Legibility of medicine labels: user studies on Chinese typefaces and font size for senior citizens in Hong Kong

Brian Sze-Hang Kwok

Keywords

Medicine label, information design, user-centered approach, usability, typography, Hong Kong

Abstract

This study examined three common Chinese typefaces to determine the optimal font and point size to increase the legibility of medicine labels for the elderly. The outcome shows different performance results in three tests. Heiti performed better in the reading test, whereas Kaiti performed well in the character recognition and searching a phrase tests. Songti performed adequately in all of the tests. Kaiti's resemblance to Chinese calligraphy significantly improved its performance. The findings indicate that simply increasing point size alone does not significantly improve legibility among the elderly.

Hong Kong medicine labels: current regulations and usability

According to Provision 22 (5) in Hong Kong's Chapter 138A Pharmacy and Poisons Regulations, medicines "shall be clearly labeled with instructions for use in English and in Chinese." Provision 38A (1) states that "no person shall sell or supply any medicine unless it is labeled with particulars printed so as to be clearly legible in English and Chinese, as to dosage and the route and frequency of administration." However, the regulations do not elaborate on what clearly labeled or legible means, or how to achieve such clarity using a specific typeface, color, leading space or font size. This lack of guidance results in a wide range of medicine label formats that spans the private pharmaceutical market.

Apart from the labels used in the private sector and at clinics, there are two types of medicine labels used in public hospitals, issued by the Department of Health and the Hospital Authority (HA), respectively (Figure 1). The HA dominates and manages over 161 public hospitals and outpatient clinics covering 18 districts in Hong Kong, and the labels they issue affect the majority of Hong Kong's citizens, especially senior citizens suffering from chronic illnesses who frequent the public hospitals for their medications. Hence, the labels issued by the HA are used in this study.

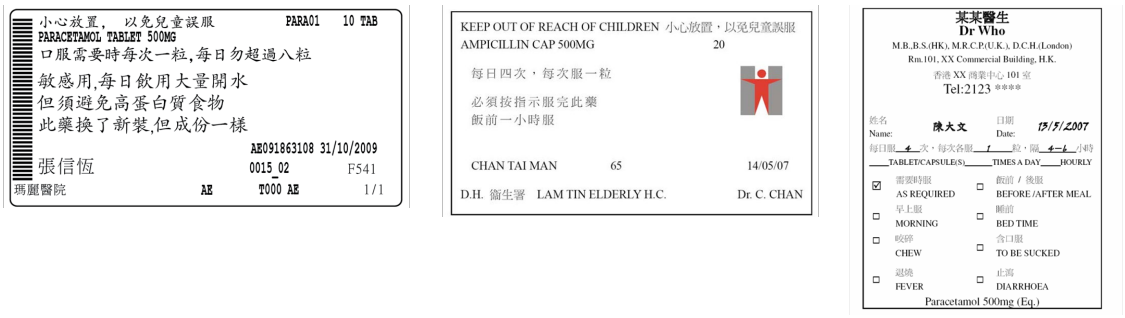

Figure 1. Hospital Authority (left), Department of Health (center) and private clinic center (right).

The HA's current labels mainly comprise seven major items of information about the medicine (The Hong Kong Medical Association, 2007): hospital's name, patient's name, date dispensed, drug name, dosage, method of administration, and precautions (Figure 2). Other information is also displayed, but in English and for internal use, such as abbreviated drug name and unit and hospital/department codes. The HA requires that all medicine information is shown clearly and legibly on labels for the public, but the current labels do not appear to fulfil these requirements, particularly for elderly users (people aged 65 or above). According to the results of a survey conducted by the Sik Sik Yuen (SSY) Ho Kin District Community Centre for Senior Citizens in November 2010,[1] nearly 97.7% of the 569 respondents were dissatisfied with the current medicine label design. About 96.1% (547) of the respondents requested a larger font size for legibility while 97.5% (555) preferred Arabic numerals over traditional Chinese characters for indicating the doses. Obviously, the current design of the medicine labels issued by the HA needs improvement as it appears not to answer to the users' needs. This reflects the importance of a user-centric design approach in information design.

