

## ABSTRACT

Two experiments were conducted to investigate the effects of minimal legible size characters on Chinese word recognition. In Experiment 1, the minimal legible size was determined empirically to be the character size necessary to attain ninety-five percent correct recognition for various Chinese characters which differed in the number of strokes comprising the character, ranging from three to twenty-seven. The results showed that the minimal legible sizes were larger for characters with more strokes. This indicates that characters with more strokes should be enlarged to attain the same recognition performance as that from characters with fewer strokes. Experiment 2 investigated recognition accuracy for a string of minimal legible size characters, versus, conventional equal size characters. The results showed that accuracy rate for the minimal legible size condition was higher than that for the conventional size condition. Although Chinese characters presented with their minimal legible size might change the present word configuration, the results suggest that minimal legible size of characters might help readers recognize words in situations where reading time is extremely short. In particular, the results suggest that minimal legible size Chinese characters may be appropriate in the design of warning or emergency signs.

## EFFECTS OF MINIMAL LEGIBLE SIZE

## CHARACTERS ON CHINESE WORD RECOGNITION

*Sheng-Hsiung Hsu  
Kuo-Chen Huang*

*Sheng-Hsiung Hsu is a professor of Human Factors in the Department of Industrial Management at National Taiwan University of Science and Technology in Taipei. He received his doctorate in Psychology from the University of Paris and is a former chairperson of the Department of Industrial Management and former president of the Directorate, Ergonomics Society of Taiwan.*

*Kuo-Chen Huang is a doctoral candidate in the Department of Industrial Management at National Taiwan University of Science and Technology. He formerly was a staff member of the Criminal Investigation Bureau at the National Police Administration, Ministry of Interior Affairs.*

National Taiwan University of Science and Technology  
43, Kee-lung Road, Section 4  
Taipei, Taiwan, 106  
*Visible Language* 35.2  
Sheng-Hsiung Hsu, 178-191  
E-mail: shhsu@im.ntust.edu.tw

© *Visible Language*, 2001  
Rhode Island School of Design  
Providence, Rhode Island 02903

# 讀員

Chapanis (1965) called attention to a very important but often neglected area of human factors: the written information associated with the tools, machines, systems and operations, which have more often been the focus of research. Negligence of this particular issue has continued over the past two decades (Chapanis, 1988). Advances in technology and automation have resulted in the introduction of increasingly complex products, machinery, operational activities and systems that require increasingly more complex written information to communicate between humans and machines. This expansion of information has contributed to the creation of brochures, warning signs, illustrations and operational manuals that are frequently incapable of being adequately understood by their users. The design of such written information for accurate and efficient human processing deserves more study.

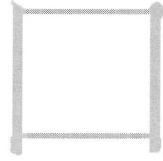
An important concern regarding the design of written information is legibility. Legibility refers to the ease of identifying individual alphanumeric characters. Poor legibility may result in the information being difficult to

read, and thus a deterrent to reading it. Important determinants of legibility are contrast, typographic style and size. Given adequate contrast, legibility is strongly affected by the type style. Much research has shown that character size significantly influences legibility (Chang and Konz, 1993; Imbeau et al., 1989; Joseph and Uhlarik, 1992; Miyao et al., 1988; Silver and Braun, 1993). Some studies (Bouwhuis, 1993; Vigilante and Wogalter, 1998) showed the effect of character size on reading performance was significant for elderly readers. In order to obtain good text legibility, an important question is how large a character must be to be identified and discriminated from adjacent characters.

Although extensive guidelines exist with respect to character size, contrast and typography, problematic issues remain concerning the spatial arrangement and characteristics of the elements within a character, in particular, those affecting discriminability of the elements, and thereby the recognition process. In English, each word may fill a different amount of space, for example, the space for "people" is twice that for "the." But, the discriminability for the letters of these two words is probably quite similar. However, written Chinese characters are square-shaped, and all characters occupy equal spaces. Some characters have few strokes, while others have twenty or more strokes. Thus, the discriminability of the strokes that form the equal size characters is different. Some investigators (Yen and Liu, 1972; Yu and Cao, 1992; Zhang and Yang, 1987) indicated that the effect of stroke number on Chinese character recognition was statistically significant and suggested that characters with more strokes are more difficult to identify than those with less strokes.

