Title:

Effects of changes in spacing on dual-script sign legibility: the role of vertical connecting spacing in bilingual (Chinese–English) traffic signs

Keywords:

Chinese-English traffic signs, dual-script typography, bilingual signage, typographic spacing, sign legibility

Abstract:

This study examined changes to typographic variables in Chinese–English traffic signs. Particularly, it considered the effect of connecting spacing – the vertical distance between a Chinese legend and its English translation – on reading performance for participants who read English but not Chinese. Participants were shown driving video simulations, featuring four connecting space measures (1/6H, 1/3H, 1/2H, 3/4H where H is the height of one Chinese character), and asked to indicate directions. A threshold method with an accuracy check was applied. The study demonstrated that connecting spacing affected participants' reading performance and is an important consideration for dual-script sign legibility.

1. Introduction

A prominent feature of Chinese traffic signs is that most of them are increasingly Chinese–English bilingual signs (Fig.1). Compared with traffic sign designs that have already developed over many decades in western countries and have cultivated a relative standardization, the evolution of Chinese traffic signs has yet to show a widely-adopted, systematic approach.

Figure 1. Chinese–English bilingual traffic signs. Photographed in Beijing by Yuchan Zhang, China, 2018

Efforts to standardize road traffic signs began in the 1980s in Mainland China. *GB5768-Road traffic signs and markings* is one of the first national standards that relates to traffic signs. It was issued in 1986 and revised in 1999 and 2009 respectively. However, there are relatively limited and inexplicit visual specifications in the standard that could support designers' decision-making. Other standards, such as *Technical Guidelines for the Replacement of National Expressway Network Related Traffic Signs* (Research Institute of Highway Ministry of Transport & Beijing Communications Highway Survey and Design Institute, 2007) and *JTC D82-Specification for Layout of Highway Traffic Signs and Markings* (Ministry of Transport of the PCR, 2009), indicate that there are visual guidelines that relate to typeface choice and the basic design of pictorial elements.[1] However, the guidelines do not sufficiently cover how to present typographic and pictorial elements in an appropriate way.

Furthermore, there is not much guidance to provide appropriate specification for the spatial presentation of combining different scripts for Chinese–English bilingual legends (to adopt Lay's 2004 term). This may be because, although there is a large amount of legibility

[1] The phrase *sign content* (Mollerup, 2013), or *sign visual communication devices* (Calori & Vanden-Eynden, 2015), is used to describe the two-dimensional graphic elements on a sign, such as textual messages, symbols, and arrows. Lay (2004) categorises textual messages as *legends*, and graphic symbols and arrows as *pictorial elements*.

research investigating monolingual signs, few studies consider sign legibility when adding another language to a sign. When they do, it seems that their considerations are limited and tend to concentrate on bilingual signs using the same script, e.g., Welsh and English (Dudek, 1991; Garvey & Mace, 1996; Jamson, 2004; Jamson et al., 2005; Rutley, 1972). The findings and solutions of these studies are concentrated on differentiating the two languages to assist users to find the information they need, so that the increased reading time caused by the doubled information could be reduced (Anttila et al., 2000; Lesage, 1981; Rutley, 1972; Smahel & Smiley, 2011). The findings and solutions, however, might not be sufficiently applicable to Chinese–English bilingual traffic signs (CEBTS) where the character sets are very different, and the type size of the Chinese legend is always larger than the size of the corresponding English legend.

There are also relatively few studies which consider dual-script typography in the specific sign scenario. Petretta (2014) considers Arabic–English signs and highlights the role of the information sequence in combining the two scripts together within a sign program because the reading direction of Arabic is different to English. Eid (2009) suggests that the differences in Arabic and English aid the users to locate the text they need quickly. But it is essential that both scripts are designed in harmony and have a balanced layout. To achieve that, it is important to consider:

1. the treatment of space – ensuring enough space to reduce clutter as the result of double information, while considering that a larger surface area might be impractical or unfeasible in some contexts or lead to additional expense);
2. scripts alignment – staggering the Arabic and English rather than typesetting them on the same line as it is read better in relatively shorter messages; and
3. the role of pictorial elements – symbols are 'bilingual' on their own since they are representations of concepts normally with a conventional meaning, so that they play an effective role in transmitting information independently of language.

Although Petretta's and Eid's works offer ideas and ways to analyse dual-script signs, their findings may not fully extend to Chinese–English typography, and to CEBTS. This is because each script has a very distinct typographic image. For example, the number of strokes for each Chinese character (Hanzi) varies from one to approximately thirty.[2] In

[2] According to the Chinese Language Research Institute (2006), the most complex Chinese character is 龖(dá) with 32 strokes.

comparison, Latin letters have fewer strokes which leads to more counter space. Furthermore, Arabic has different proportions, more connectivity and requires more spacing (Chahine, 2012; Petretta, 2014). Much of the time, these differences result in Arabic and Chinese text appearing heavier (or blacker) than Latin text when they coexist on a sign.

Furthermore, as scripts with more strokes generally appear larger than scripts with fewer strokes (Lu & Tang, 2016), Hanzi characters appear larger (optical size) than Latin letters when they are set in the same type size. On CEBTS, the type size of a Chinese legend is usually set larger than English ones (Fig.1).[3] This difference in visual saliency may increase the reading time of English readers, because the English type size is relatively small, and the Chinese legend potentially attracts more attention (as they appear much heavier and larger than the English legend). It also infers a slight visual hierarchy even though the presentation in other respects seemingly corresponds with what Tam (2012: 40) describes as "parallel bilingualism".

In addition, English is *alphabetic* and Chinese is *logographic* (Tam 2012: 44) – in English, one or several individual letters combine a semantic unit, this creates various word widths and shapes. However, in the field of Chinese typography, each Hanzi character is a semantic unit, which yields the same width in a line or block of text. Thus, Hanzi characters encode more information in a glyph than Latin characters do. Furthermore, Chinese locations tend to be shorter than their accompanying English translations. Yanqun Yang et al. (2020) suggest increasing the type size of letterform to improve CEBTS legibility. However, this would likely lead to longer English legend lengths, which may require an increased area for signs and, potentially, an increased cost.