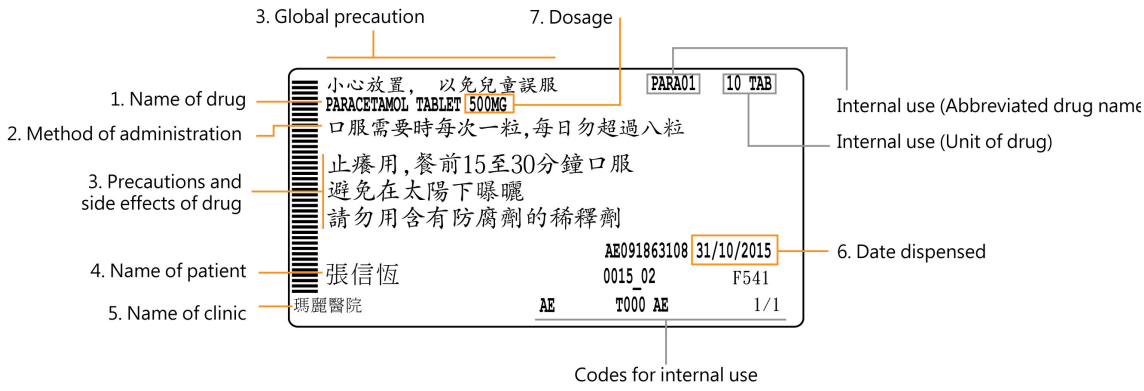

Figure 2. Seven major items and internal use information on Hospital Authority labels.

In its 2007 Global Age-friendly Cities Guide, the World Health Organization (WHO) also suggested improving the legibility of medicine labels. It stated that “font size on text materials, mainly hard copy ... is too small to read. Product labels and instructions, particularly for medications, are hard to decipher” (WHO, 2013). This is especially true for senior citizens (ibid, p. 63).

Research questions: font sizes, typefaces and legibility

In accordance with the WHO’s urgent call to develop age-friendly cities and the results of the aforementioned survey, the SSY recommended that the HA increase the current label’s font size (from 12.5 to 16 points or above) and use Kaiti (a Chinese typeface) to enhance the legibility level for the elderly. Based on the results of the SSY’s survey, the following research questions were raised:

- Is a 16-point font size feasible for application to the design of the medicine labels currently issued by the HA?
- If not, what is the optimal font size capable of striking a balance between the constraints of small medicine label space and senior citizens’ need for a higher level of legibility?
- Why is Kaiti considered a more legible typeface for seniors? Are there any alternative Chinese typefaces that could achieve the desired legibility?
- How does age influence legibility levels?

Even thought the SSY suggested that increasing the point size will improve legibility, Miles Tinker (1963) had stated that “[o]ptimal legibility of print is achieved by a typographical arrangement in which shape of letters and other symbols, characteristic word forms, and all other typographical factors such as type size, line width, leading, etc., are coordinated to produce comfortable vision and easy and rapid reading with comprehension.” The research has therefore not only considered point sizes when testing legibility. It has taken into account other factors such as negative space and the anatomy of Chinese characters.

Methods for usability test

There is a number of ways to measure legibility, such as reading tests, searching word tests, user preferences studies, comprehension tests and eye tracking. However, there is no single method that produces sufficiently useful results (Beier, 2012; Spencer, 1983). One of the objectives of this study is to examine the feasibility of using a 16-point font size on the medicine labels issued by the HA, as recommended after the 2010 SSY survey. The drawback of that survey was that it did not reflect users’ daily lives as it merely presented survey takers with a list

of dichotomous questions. Thus, user tests are needed to learn how medicine labels perform in the area of user comprehension.

Although the survey findings highly recommended the use of a 16-point or larger font size for greater legibility, this is actually impractical given the innate limitations of the size of the label, spatial arrangement and number of Chinese characters.[2] If the font size were increased from the current 12.5 points to 16 points, the 2 x 4-inch label would not be able to accommodate the required information. More importantly, mandating a 16-point or above font size would require a tremendous revision of the entire label system, which would seriously affect the HA. Considering these practical issues, this study maintains the size of the current medicine labels and uses as large a font size as possible for the prototypes in conducting user tests. One of the typefaces used in the test has significantly more negative space around its characters, however. To better understand the relation between font size and legibility, this typeface has been manually enlarged from its standard maximum font size to a larger face size with less negative space.