For instance, although the characters 元 (*yuán*, dollar) and 寶 (*bao*, treasure) are the same size, in terms of space occupied, they differ significantly in number of strokes and the discriminability of their strokes.

Based on the findings presented above, recognition performance should be affected by the number of strokes within a character and character size. If this is true, then how do we redesign characters, such as 元 (*yuán*, dollar) and 寶 (*bao*, treasure), so that they are more equally legible? Perhaps, 元 (*yuán*, dollar) should be drawn a little smaller, and 寶 (*bao*, treasure) should be drawn a little larger to produce equally legible characters, and thus improve overall recognition performance. The purpose, then, for Experiment 1 was to determine, for Chinese characters having different numbers of strokes, the minimal legible size which produced a high (ninety-five percent) correct recognition accuracy. Experiment 2 evaluated these minimal legible size characters presented within a string of Chinese text.



### **EXPERIMENT 1: SPECIFY THE MINIMAL LEGIBLE SIZE FOR CHINESE CHARACTERS**

The purpose of this experiment was to specify the minimal legible size for Chinese characters comprised of different numbers of strokes under extremely short exposure durations.

#### **METHOD**

##### **SUBJECTS**

Thirty college students between 20 and 22 years old ( $M=20.8$ ,  $SD=0.82$ ) served as subjects. All had 18/20 corrected visual acuity or better. The subjects were paid for their participation.

##### **MATERIALS AND DESIGN**

Two hundred fifty characters with 10 occurrences per 800,000 characters in Chinese literature (Wu, 1987) were used in this experiment. The 250 characters were divided into five levels based on the number of strokes per character: Level 1 (3 to 7 strokes), Level 2 (8 to 12 strokes), Level 3 (13 to 17 strokes), Level 4 (18 to 22 strokes), and Level 5 (23 to 27 strokes). Each level contained fifty characters, ten characters per stroke number.

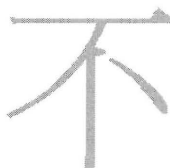
Ten levels of character size were used: The largest characters were 19 mm by 19 mm; the smallest characters were 8.2 mm by 8.2 mm. The increment between levels was 1.2 mm.

The selection of 19 mm x 19 mm as the largest size character was based on the results from a preliminary study using the method-of-limits (Gescheider, 1976; Liu, 1976). In this preliminary study, 10 subjects (who did not participate in the formal experiment) were presented 50 words consisting of 27-stroke characters for 250 msec at a viewing distance of 200 cm. The result was that a 19 x 19 mm character size was required to achieve 95% correct recognition.

The design consisted of a 5 (Stroke Number Level) x 10 (Character Size) mixed design, with character size as the between-subjects factor and stroke number level as the within-subjects factor. Subjects practiced ten characters before their performance data were collected on the experimental characters. Each subject received five blocks of trials in about 80 minutes. A block consisted of 50 trials with stimulus characters of the same stroke number level. There was a 5 minute break between blocks. The sequence of the five stroke number levels was randomized.

#### **TASK SPECIFICATION**

All subjects were tested individually. A Model G1290 tachistoscope (by Gerbrands Corporation) was used to present the task. Prior to the presentation of each stimulus, subjects were instructed to fixate their eyes at the center of the screen. Two seconds later, a stimulus was presented for 250 msec. exposure duration. Then, the subjects reported the character just presented. Subjects were instructed to perform the task as accurately as possible.



#### **PERFORMANCE MEASURE AND ANALYSIS**

Accuracy rate (percentage correct) were obtained for each block of 50 trials and entered into the analysis of variance performed using SAS.

#### **RESULTS AND DISCUSSION**

The means for accuracy rate under the different experimental conditions are shown in table 1. The results of analysis of variance showed a main effect of stroke number ( $F_{4,80}=629.3$ ,  $p<.001$ ) on accuracy rate. Multiple comparisons with the Duncan test showed that the differences among Level 1 ( $M=90.9$ ,  $SD=12.8$ ), Level 2 ( $M=84.8$ ,  $SD=18.1$ ), Level 3 ( $M=77.0$ ,  $SD=27.0$ ) and Level 5 ( $M=58.9$ ,  $SD=30.9$ ) were significant. The difference between Level 3 and Level 4 ( $M=76.1$ ,  $SD=24.9$ ) was not statistically different. The results indicated that the more strokes a character has, the more difficult it is to recognize the character.