As an alternative, this study investigates how to improve CEBTS legibility through an appropriate arrangement of typographic spacing without increasing text size. In particular, because the two languages are often vertically stacked on CEBTS with the Chinese set above its English translation (Fig.1), the main purpose of this study is to find out how the connecting spacing (Fig. 4) – the vertical spacing between a Chinese legend and its English

---

[3] The way of the measurement of type size in the reviewed standards is inexplicit and inappropriate. The standard *GB5768(1999)-Road traffic signs and markings* specifies that the size of capitals should be 1/2H (H is the height of one Chinese character) and lowercases should be 1/3H. The guidance does not specific whether H refers to the height of the body frame or surface frame (see footnote 5). In addition, it seems inappropriate to measure them by using the height of capitals and lowercases. This is because the height of lowercases with ascenders and descenders varies from one font to another and in some typefaces the capital height is lower than the ascender height. Although the height range of the Latin scripts (1/3H to 1/2H) is pointed out in standard *GB5768 (2009),* it also seems not to make sense because it does not specify what this range refers to, initial capital or x-height.

translation – affect CEBTS legibility.[4] This study builds on and is intended to complement a previous study (Zhang, 2021) that considered the effects of changes in separating space on CEBTS – the space between bilingual units of information (e.g. the spacing between different locations on a sign) rather than the space within bilingual, rhetorically-equivalent units of information (e.g. a Chinese place name and its English translation).

2. Defining variables

2.1 Four measures of connecting spacing

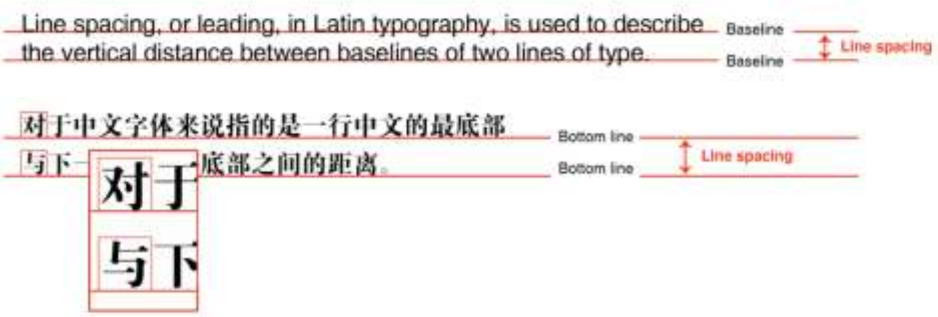
In Latin typography, the term line spacing, or leading, is used to describe the spacing of two lines of Latin text. Specifically, it is the vertical distance between baselines of two lines of type (Highsmith, 2012). Tam (2012) notes that many graphical attributes of Chinese and English typography do not have equivalents in the other script but line spacing is used in both. Line spacing in Chinese typography considers the vertical distance from the bottom line of characters to the next bottom line. Line spacing could also be measured by the top line of two lines of Chinese text. However, it is important that both the bottom and top line are relative to the body frame of a Hanzi character, i.e., the box that contains each character that includes a small amount of space around the character for attaching to the adjacent character, rather than the surface frame.[5] Figure 2 illustrates line spacing in both English and Chinese contexts.

When Chinese and English are combined into two lines of bilingual text, an issue of accurately defining the spacing arises. For our research, the term *connecting spacing* is adopted to refer to the distance between two lines of equivalent Chinese–English text to prevent potential ambiguity. As Figure 3 illustrates, connecting spacing is the distance from the *bottom line* of the Chinese legend to the *x-height line* of the English legend. The bottom line includes the lower space that is part of the body frame (so in this respect is different to

[4] There are signs indicating multiple destinations that are vertically stacked (e.g., the middle sign in Fig. 1), resulting in vertical distance separates bilingual location names. In such scenarios, the term 'separating spacing' will be used to distinguish 'connected spacing'. This paper, however, only focuses on the effect of connecting spacing on CEBTS's legibility. Please see "Effects of text space of Chinese–English bilingual traffic sign on driver reading performance" *Displays* 67 (2021) for the first author's work on separating spacing.

[5] Both the body frame and surface frame are relative to the size of Chinese characters (Lu & Tang, 2016). *Body frame* is similar to the concept of *glyph space*, or *body* (Highsmith, 2012), in the field of Latin typography, which refers to the box that contains each character and the small amount of space around the character for attaching it to the adjacent character. Generally, the body frame is fixed in a certain font. The *surface frame*, however, is the actual boundary box of a Hanzi which is smaller than the body frame. Surface frames vary from Hanzi to Hanzi in each font. Thus, the height of the body frame and the surface frame are different.

the term ‘baseline’ which Tam notes does not have an equivalent in Chinese). The bottom line of Chinese texts, rather than the top line, was used because the Chinese legend is always set



above its English translation on CEBTS. If the top line was applied, the connecting spacing would be influenced by type size.

x-height, in Latin typography, refers to the height of the lowercase x in a given typeface and it provides a way of describing the general proportions of any typeface. Because proportions vary from typeface to typeface, to prevent the connecting spacing being influenced by the different typeface used for the English legend, the x-height line, rather than the baseline, was used. For the same reason, in some typefaces the capital height is lower than the ascender height, if using *caps or ascender line* to define connecting spacing, the spacing will interact with the typeface in use, necessitating typeface-control considerations when specifying connecting spacing.

Figure 2. Line spacing in both English and Chinese contexts. In a Chinese context, the bottom line is relative to the body frame. Drawn by Yuchan Zhang

Figure 3. Connecting spacing of a Chinese–English legend. Connecting spacing is the vertical distance from the bottom line of the Chinese legend to the x-height line of the English legend. Drawn by Yuchan Zhang

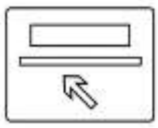
As the definition of line spacing is inexplicit in the reviewed standards (General Administration of Quality Supervision et al., 2009; Research Institute of Highway Ministry of Transport & Beijing Communications Highway Survey and Design Institute, 2007; Research Institute of Highway Ministry of Transport et al., 2017), it potentially causes difficulties in determining the measures of connecting spacing that might be tested based on the relevant standard guidance. Accordingly, the appropriate measures of connecting spacing are identified by looking at real CEBTS samples.[6] Three measures, the closest (1/6H), medium (1/2H), and the widest space (3/4H) where H refers to the height of one Chinese character are frequently used in the samples. 1/3H is the recommended 'line-spacing' in the reviewed standards (though the concept of it is ambiguous), and thus is added as an additional connecting spacing measure which may establish a metric for this study. Thus, in total, there are four measures of connecting spacing evaluated in this study.