Chinese typography: styles, font sizes and strokes

Three common Chinese typefaces

In Chinese typography, there are three commonly used typefaces for text writing that are similar to those for the Latin alphabet: Songti, Heiti and Kaiti (Figure 3). The Songti style is similar to that of Roman letters, which have serifs at the ends of the strokes. Its vertical strokes are thicker than its horizontal strokes and the structure of the style appears as high thick–thin contrast. The Kaiti style resembles traditional Chinese calligraphy with brush strokes and is similar to that of a script font in Roman characters. The horizontal and vertical strokes possess similar weights and each horizontal stroke is gently tilted upward to the right. The Heiti style is similar to the san serif typeface used in Latin typography. It is monolinear, with little thick–thin contrast in the horizontal and vertical strokes. To find the most legible and accessible typeface for seniors, these three typefaces were used in the legibility tests.

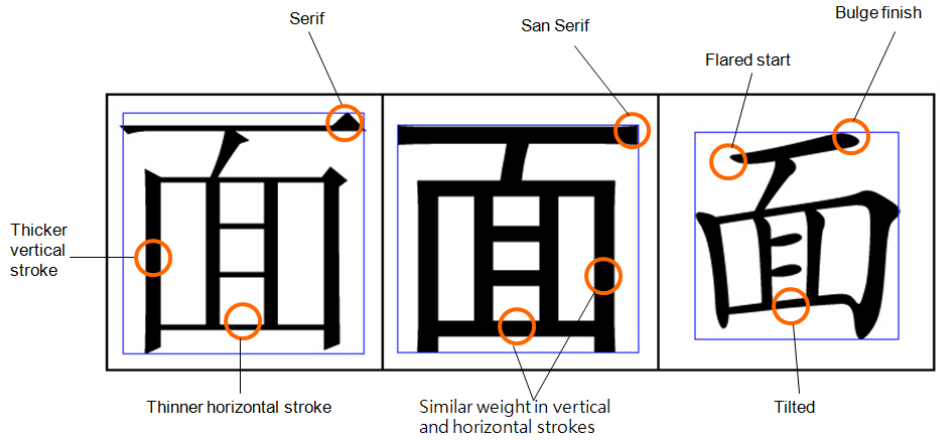

Figure 3. Songti (left), Heiti (center), and Kaiti (right) appear in the user tests. Each Chinese typeface possesses some unique features that are similar to the Roman letter, san serif and script fonts in Latin typography.

Font sizes

As mentioned earlier, the tests generate different character heights based on the different typefaces. As Figure 4 shows, the characters of the Kaiti style are shorter and smaller (2.9-4.2 mm) than those of Heiti (4.1-4.5 mm) and Songti (4.3-4.6 mm), although all three samples are set at 14 points. This is because the anatomy of the Kaiti characters is rooted in the center and developed inward, leaving more negative space around it. Heiti and Songti, in contrast, are developed outward and take up more space. However, a close examination of Figure 4 reveals subtle differences between Heiti and Songti. From edge to edge, the shadow of the Songti character takes up more vertical and horizontal space than that of the Heiti character. In fact, each Chinese character is uniquely influenced by both the stroke density and font skeleton.

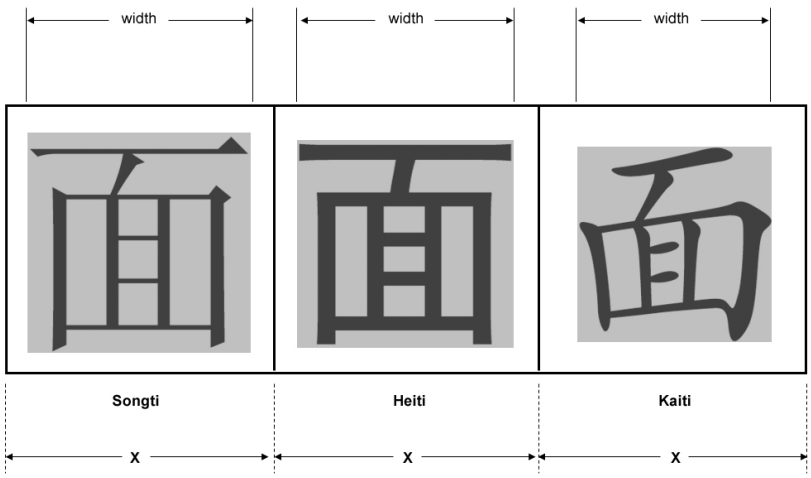

Figure 4. The anatomy of Kaiti (right) characters is the smallest of the three typefaces at the 14-point size. Kaiti characters are usually surrounded by more white space. The shadows indicate the EM space for each typeface. Although they are all set in a 14-point font size, the amount of white space surrounding them differs.