Table 1

## Stroke Number

	Level 1		Level 2		Level 3		Level 4		Level 5		Subtotal	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
<b>Character Size</b>												
8.2	66.27	1.94	47.45	1.24	27.10	3.00	36.26	1.19	13.92	3.52	38.20	18.57
9.4	72.44	2.30	64.42	3.59	40.74	1.48	38.23	2.80	18.04	1.73	46.77	20.31
10.6	78.27	3.85	68.17	1.94	49.48	6.59	49.57	2.44	27.34	1.41	54.57	18.43
11.8	95.38	3.87	83.60	2.19	74.59	3.82	72.60	3.59	42.48	1.91	73.73	18.39
13.0	96.55	2.81	88.67	6.44	84.38	3.10	79.58	2.61	52.40	4.88	80.32	15.96
14.2	100	0	97.27	1.70	96.95	1.49	91.60	1.72	74.14	3.43	91.99	9.81
15.4	100	0	98.29	2.99	98.17	0.30	94.63	1.25	84.44	2.05	95.11	5.99
16.6	100	0	100	0	98.86	0.26	98.89	1.02	87.51	1.61	97.05	5.02
17.8	100	0	100	0	99.77	0.39	99.84	0.28	91.92	1.25	98.31	3.34
19.0	100	0	100	0	100	0	100	0	96.57	1.73	99.31	1.56
<i>Subtotal</i>	<i>90.89</i>	<i>12.9</i>	<i>84.79</i>	<i>18.2</i>	<i>77.00</i>	<i>27.0</i>	<i>76.12</i>	<i>24.9</i>	<i>58.88</i>	<i>30.9</i>		

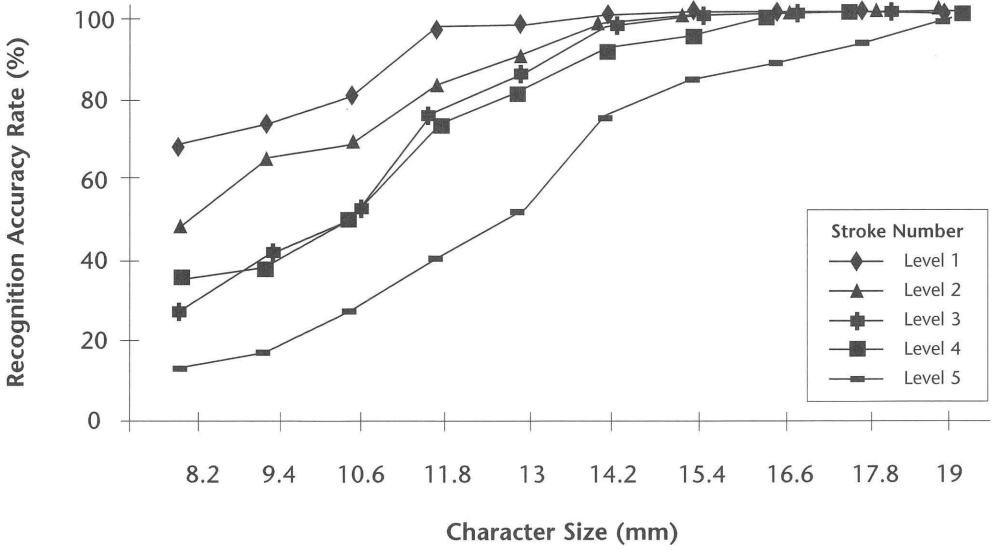
Means and Standard Deviations of Accuracy Rate (%) under Different Conditions of Exp. 1 (n=3).

絕

The main effect of character size ( $F_{9,20}=2261.7$ ,  $p<.001$ ) on accuracy rate was significant. Multiple comparisons with the Duncan test showed that the differences among 8.2 mm ( $M=38.20$ ,  $SD=18.57$ ), 9.4 mm ( $M=46.77$ ,  $SD=20.31$ ), 10.6 mm ( $M=54.57$ ,  $SD=18.43$ ), 11.8 mm ( $M=73.73$ ,  $SD=18.39$ ), 13.0 mm ( $M=80.32$ ,  $SD=15.96$ ), 14.2 mm ( $M=91.99$ ,  $SD=9.80$ ), 15.4 mm ( $M=95.11$ ,  $SD=5.99$ ), 16.6 mm ( $M=97.05$ ,  $SD=5.02$ ) and 19 mm ( $M=99.31$ ,  $SD=1.56$ ) were significant. The differences between 16.6 mm and 17.8 mm ( $M=98.31$ ,  $SD=3.34$ ), and between 17.8 mm and 19 mm were not statistically different. The interaction between stroke number and character size was also significant ( $F_{36,80}=32.0$ ,  $p<.001$ ). Figure 1 shows this interaction, in which the effect of stroke number on accuracy rate was greater for smaller characters than for larger characters. However, when the character size exceeded 15.4 mm, the effect of stroke number on accuracy rate was less significant.