2.2 Two levels of sign complexity

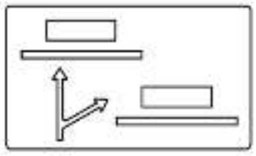
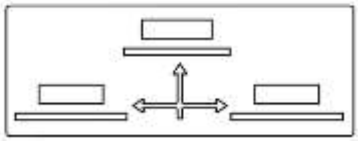
Traffic signs may show one or more locations. Accordingly, it is important to examine the complexity of signs to ensure the research findings have good application to a range of signs in practice. It is also valuable to consider whether any effects of the connecting spacing may change along with sign complexity. Sign complexity here refers to the sign's graphic and informational complexity (whereas the amount and variation in information is what leads to graphic content).

From the review of the standards and the preliminary observation of the photographed samples, it is evident that the categories of CEBTS in the standards do not correspond to the actuality of signs in practice. Hence, the sign complexity was classified based on the photographed samples as these reveal the actual physical format of CEBTS. Two levels of sign complexity, simple and complex, were identified and tested. The two levels can be presented in terms of the numbers of the directions, which are shown in Table 1. The two complexities can cover a range of signs within the collected sign samples.

Table 1. Two levels of sign complexity categorised in terms of the number of the direction.

Simple:	one direction		

[6] The real sign samples were photographed in four cities in China: Beijing, Shanghai, Wuxi, and Dalian by the first author between 2017 and 2019. The sample aimed at analyzing the design of traffic signs in practice and exploring how visual guidelines are applied to real applications.

Complex:	two directions		
	three directions		

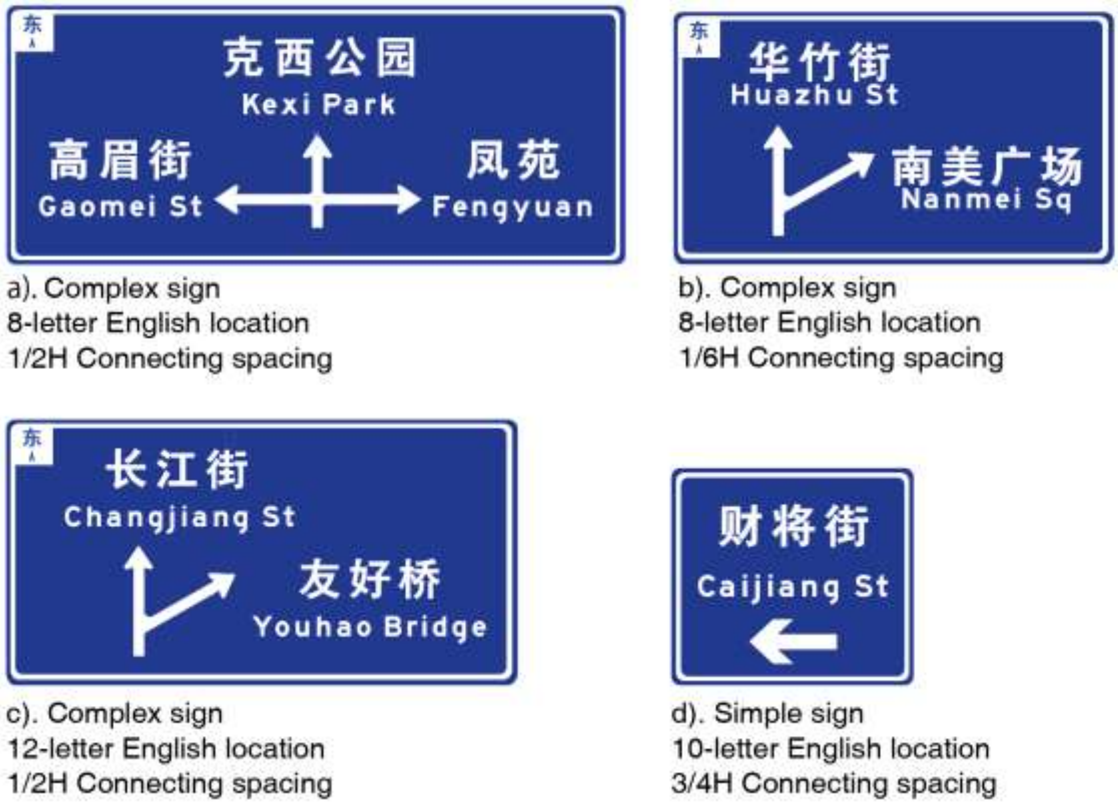
2.3 Three lengths of English information

To represent the fact that locations in the real world vary in length, in this study, the English locations on the stimuli were set into three levels: 8 letters, 10 letters and 12 letters. Because line spacing interacts with type size and line length in Latin typography (Hochuli, 2008; Luna, 1992), the three length variations also set to identify if there was an interaction between the connecting spacing and the length of an English legend on CEBTS.

In summary, the research questions of this study are:

1. whether changing connecting spacing affects participants' ability to read dual-script legends presented on CEBTS;
2. if it has an impact, to identify how large the connecting spacing should be to improve their speed and accuracy of reading;

3. how the connecting spacing interact with the sign complexity and the length of English information?



A within-subject design was used. This study was evaluated under two sign complexities (simple and complex). In each complexity, three lengths (8 letters, 10 letters, and 12 letters) of English information were tested, and each length was tested by using four levels of connecting spacing (1/6H, 1/3H, 1/2H, and 3/4H). In total, there were 24 (2×3×4) variations and each combination was presented four times in a different random order for each participant,[7] resulting in a total of 96 stimuli to be presented to each participant. Figure 4 provides four stimuli examples used in the study.

3. Study design

Figure 4. Four stimuli vary in sign complexity, length of English location, and connecting spacing used in the study

3.1 Method

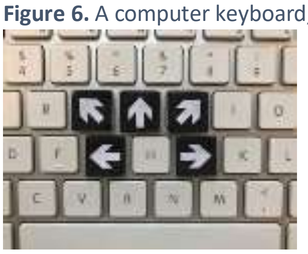
In the study, participants were shown a series of video stimuli. All video stimuli were 3D graphics rendered and displayed on a monitor. The stimuli simulated the view a driver would

[7] Thanks to E-Prime software (see section 3.3), which enabled the automatic generation of randomized trial sequences for the study.

have on a road in which they are driving towards a road sign (with legend and arrows indicating directions) at a consistent speed (Fig. 5). Thus, the signs gradually enlarged on the display until the participants were able to identify them and enter a response.