To eliminate any optical bias in testing the differences among font sizes, we adjusted the heights of the three typefaces so that all were optically similar (Figure 5). Specifically, the Heiti and Songti characters were set at 14 points with font heights of approximately 4.4 mm while the Kaiti characters were enlarged to 15.5 points to achieve an average height of 4.4 mm.

Kaiti	**Songti**	**Heiti**	**Kaiti**
14 points	14 points	14 points	15.5 points
(2.9-4.2mm)	(4.3-4.6mm)	(4.1-4.5mm)	(3.2-4.5mm)

Figure 5. The Kaiti characters are deliberately scaled up to align with the similar heights of the Songti and Heiti characters. The average height of Kaiti characters at 15.5 points is 4.4 mm while those of Songti and Heiti characters at 14 points are about 4.3 mm and 4.2 mm, respectively.

As mentioned, the current medicine label size issued by the HA cannot accommodate the required information when a font of 16 points or above is used, because the longest pharmaceutical information or precaution message on such labels is 18 Chinese characters per line, including punctuation. Moreover, three lines are reserved in the current label for the indication of precautions or side effects. To increase the point size and display up to 18 Chinese characters per line, 14 points is the maximum and optimal font size for all three of the Chinese typefaces studied. The Kaiti characters, however, could be increased to 15.5 points due to its inward anatomy (Figure 6). Given the current labels' space constraints, this study maintained the current medicine labels' size and changed the maximum font size to 14 points (approximately 4.4 mm), plus 2 points of leading space, in our test materials.

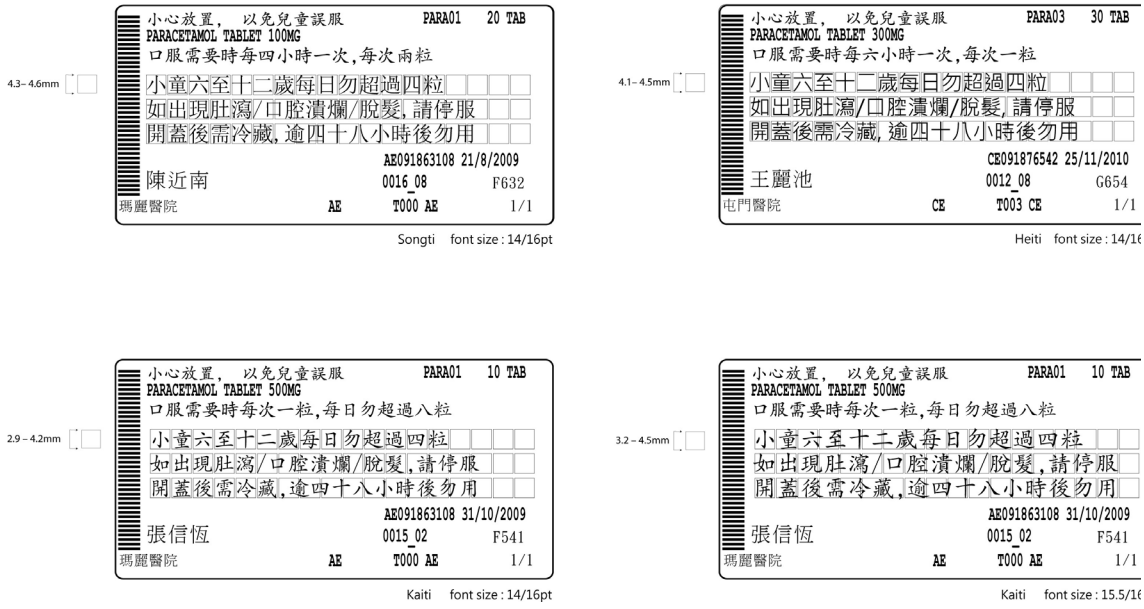

Figure 6. A 14-point font is the maximum and optimal size that can accommodate 18 Chinese characters in a line. The character heights at 14 points are about 4.3-4.6 mm for Songti (top left), 4.1-4.5 mm for Heiti (top right), and 2.9-4.2 mm for Kaiti (bottom left). The character height at 15.5 points is about 3.2-4.5 mm for Kaiti (bottom right).