The purpose of this experiment was to specify the minimal character size necessary to attain a recognition level greater than ninety-five percent correct, for characters with different stroke numbers. These minimal sizes referred to herein as the minimal legible sizes are presented in table 1, and are as follows: 11.8mm, 14.2mm, 14.2mm, 16.6mm, and 19.0mm, for stroke number Levels 1 through 5, respectively. These results indicate that characters with more strokes must be enlarged to obtain the same recognition accuracy as that for those with fewer strokes.

Figure 1



Recognition accuracy as a function of character size for characters which varied in number of strokes per character.  
(Note significant interaction between stroke number and character size).

## EXPERIMENT 2: EFFECTS OF MINIMAL LEGIBLE SIZE CHARACTERS ON RECOGNITION

The purpose of this experiment was to investigate the effects of minimal legible size characters on recognition performance in a task using a string of Chinese characters.

### METHOD

#### SUBJECT

Thirty-six college students (6 women and 30 men) between 20 and 24 years old ( $M=22.4$ ,  $SD=1.02$ ) served as subjects. All had 16/20 corrected visual acuity or better. The subjects were paid for their participation.

#### MATERIALS

Ninety-six four-character Chinese strings, including 32 high-frequency, 32 low-frequency words and 32 nonwords were used in this experiment. The high-frequency words appeared over six times per 800,000 characters, and the low-frequency words appeared only one time per 800,000 characters in Chinese literature (Wu, 1987).

Two presentation conditions for the character strings were employed. In the redesign condition, each character was presented at its minimal legible size as obtained from Experiment 1. Thus, the sizes of the four characters within any string could be different. In the conventional condition, the size of all four characters of a string was the same; it was the average of the minimal legible sizes of the string's comprising characters (see *figure 2* for illustration).

#### DESIGN

The design consisted of a 3 (Word Frequency)  $\times$  2 (Presentation Condition) mixed design, with presentation condition as the between-

subjects variable. Subjects practiced with ten training strings before their performance data on the experimental strings were collected.

The stimulus exposure duration was decided in a preliminary study, in which 12 persons (who did not participate in the formal experiment) received 30 stimuli in a recognition task. The results showed that a 400 msec. exposure duration was necessary to produce a mean accuracy rate of 52.2% ( $SD=9.2$ ) at a viewing distance of 200 cm. Thus, the exposure duration of the stimulus was set at 400 msec. in the formal experiment.

#### TASK SPECIFICATION

Subjects were tested individually and informed that the character string was to be recognized in the left-to-right direction. A Model G1290 tachistoscope (by Gerbrands Corporation) was used to present the task. Prior to the presentation of each stimulus, subjects were instructed to fixate their eyes at the center of the screen. Two seconds later, a stimulus was presented, then, they reported the string they had just read. Each subject received three blocks of 32 trials in about 40 minutes. There was a 5 minute break between blocks. Subjects were instructed to perform the recognition task as quickly and accurately as possible.

#### PERFORMANCE MEASURE

Accuracy rates (percentage correct) were obtained for each block of 32 trials and entered into the analysis of variance performed using SAS.