Figure 5. The screenshot of a video stimulus developed for the study. It presents a complex sign containing 8-letter English locations with the 1/6H connecting spacing

Figure 6. A computer keyboard, adjusted to provide five directional arrows enable participants to enter their response.


The study procedure shows that participants were asked to answer a question before being shown each video stimulus

A threshold method measures the first point at which an observer can detect and identify a target (Dyson, 2019). Accordingly, participants were asked to identify what

direction they might take by viewing a series of video stimuli and making an immediate response when they had identified each target. The time they took to look up a target together with the accuracy of the response were analyzed as an indication of the relative legibility of the different conditions.

The participants were cued by the researcher when the task was ready. Using the display described in Section 3.3, participants were shown video stimuli. For each one, they were asked to answer a question in the form of ‘what direction should be taken to destination xxx?’. The participants were asked to read out the question, aimed to help them to carefully read the destination and reduce the temptation to skim through the words. After that, a computer keyboard (specifically using the SPACE key) allowed the participants to self-pace when they were ready to engage with watching the video. The participants needed to find the answer by reading the sign they saw in the video. When the participants had identified the direction, they were able to indicate their response by pressing the directions on the keyboard (Fig.6), which also stopped the video and caused the screen to go to the next question slide. The participants repeated the same procedure until all stimuli had been displayed (Fig.6).

Before the main session, a small pilot was conducted to identify and adjust for any problems with the main session. In the pilot, participants viewed all designed stimuli (24 variations) without repetition (in the main session, each variation was presented four times to each participant in a different randomized order, resulting in a total of 96 stimuli to be presented to each participant). The pilot recruited six participants and took ten minutes per participant. The findings of the pilot session informed the decisions made for equipment setting that were discussed in Section 3.3.

3.2 Materials

It is important to ensure the materials have reasonable ecological validity. Therefore, the CEBTS shown in all video stimuli were designed in according with the related regulations (General Administration of Quality Supervision et al., 2009; Ministry of Housing and Urban-Rural Development of the PRC, 2015) to match the road signs that users would read in China. It covered typeface and size specifications, graphic elements guidance such as arrows and borders, as well as the spatial values such as the distance between legend (text) and pictorial elements.

Additionally, all video stimuli were developed to realistically simulate the actual driving experiences in China. In each video, the car was driven on the right side of the road, having the steering wheel on the left side that was parallel with the right-hand traffic in China. The lane width was 3.5 meters and the posted speed limit was 40 km/h that is in line with the rules of road in China. The height of the visual horizon in the videos was set to 1.2 meters above the lane based on the actual average height of a person sitting in a car

(Capaldo, 2012). The placement, size, height, and construction of CEBTS shown in the stimuli, follow the standard specifications.

In addition to ensuring the materials have ecological validity, it is important that the materials can be sufficiently controlled so that all tested variants can be compared under equivalent conditions. Accordingly, all video stimuli were 3D graphics rendered in Lumion and the two-dimensional sign surface shown in the stimuli was drawn in Adobe Illustrator 2019, rather than using real signs and actual driving videos.

The above settings prevented easy guessing and minimized any effects caused by familiarity. The bilingual text shown on the CEBTS was carefully designed to exclude the location names that are commonly used in practice. That is because many studies suggest that familiarity assists in reading signs (Lay, 2004; Sanocki, 1992; Zineddin et al., 2003). Accordingly, the Chinese locations were formed with characters that were randomly combined, and they had no semantic meaning. The characters were selected from the *Basic Vocabulary Table of Modern Chinese Characters* and were within the high-frequency category of usage (The State Language Commission, 1989). The lengths of Chinese locations ranged from two to four characters. All English locations were translations of the Chinese ones based on the relevant translation rules (*Standardization administration of China & Inspection and Quarantine of the People's Republic of China*, 2017). Furthermore, each bilingual legend was only presented once, resulting in different location names displayed across 24 variations.

All stimuli only indicated one destination per direction. That was set to control the potential influence of separating spacing (the vertical spacing used to separate bilingual legends in different groups) on sign legibility (Zhang, 2021). Furthermore, there were no passing vehicles, lane changes and slowdowns in the videos, so as not to distract participants from reading the signs. All contextual parameters were kept consistent.

3.3 Equipment and site

The stimuli were shown on a 75-inch monitor with a resolution of 1280×1024 pixels. E-Prime 2.0 software, which controlled the time, presented the stimuli, and stored the data in a spreadsheet, ran on a laptop. The findings of the pilot test informed the decisions made for the equipment settings. Some adjustments were made after the pilot that were based on asking participants' feelings and suggestions after their engagement. Based on their composite opinions, the changes made are listed below:

1. Keyboard. Changed the position of the arrows on the keyboard from the right bottom to the centre of the keyboard (Fig.6).
2. Viewing distance. Participants were seated 1.6 meters away from the monitor and it was indicated by most participants as a comfortable distance.
3. Height of the monitor. The monitor was set 1 meter above the floor and most participants felt comfortable at this height.

In line with the adjustments made after the pilot, during the main session the participants sat behind a 0.8m high table which was 1.6m away from the monitor. An adjustable chair was provided for the participants' comfort throughout the test. A computer keyboard, adjusted to provide three directional arrows for participants to respond (Fig.6), was provided on the table. Figure 7 shows a participant doing the study, using the equipment involved in the study. To keep the conditions controlled, both the pilot and main sessions used the same room.

Figure 7. A Participant doing the study

4. Participants

Participation was voluntary. This study (including the pilot) was conducted in compliance with the research ethics procedures of the University of Reading and all participants gave their consent. The participants were students and staff recruited from the university campus and met the following screening requirements:

1. Have normal or corrected vision, because eyesight has a significant impact on participant reading performance (Mandelbaum & Sloan, 1947).
2. Do not read Chinese and use English as their first or second language. The users of CEBTS vary by language ability and they can be divided into three groups: Chinese drivers, bilingual drivers who can read both Chinese and English, and drivers without Chinese reading ability (Yang et al., 2020). In this presented study, only participants who cannot read Chinese and would only rely on the English information and pictorial cues (arrows) to inform their responses took part. That was because, on CEBTS, the very different appearances of the two languages, as well as the much larger type size of Chinese locations, potentially aid Chinese and bilingual drivers to locate Chinese information faster (Eid, 2009; Yang et al., 2020). Therefore, this screening question was used to minimize a potential confound from participants' language ability.
3. Have driving experience and are between the ages of 25 to 55 years old. Driving experience and age factor also have an impact on reading road signs (Cantin et al., 2009; Kline & Fuchs, 1993; Ng & Chan, 2008). This screening criteria was introduced to recruitment to minimize the influence of potential confounds on the results.