Number of strokes

Other factors that may affect the legibility of Chinese characters are the number of strokes or their density. Chinese characters are composed of complex combinations that are completely different from the characters in the Latin alphabet, which are constructed from left to right in one dimension. In contrast, Chinese characters are presented in a two-dimensional way and can comprise up to three different components. The first is characters composed of a radical, the basic unit, of which there are 213. The second is characters composed of a radical and a stem. The third is complex characters composed of a radical combined with another radical and a stem (Figure 7). Even a simple Chinese character composed of a radical can be formed using a varied number of strokes, ranging from 1 to as many as 32 (Stallings, 1976; Gu, 1994).

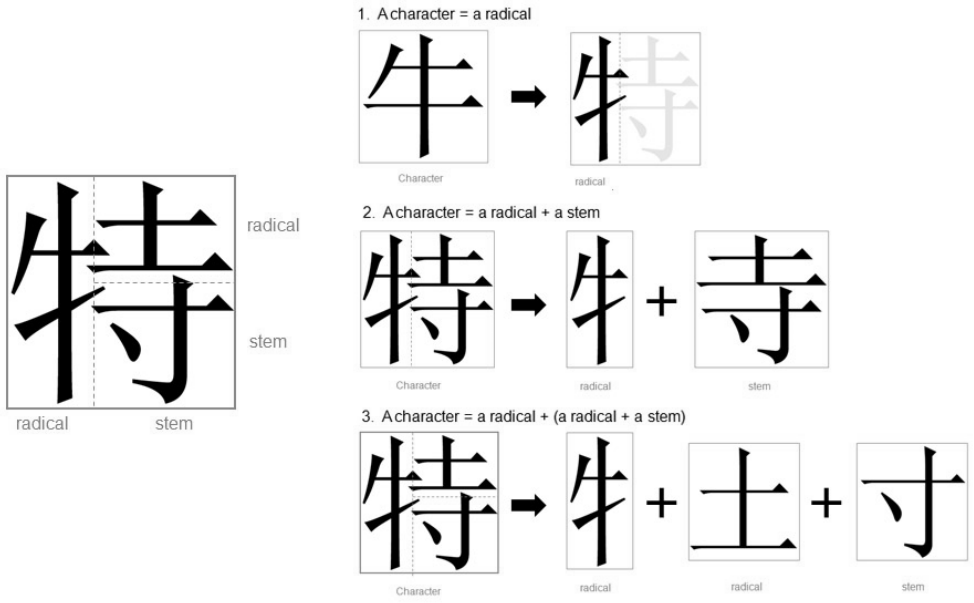

Figure 7. The basic unit of a Chinese character is composed of both a radical and a stem.

To achieve the same number of strokes in every test, 315 full sentences of pharmaceutical instructions from the HA's labels were used in the tests. After filtering out repeated characters, the remaining unique characters were classified into four groups according to the number of strokes (Table 1). Group 1 comprises 108 characters ranging from 1 to 6 strokes, group 2 comprises 200 characters ranging from 7 to 12 strokes, group 3 comprises 112 characters ranging from 13 to 18 strokes, and group 4 comprises 23 characters ranging from 19 to 24 strokes. The number of Chinese characters used in the medicine label information is 443, and all were used in our prototype designs produced for the legibility tests.