#### RESULTS AND DISCUSSION

The mean accuracy rates under the different experimental conditions are shown in table 2. The results of analysis of variance showed a

# 筆劃

Figure 2

	Redesign condition	Conventional condition
High-frequency word	讚不絕口	讚不絕口
Low-frequency word	波瀾壯闊	波瀾壯闊
Nonword	瀨生通水	瀨生通水

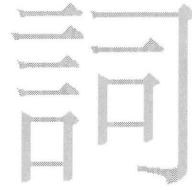
*Some examples of character size design in Experiment 2. In the redesign condition, the sizes of characters for the high-frequency word were 19.0, 11.8, 14.2, and 11.8mm; those for the low-frequency word were 14.2, 16.6, 11.8, and 14.2mm; those for the nonword were 19.0, 11.8, 14.2, and 11.8mm, respectively. In the conventional condition, the size of all characters were the average of the sizes of the characters in the redesign condition: 14.2mm.*

main effect of presentation condition ( $F_{1,34}=37.9, p<.001$ ) on accuracy rate. The accuracy rate for the redesign condition ( $M=50.8, SD=8.3$ ) was significantly higher than that for the conventional condition ( $M=47.3, SD=8.7$ ). The main effect of word frequency ( $F_{2,68}=407.1, p<.001$ ) on accuracy rate was also significant. Multiple comparisons using the Duncan test showed that the differences among high-frequency words ( $M=56.2, SD=3.0$ ), low-frequency words ( $M=53.1, SD=3.0$ ), and nonwords ( $M=38.0, SD=4.0$ ) were significant. Interaction of the two factors was not statistically significant.

The results indicate that the accuracy rate was higher when the characters of a word were presented at their minimal legible sizes than when they were presented at their conventional size. The results also showed that word frequency affects accuracy rate. High-frequency words were correctly recognized more often than were low-frequency words, which in turn, were recognized better than meaningless nonwords.

## CONCLUSIONS

The results of Experiment 1 indicate that recognition accuracy was higher for Chinese characters with fewer strokes than for characters with more strokes, a finding consistent with other researchers (Yen and Liu, 1972; Yu and Cao, 1992; Zhang and Yang, 1987). According to Just and Carpenter (1987), the average focus time for each stroke is 4.6msec. Hence, subjects need more time to identify the characters with more strokes. In addition, characters in Chinese text have the same size regardless of number of strokes. Thus, the characters with more strokes would necessarily have a higher stroke density, and would be



more difficult to recognize. One implication is that the size of Chinese characters should vary directly with character stroke number, in order to maintain the same level of discriminability for each character.

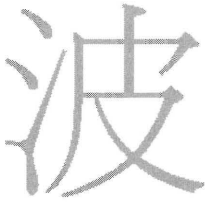
This implication was confirmed in Experiment 2, in which recognition accuracy was higher for words consisting of minimal legible size characters, versus, the conventional equal size. One possible reason for this result is that when the size of the character with more strokes was enlarged, it became easier for the subjects to recognize and the accuracy rate increased. However, perhaps the smaller size for characters with fewer strokes did not degrade recognition performance because the size was still within the minimal legible range. The recognition accuracy rate for high-frequency words was greater than that for low-frequency words. This result is similar to the results of other studies (Clark et al., 1994; Duchek and Neely, 1989; Howes and Solomon, 1951; Johnston, 1978; Scarborough et al., 1977). The effect of word frequency was still

Table 2

## Word Frequency

	High-frequency		Low-frequency		Nonword		<i>Subtotal</i>	
	M	SD	M	SD	M	SD	M	SD
<b>Design</b>								
Redesign condition	57.87	2.56	54.54	2.50	40.05	3.44	50.82	8.30
Conventional condition	54.47	2.43	51.57	2.75	35.92	3.59	47.32	8.72
<i>Subtotal</i>	56.17	3.00	53.06	2.99	37.99	4.04		

Means and Standard Deviations of Accuracy Rate (%) under Different Conditions of Exp. 2 (n=18).



significant when the characters of a word were individually designed with their minimal legible sizes.

The minimal legible size design varies the sizes of the characters according to their stroke numbers: Character with fewer strokes required less size and vice versa. Different character sizes may not be plausible in the same text due to editing and aesthetic considerations, but for “signs” with only a few characters, character sizes could be designed according to stroke number. If the sizes of the characters in a sign must be equal, the appropriate size should be equal to or larger than the minimal legible size of the character with the highest stroke number. These character size recommendations may prove appropriate for emergency situations where recognition accuracy is critical and time is short.

The findings of this study might be useful for a designer in creating a more easily recognized Chinese character string. However, the experimental stimuli used in this study were short character strings; future research is needed to explore the effects of minimal legible size design on reading sentences or text. This design also changes the word configurations. The effect of these changes on recognition accuracy needs further study.