In total, 20 participants were recruited in the main session. The study took around 40 minutes per participant, including short breaks. Each participant first completed a practice consisting of five trials, followed by a series of 96 experimental trials presented in random order. The practice trials were necessary to help participants become familiar with the equipment and procedure so that (1) they felt comfortable and could raise any queries if they needed to and (2) to ensure that the data for the first few stimuli shown was not affected by a lack of familiarity.

5. Result

5.1 Response time

Table 2 lists the mean and standard deviation (SD) of 20 participants' response times for each combination. A three-way repeated measure ANOVA was conducted to determine the effects of sign complexity, line length of English information, and connecting spacing on time taken to read the dual-script locations. Four outliers were detected that were more than 1.5 box-lengths from the edge of the box in a boxplot (Fig.8). Inspection of their values did not reveal them to be extreme and they were kept in the analysis. The response time was approximately normally distributed ($p > .05$) except for two combinations (simple sign contains 8 English letters with 1/2H connecting spacing, $p = .028$ and complex sign contains 12 English letters with 3/4H connecting spacing, $p = .032$), as assessed by Shapiro-Wilk's test of normality. The original data had been kept for analysis because there were no meaningful differences changed in statistical conclusions by running three-way repeated ANOVA on the transformed and non-transformed data. There was homogeneity of variances, as assessed by Levene's test for equality of variances, $\chi^2(2) = 23.646$, $p = .266$.

Table 2. Mean and SD of the response times (in seconds) for four levels of connecting spacing across three lengths of English locations on both simple and complex signs

			1/6H	1/3H	1/2H	3/4H
Simple sign	8 letters	Mean	3.824	4.372	3.773	3.809
		SD	1.237	1.308	0.890	1.236
	10 Letters	Mean	3.352	3.133	3.118	3.150
		SD	1.235	0.998	1.196	1.095
	12 letters	Mean	3.917	3.723	3.410	3.104
		SD	1. 395	1.077	1.108	1.044
Complex sign	8 letters	Mean	4.764	3.695	3.778	3.482
		SD	1.481	0.887	1.289	1.164
	10 letters	Mean	3.743	2.919	3.099	3.966
		SD	1.138	0.956	1.306	1.446
	12 letters	Mean	3.549	4.814	2.700	2.727
		SD	0.994	1.505	0.913	0.852

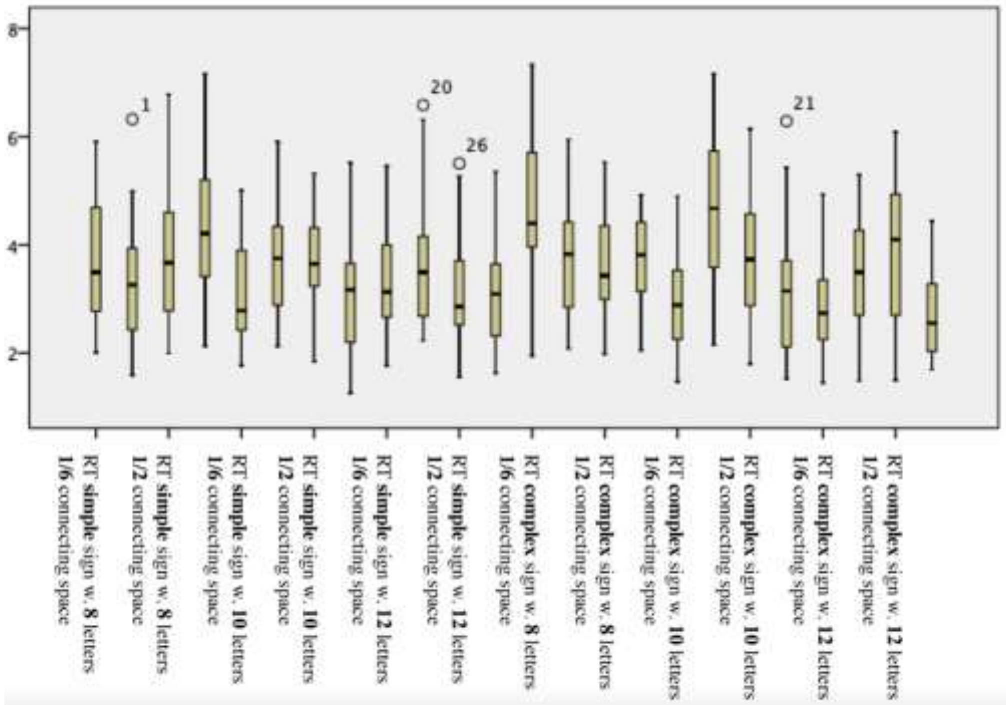


Figure 8. Boxplot of 20 participants' response times for each combination. There were four outliers, which were kept in the analysis.

The results showed a statistically significant three-way interaction between the three variables, $F(6, 114) = 15.451$, $p < .001$. There was also a significant two-way interaction between the line length and connecting spacing in both simple signs, $F(6, 114) = 2.977$, $p = .01$ and complex signs, $F(3.862, 73.376) = 26.343$, $p < .0005$, $\varepsilon = .644$. In simple sign conditions, the different levels of connecting spacing elicited a significant mean difference in response times when the length of English legend included 12 letters, $F(2.048, 38.907) = 8.924$, $p = .001$, $\varepsilon = .683$, but neither for the length of 8 letters, $F(3, 57) = 2.750$

$p = .051$, nor for the 10 letters, $F(3, 57) = 1.149$, $p = .337$. However, in complex sign conditions, the four levels of connecting spacing under all three lengths of English locations had a significant mean difference in reading speed:

8 letters: $F(3, 57) = 18.3$, $p < .001$;

10 letters: $F(3, 57) = 12.166$, $p < .001$;

12 letters: $F(1.476, 38.907) = 51.847$, $p = .001$, $\varepsilon = .492$.