No. of strokes	No. of characters	Chinese characters
1 – 6 strokes group (108 characters)		
1	1	一
2	10	了七又九二力入人八十
3	10	丸士三大上才口上小下
4	30	午匹化公毛片太心方廿孔五牙中之勻六分升及天手月內止不氏水日勿
5	25	由孕包充汁它生本末外他奶出白甲必未四皮以加示只可用
6	32	先向肌肉份吐色早全各污衣任地合西血光有耳冰多在年成次如同至再存此
7 – 12 strokes group (200 characters)		
7	25	尿位低防皂肘抗沖肛完何足即低肚免吞含身但吸冷忌作每
8	36	近亞直炎況受芝和泠其肺並肥易空肩房底板門供乳肢析定命的始兩物放注或使於服
9	32	急室柔封持胃紅科限品背疣染面致保柚重查咬為便要指按洗胃星食度前後
10	30	凍破消粉捐徑能除效病高特胸配送酒疹秘唇浸脊原格高退倍個脈射時
11	44	敖被剪清動敏情痕途脫得理淡睌蛋連眼莖軟混涼控械現匙透接第混術毫趾乾爽將液停患雪陰處部粒啟
12	33	稍替診教發無減週視短陽掌腕斑腋腔換等量間貼童進痛鈣飲超最程稀須期開
13 – 18 strokes group (112 characters)		
13	38	暖塊痰經解煙會照腸跳暗腰跟零傷微道損暈感新樣歲準逾腹意裝碎搖逾腳溫置當過溶採
14	15	維漱腐聚輕管腿鼻瞳滴嘔製需與蓋
15	25	潔層潰暫模糊適瘡膝踝瀾影竊靛膠標遮質鋅髮臏膚調質請
16	17	澡磨器燃頰凝樽糖操機頸器剛燒頭餐靜
17	11	癮薄臨壓臂臀濕檢避應療
18	6	覆顎雖瀉櫃藏
19 – 24 strokes group (23 characters)		
19	6	懷壞曠邊類藥
20	4	癢嚴釋鐘
21	5	爛醬鐵續攝
22	1	囊
23	4	鑼體髓變
24	3	靈癮鑽
Total:	443	

Table 1. The 443 most frequently used characters on medicine labels issued by the Hospital Authority, classified into 4 groups according to the number of strokes.

Participants

Eighty Hong Kong Chinese seniors (36 female and 44 male) aged between 60 and 91, with a mean age of 73.4, volunteered to participate in this legibility testing. All of the participants were Cantonese speakers recruited through five centers for the elderly in Hong Kong. The selection criteria were the ability to recognize Chinese characters, and a lack of visual impairment that might keep them from completing the legibility testing. All of the participants were also required to have previous experience administering medicines and to not have participated in any other similar type of legibility test for medicine labels in the past.

Procedure

The legibility tests were conducted in the elderly centers between August and September 2012. Each participant was tested individually and the purpose of the study was explained at the

beginning of the testing. Testing took about 20 minutes for each participant. First, the participants were required to pass the pre-test, in which they read a paragraph that was randomly selected and extracted from the daily news, to ensure they did not have any difficulty recognizing Chinese text.

Each test and its instructions was printed on an individual A4-size sheet. The test sheets were only given to the participants who passed the pre-test. To compare the three typefaces and find which was the most legible for the elderly, three test worksheets with the content written in the three typefaces (Heiti, Songti, and Kaiti) were prepared. Each worksheet contained a reading test, a word recognition test, and a searching a phrase test. To complete the testing process, the participants were required to finish all three tests written in each typeface.

In the reading test, the participants were asked to read the three lines (72 characters) of Chinese text inside each label aloud (Figure 8a). The text had no meaning because it was not arranged in a logical order. Instead, it was randomly selected from the 4 pre-defined groups of strokes from the 443 characters most frequently used in pharmaceutical instructions (Table 1) to maintain the same number of strokes for each session. Thus, the participants had to pay extra attention and read each single word carefully. In the word recognition test, the participants were asked to locate and circle a specific character in a text with a color pen (Figure 8b). The specific character was shown at the bottom of the test sheet so that the participant could refer to it. In the searching a phrase test, the participants were required to locate a phrase (e.g., “Avoid alcoholic drink”) in the context of the medicine label, instead of illogical text (Figure 8c).

