(8), 311-327.

Strasburger, H., Harvey, L. O., & Rentschler, I. (1991). Contrast thresholds for identification of numeric characters in direct and eccentric view. *Perception & Psychophysics*, *49*(6), 495-508.

van den Boer, M., & Hakvoort, B. E. (2015). Default spacing is the optimal spacing for word reading. *The Quarterly Journal of Experimental Psychology*,*68*(4), 697-709.

Venables, W. N. and Ripley, B. D. (2002) *Modern Applied Statistics with S*. Fourth Edition, Springer.

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(6), doi:10.1167/11.6.8

Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin & Review*, 8, 221-243. doi: 10.3758/BF03196158

Yap, M. J., & Balota, D. A. (2007). Additive and interactive effects on response time distributions in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(2), 274.

Yu, D., Cheung, S., Legge, G. E., & Chung, S. T. L. (2007). Effect of letter spacing on visual span and reading speed. *Journal of Vision*, 7(2):2, 1–10.

Zorzi, M., Barbiero, C., Facoetti, A., Lonciari, I., Carrozzi, M., Montico, M., Bravar, L., George, F., Pech-Georgel, C., & Ziegler, J. C. (2012). Extra-large letter spacing improves reading in dyslexia. *PNAS Proceedings Of The National Academy Of Sciences Of The United States Of America*, 109(28), 11455-11459. doi:10.1073/pnas.1205566109

Acknowledgments

The research reported here was supported by a grant from the Microsoft Corporation to Keith Rayner and the first author. We would also like to thank to Keith Rayner for his helpful comments on a much earlier draft. Correspondence should be sent to Timothy J. Slattery (tslattery@bournemouth.ac.uk).