Table 3. Pairwise comparisons between connecting spacing levels under both simple and complex sign conditions with the three lengths. Only the significant mean differences are presented

				Mean Difference	95% Confidence Interval for difference[b]	Sig.[b]
Simple	12 letters	1/6H	3/4H	0.813s	95% CI [0.239, 1.387]	p=0.003
		1/3H	3/4H	0.620s	95% CI [0.270, 0.968]	p<0.001
Complex	8 letters	1/6H	1/3H	1.069s	95% CI [0.433, 1.704]	p=0.001
			1/2H	0.986s	95% CI [0.444, 1.528]	p<0.001
			3/4H	1.336s	95% CI [0.746, 1.927]	p<0.001
	10 letters	1/6H	1/3H	0.823s	95% CI [0.398, 1.248]	p<0.001
		1/3H	3/4H	1.047s	95% CI [0.420, 1.673]	p=0.001
		1/2H	3/4H	0.867s	95% CI [0.205, 1.529]	p=0.006
	12 letters	1/6H	1/3H	1.265s	95% CI [0.478, 2.052]	p=0.001
			1/2H	0.849s	95% CI [0.543, 1.155]	p<0.001
			3/4H	0.822s	95% CI [0.461, 1.183]	p<0.001
		1/3H	1/2H	2.114s	95% CI [1.430, 2.798]	p<0.001
			3/4H	2.087s	95% CI [1.317, 2.857]	p<0.001

Post hoc analysis with a Bonferroni adjustment revealed that, in simple signs with the English locations including 12 letters, the reading speed was significantly faster by using 3/4H connection than using the 1/6H and 1/3H spacing levels. The difference between 3/4H and 1/2H was not significant. In complex signs with the English locations including 8 letters, there was no significant mean difference in the response time between 1/3H, 1/2H, and 3/4H. However, all the three spacing levels achieved a significantly faster response time than using the 1/6H spacing. In regard to the combination of complex signs including 10 letters, the participants' response time was significantly decreased by using 1/3H and 1/2H connecting spacing than using the other two levels, but the difference between 1/3H and 1/2H was insignificant in reading speed. Under the 12-letter line length, although the difference between the 1/2H and 3/4H connections was insignificant, both spacing levels caused a significant faster reading time than the other two levels. Table 3 presents the pairwise comparison between variables for the pairs that had a significant mean difference in reading speed.

5.2 Accuracy

Accuracy data was classified into two groups: the non-error group that refers to 100% accuracy and the error group. The non-error group occupies 86.04%. Generalized estimating equations (GEE) were used to examine if the three variables had an impact on the accuracy of reading signs. The result showed that the sign complexity, $p = .792$, English legend length, $p = .326$, and connecting spacing, $p = .508$ had no significant impact on the accuracy.

It was also important to consider if participants tried to respond slowly because they were aiming to be accurate (or vice versa). It is difficult to look at a continuous dependent variable together with a categorical dependent variable simultaneously (in terms of no statistical method has been found). Accordingly, based on the average response time, the data was classified into the fast-response group (above average) and the slow-response group (below average). In such a way, it may be able to look at the relationship between speed and accuracy. Accordingly, a Mann-Whitney U test was used to determine if there were differences in accuracy scores between the fast-and slow-response groups. The distributions of the accuracy scores for the two groups were similar, as assessed by visual inspection. The median accuracy score was not significantly different between the two groups, U = 113416, z = -.172, $p = .864$, which indicated that there was no significant difference between the response time and accuracy. In other words, this finding may suggest that the participants who responded slower did not necessarily seem to be more accurate than those who responded faster (or vice versa).

6. Discussion

The purpose of this study was to investigate whether changing the connecting spacing affects CEBTS and, if it has an impact, to identify how large the connecting spacing should be to improve the legibility, and whether the connecting spacing changes along with the sign complexity and the line length of English legends. The findings demonstrated that 1/2H connecting spacing achieved faster reading times regardless of changes in sign complexity and the length of English locations. However, the adjustment of connecting spacing did not elicit a significant difference in accuracy.

6.1 Response time

The connecting spacing affected how quickly participants read dual-script legend on CEBTS, and this effect interacted with the sign complexity and length of English locations.

Specifically, the four levels of connecting spacing did not have a significant impact on the speed of reading the simple signs, especially for the simple signs that only contained shorter English locations (8 and 10 letters). However, when the English location was longer (12 letters), the wider connection (1/2H and 3/4H) achieved a faster response times than the tighter spacing levels (1/6H and 1/3H). Although the difference between the two wider spacing was not significant, the 3/4H response time led to a more significant difference from the two tighter connections than 1/2H connecting spacing. It may indicate that, on a simple sign that only indicates one direction, the longer the English information (compared with 8 and 10 letters), the wider the connecting spacing (3/4H) might slow down the response time. This result aligns with Hochuli's (2008) and Highsmith's (2012) statements that the longer the line, the more line spacing it needs (in continuous reading) for comfortable reading.

The different measures of connecting spacing affected the speed of reading complex signs significantly. The connecting spacing 1/3H,1/2 H, and 3/4H did not have a significant difference between each other. But all achieved faster reading speed than the tightest spacing (1/6H) in a combination of complex signs containing 8 English letters; both 1/3H and 1/2H performed faster response times than other two connections on complex signs having 10 English letters; for complex signs including 12 English letters, both 1/2H and 3/4H worked better than the others. Table 4 illustrates the connecting spacing that achieved faster response times in both simple and complex signs under three lengths of English legend.

Table 4. The connecting spacing achieved significant fast response times (marked *) in both simple and complex signs with three lengths of English locations. The highlighted grey column shows that the 1/2H spacing performed well across all lengths in the complex signs, and it also worked well in simple signs containing 12 letters

Simple signs	**1/6H**	**1/3H**	**1/2H**	**3/4H**
8 letters	No significant difference			
10 letters	No significant difference			
12 letters			*	*
Complex signs	**1/6H**	**1/3H**	**1/2H**	**3/4H**
8 letters		*	*	*
10 letters		*	*	
12 letters			*	*

Table 4 shows that 1/2H connecting spacing (shaded in grey) performed well across all lengths of complex signs. In addition, it also worked well on simple signs that contain a longer English translation (12 letters). This result may suggest that using the connecting spacing of ½ height of one Chinese character could improve the reading speed of CEBTS regardless of sign complexity and the length of an English legend.