#### **ACKNOWLEDGMENT**

The authors would like to thank the National Science Council of the Republic of China for financial support, Contract No. NSC 87-2213-E-011-031. We also thank Dr. K. K. Shieh and Dr. Tommy. R. Morrison for useful comments and suggestions.

## REFERENCES

- Bouwhuis, D. G. 1993. "Reading rate and letter size." In *IPO Annual Progress Report 28, Institute for Perception Research*, 30-36.
- Chang, P. T., and Konz, S. 1993. "Character size vs viewing distance on VDTs." In H. Luczak, A. Cakir and G. Cakir, editors. *Work with Display Units 92*. North-Holland, 143-147.
- Chapanis, A. 1965. "Words, words, words." *Human Factors*, 7:1, 1-17.
- Chapanis, A. 1988. "Words, words, words' revisited." *International Reviews of Ergonomics*, 2:1, 1-30.
- Clark, S. E., and R. E. R. Burchett. 1994. "Word frequency and list composition effects in associative recognition and recall." *Memory & Cognition*, 22:1, 55-62.
- Duchek, J. M., and J. H. Neely. 1989. "A dissociative word-frequency x levels-of-processing interaction in episodic recognition and lexical decision tasks." *Memory & Cognition*, 17, 148-162.
- Gescheider, G. A. 1976. *Psychophysic: Method and theory*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 28-34.
- Howes, D. H., and R. L. Solomon. 1951. "Visual duration threshold as a function of word probability." *Journal of Experimental Psychology*, 41, 401-410.
- Imbeau, D., W. W. Wierwille, L.D. Wolf and G. A. Chun. 1989. "Effects of instrument pane luminance and chromaticity on reading performance and preference in simulated driving." *Human Factors*, 31:2, 147-160.
- Johnston, J. C. 1978. "A test of the sophisticated guessing theory of word perception." *Cognitive Psychology*, 10, 123-153.
- Joseph, K. M., and J. Uhlarik. 1992. "The effects of legibility and display size on performance for RAPCOM and spatial display formats." In *Innovations for Interactions. Proceedings of the Human Factors Society 36th Annual Meeting, Atlanta, Georgia, October 12-16, 1992*. Santa Monica, California: The Human Factors Society 2, 1388-1392.
- Just, M. A. and P. A. Carpenter. 1987. *The Psychology of Reading and Language Comprehension*. Newton, Massachusetts: Allyn and Bacon, 25-43.
- Liu, I. M. 1976. "Evidence of sensory conditioning with pigeons." *Acta Psychologica Taiwanica*, 18, 19-24.
- Miyao, M., J.S. Allen, S. S. Hacidalihzade, S. A. Cronin and L. W. Stark. 1988. "The effect of CRT quality on visual fatigue." In F. Aghazadeh, editor. *Trends in Ergonomics/Human Factors V*. North-Holland, Amsterdam, 297-304.
- Scarborough, D. L., C. Cortese and H. S. Scarborough. 1977. "Frequency and repetition effects in lexical memory." *Journal of Experimental Psychology: Human Perception and Performance*, 3, 1-17.
- Silver, N. C. and C. C. Braun. 1993. "Perceived readability of warning labels with varied font sizes and styles." *Safety Science*, 16(5-6), 615-625.
- Vigilante, W. J. and M. S. Wogalter. 1998. "Older adults' perceptions of OTC drug labels: Print size, white space and design type." In S. Kumar, editor. *Advances in Occupational Ergonomics and Safety 2*. IOS Press, 599-602.
- Wu, J, T. 1987. *Exploring the phonetic and semantic features of Chinese words: Part 1. Frequency count of Chinese words, characters, and monosyllables* (Grant No. NSC75-0301-H002-24. Taipei, Taiwan: National Science Council (in Chinese).
- Yen, J. S. and I. M. Liu. 1972. "Factors affecting recognition thresholds of Chinese characters." *Chinese Journal of Psychology*, 14, 113-117 (in Chinese).
- Yu, B. and H. Cao. 1992. "A new exploration on the effect of stroke number in the identification of Chinese characters." *Acta Psychologica Sinica*, 24(2), 120-126 (in Chinese).
- Zhang, W. and D. Yang. 1987. "The effect of Chinese character stroke numbers on the capacity of STM." *Acta Psychologica Sinica*, 19(1), 79-85 (in Chinese).