6.2 Response time versus accuracy

The results also showed that the majority of participants answered correctly, and the tested variables had no impact on accuracy. Additionally, a Mann-Whitney U test compared the median accuracy score between the fast-response group and the slow-response group, and the results indicated no significant differences between the accuracy and the speed. It may indicate that, in the study, though the accuracy was high, the participants may not sacrifice their speed to enhance it.

It is important to ensure participants prioritised speed rather than accuracy while they were doing the study. That is because speed is the paramount factor of a threshold method that indicates the sign legibility, while the accuracy check is only a supplement way to look at the results. However, in this study, it was difficult to determine if participants tried to respond more slowly because they were trying to be accurate (or vice versa), although there was a statistical analysis to examine the correlation between speed and accuracy (as Section 5.2 did). This is a potential limitation of this study and this could be addressed in the further relevant research by (1) limiting the duration of the video stimuli to require participants to identify test material for a limited period; (2) emphasizing the importance of how quickly to respond before the study; and (3) providing feedback (time they used and response accuracy) at the end of each video.

7. Contribution

The results of this study showed that the vertical spacing between a two-line of dual-script legend had a significant influence on the performance of participants reading traffic signs. This demonstrated that the typographic spacing in two different scripts is a key consideration for the legibility of a dual-script sign. The findings could contribute to guiding future design practice and standards.

Additionally, the results reinforce that connecting spacing is a useful design attribute that sign designers can utilise to organise information for a legible purpose. It provides clues as to what information belongs to which group. Specifically, on a dual-script sign, it is important to ensure that the connecting spacing can group both scripts into a bilingual unit as a whole, to convey the same meaning to their potential users. But this spacing should neither be so tight that it increases the risk of clutter due to double information, nor be so far away that it seems unrelated.

The above suggestion can be supported by the grid theory in information design and the Gestalt theory of proximity. Information designers use a grid to organise space to create structure and direct the eye flow. Samara (2017) states that a common way is to divide space based on content: like information is grouped together, disparate information is separated, and Elam (2007) advocates that the line break (and line spacing) is a useful approach to group and separate content. Using vertical spacing to organise dual-script legends based on semantic meaning of bilingual contents, therefore, improves sign structure and guide driver's eye flow, which, as a result, can increase reading speed. Among the various types of Gestalt groupings, proximity groups objects in terms of physical space (Wertheimer, 1950), which serves to bring together objects that are closer from one another than from others (Frascara, 2015). According to this, the connecting spacing on CEBTS serves to connect English translations to their right Chinese legends.

The simplified laboratory approach to understanding driver performance has often been criticised because it can hardly mimic the real-life complex environment. However, Waller (2007, p.3) suggests that testing signs in situ and in real settings "would be impracticable for several reasons, including the high cost of mounting signs with multiple factors in turn, and the difficulty in obtaining judgements in consistent conditions." Although the experimental findings of this research were obtained through participants sitting in a room, reading from a monitor display without the stress of driving, the important thing was that all test variables were compared under equal conditions. As the effects of connecting spacing have not been tested before, this study provides a cost-effective way that might help identify an appropriate range of attributes for further testing using driving simulation centres. This is particularly useful given that typographic variables are interrelated and testing different combinations can become expensive and time-consuming.

Nonetheless, ecological validity was important, and the tested variables were considered to be reasonably representative of the real-world context. In this research, the material design was informed by both visual analysis of real CEBTS samples and systematic analysis of existing standards. For example, the video stimuli and CEBTS shown in the stimuli were controlled in line with the traffic rules in China so that it was able to simulate the actual driving experience in China as much as possible. CEBTS were gradually enlarged in the display and participants were asked to perform a search task. As when driving, the sign appears to expand as the driver approaches it and drivers need to look for a destination from a sign encountered along the route, and so, the test was able to simulate the navigation activities whilst driving.

Overall, this research makes a methodological contribution through demonstrating how using a monitor to display stimuli in empirical studies can ensure that variables are sufficiently controlled and compared under equal conditions, and at the same time, to ensure these variables and findings have reasonable ecological validity. In addition, the experimental design performs a low-risk study, because participants sat in a quiet room to make responses, which enabled their safety. There is still a need for testing signs through a fully interactive driving simulator (Cantin et al., 2009; Jamson et al., 2005; Tejero et al., 2018; Yang et al., 2020). However, accessing a full integrative driving simulator is an expensive process and there are limited simulation centres that can provide research services. In contrast, the method used in this study is relatively cost-effective and easier to access. The results of this study could be used to develop appropriate materials to test in road simulation experiments (where the materials would be shown at actual size and participants might be driving a car). Therefore, this research could inform which variables would be best for researchers to test further using a fully interactive driving simulator.

Funding information

This study was supported by the China Scholarship Council. Grant number: 201808060296.

Acknowledgments

Our thanks go to Prof. Alison Black and Dr. Matthew Lickiss who shared their knowledge and provided their invaluable advice on the research undertaken for this paper. We also would like to thank the participants of our empirical study, who must remain nameless.

References

Anttila, V., Luoma, J., & Rämä, P. (2000). Visual demand of bilingual message signs displaying alternating text messages. *Transportation research part F: traffic psychology and behaviour*, *3*(2), 65–74.

Calori, C., & Vanden-Eynden, D. (2015). *Signage and wayfinding design: a complete guide to creating environmental graphic design systems*. John Wiley & Sons.

Cantin, V., Lavallière, M., Simoneau, M., & Teasdale, N. (2009). Mental workload when driving in a simulator: Effects of age and driving complexity. *Accident Analysis & Prevention*, *41*(4), 763–771.

Capaldo, F. S. (2012). Driver eye height: Experimental determination and implications on sight distances. *Procedia Social and Behavioral Sciences*, *43*, 375–383.

Chahine, N. (2012). *Reading Arabic: legibility studies for the Arabic script*. PhD thesis. University of Leiden.

Dudek, C. L. (1991). Guidelines on the use of changeable message signs. TRID database (Accession Number: 01405876)

Dyson, M. C. (2019). *Legibility: how and why typography affects ease of reading*. Centro de Estudios Avanzados de Diseño.

Eid, S. (2009). *Bilingual Information Design in Saudi Arabia*. (Unpublished master's thesis). University of Reading.

Elam, K. (2007). *Typographic Systems of Design: Frameworks for Type Beyond the Grid (Graphic Design Book on Typography Layouts and Fundamentals).* Princeton Architectural Press.

Frascara, J. (2015). *Information Design As Principled Action : Making information accessible, relevant, understandable, and usable*. Illinois.

Garvey, P. M., & Mace, D. J. (1996). *Changeable message sign visibility*. Federal Highway Administration.

General Administration of Quality Supervision, Inspection and Quarantine, & Standardisation Administration of the PRC. (2009*). Road traffic signs and markings (道路交通标志和标线).* (GB5768.3-2009). https://www.wdfxw.net/doc11265607.htm

Highsmith, C. (2012). *Inside Paragraphs: Typographic Fundamentals*. Font Bureau.

Hochuli, J. (2008). *Detail in typography: letters, letterspacing, words, wordspacing, lines, linespacing, columns.* Hyphen Press.
Jamson, S. (2004). Evaluation of techniques to improve the legibility of bilingual Variable Message Signs. *Advances in Transportation Studies*, *4*, 71–88.
Jamson, S., Tate, F., & Jamson, A. H. (2005). Evaluating the effects of bilingual traffic signs on driver performance and safety. *Ergonomics*, *48*(15), 1734–1748.
Kline, D. W., & Fuchs, P. (1993). The Visibility of Symbolic Highway Signs Can Be Increased among Drivers of All Ages. *Human Factors*, *35*(1), 25–34. https://doi.org/10.1177/001872089303500102
Lay, M. G. (2004). Design of traffic signs. In C. Castro & T. Horberry (Eds.), *The human factors of transport signs* (pp. 25-48). CRC Press, Taylor & Francis Group.
Lesage, P. B. (1981). *Design and comprehension of bilingual traffic signs*. TRID database (Accession Number: 00370841).
Lu, X., & Tang, T. (2016). Elements of Chinese typeface design. In Mary C Dyson & Ching Y Suen (Eds.), *Digital Fonts and Reading* (pp. 109–130). World Scientific Publishing.
Luna, P. (1992). *Understanding type for desktop publishing*. Chapman & Hall, Ltd.
Mandelbaum, J., & Sloan, L. L. (1947). Peripheral Visual Acuity: with special reference to scotopic illumination. *American Journal of Ophthalmology*, *30*(5), 581–588.
Ministry of Housing and Urban-Rural Development of the PRC. (2015). *Code for layout of urban road traffic signs and markings (城市道路交通标志和标线设置规).* (GB 51038—2015). https://www.wdfxw.net/doc20970684.htm
Ministry of Transport of the PRC. (2009). *Specification for Layout of Highway Traffic Signs and Markings (公路交通标志和飚线设置规范).* (JTC D82—2009). https://www.waizi.org.cn/bz/116644.html
Mollerup, P. (2013). *Wayshowing-> Wayfinding*. BIS Publishers.
Ng, A. W. Y., & Chan, A. H. S. (2008). The effects of driver factors and sign design features on the comprehensibility of traffic signs. *Journal of Safety Research*, *39*(3), 321–328.
Petretta, J. (2014). Arabic sign design: Right to left and left to right. *Information Design Journal*, *21*(1), 18–23. https://doi.org/10.1075/idj.21.1.04pet
Research Institute of Highway Ministry of Transport, & Beijing Communications Highway Survey and Design Institute. (2007). *Technical Guidelines for the Replacement of National Expressway Network Related Traffic Signs (国家高速公路网相关标志更换工作实施技术指南).* China Communications Press. http://www.jianbiaoku.com/webarbs/book/13627/676677.shtml
Research Institute of Highway Ministry of Transport, Highway Administration of Anhui Province, & Beijing Highway Survey and Design Institute. (2017). *Technical Guidelines for the Adjustment of National Highway Network Traffic Signs (国家公路网交通标志调整工作技术指南).* https://www.wdfxw.net/doc100277416.htm?fs=1
Rutley, K. (1972). *An investigation into bilingual (Welsh/English) traffic signs.* Transport and Road Research Laboratory.
Sanocki, T. (1992). Effects of font-and letter-specific experience on the perceptual processing of letters. The American journal of psychology, 105(3), 435–458.
Samara, T. (2017). *Making and breaking the grid, updated and expanded: A graphic design layout workshop.* Rockport Publishers.
Smahel, T., & Smiley, A. (2011). Evaluation of bilingual sign layout and information load before implementation of new signing system. *Transportation research record*, *2248*(1), 37–44.
Standardization administration of China, & Inspection and Quarantine of the People's Republic of China. (2017). *Guidelines for the use of English in public service areas (公共服务领域英文译写规范).* (GB/T 30240.2 —2017). http://wb.beijing.gov.cn/home/wswm/yyhj/bscx/gjbz/202002/P020200211586150202305.pdf
Tam, K. C. (2012). A descriptive framework for Chinese–English bilingual typography. *Typographische Monatsblätter*, *4*(5), 38–46.
Tejero, P., Insa, B., & Roca, J. (2018). Increasing the default interletter spacing of words can help drivers to read traffic signs at longer distances. *Accident Analysis & Prevention, 117*, 298–303.
The State Language Commission（国家语言文字工作委员会）(1989). *High-frequency Category of Usage of Common Chinese Characters (现代汉语常用字频度统计).* Language Press (语文出版社).
Waller, R. (2007). Comparing typefaces for airport signs. *Information Design Journal, 15*(1), 1–15.
Wertheimer, M. (1950). Gestalt theory. In W. D. Ellis (Ed.), *A sourcebook of Gestalt psychology* (pp. 1–11). Humanities Press.

Yang, Y., Chen, J., Easa, S. M., Zheng, X., Lin, W., & Peng, Y. (2020). Driving simulator study of the comparative effectiveness of monolingual and bilingual guide signs on Chinese highways. *Transportation research part F: traffic psychology and behaviour, 68, 67*–78.

Zhang, Y. (2021). Effects of text space of Chinese-English bilingual traffic sign on driver reading performance. *Displays, 67*, 102002. https://doi.org/10.1016/j.displa.2021.102002

Zineddin, A. Z., Garvey, P. M., Carlson, R. A., & Pietrucha, M. T. (2003). Effects of practice on font legibility. *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting, 47*(13), 1717–1720. https://doi.org/10.1177/154193120304701326