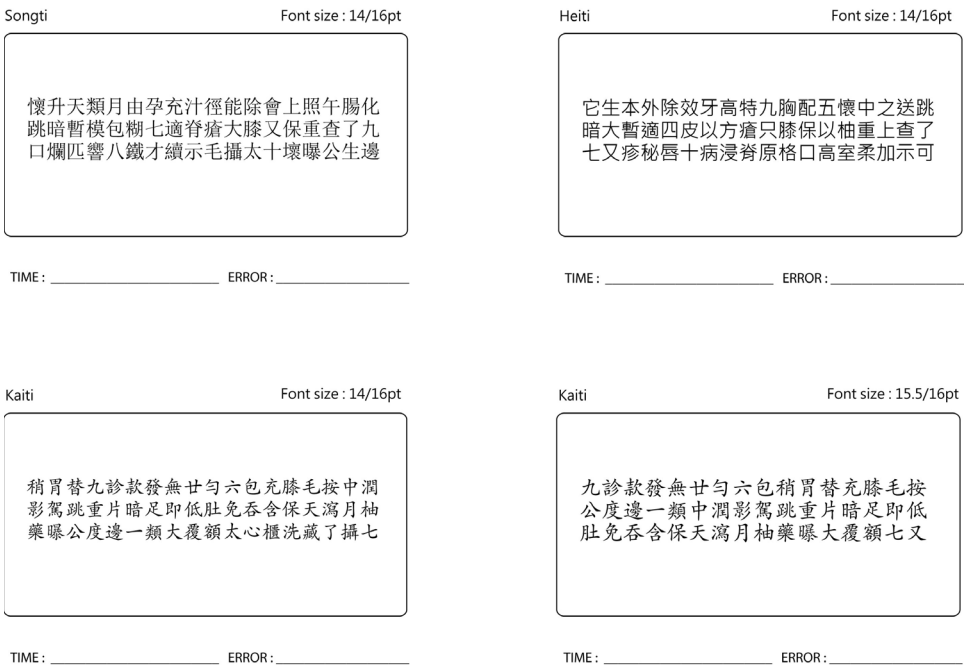

Figure. 8a Reading test. Each set of labels was given to the participant individually and they were asked to read the text aloud. Response times and errors were recorded.

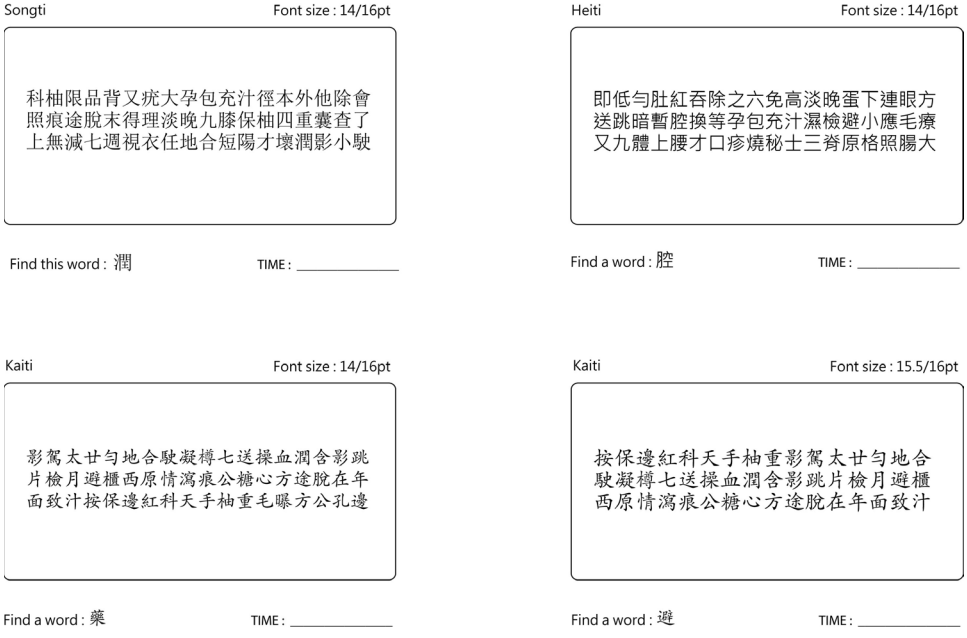

Figure. 8b Word recognition test. The participants were asked to locate a specific character in a text with a color pen. The specific character was shown at the bottom of the sheet for easy referral. Response times were recorded.

Figure. 8c Searching a phrase test. The participants were required to locate a phrase in the context of a medicine label. The specific phrase was shown at the bottom of the sheet for easy referral. Response times were recorded.

To avoid any bias that might stem from the arrangement of the typefaces, the order of the test sheets was randomly changed. Moreover, the characters in the word recognition and search a phrase tests and the locations of the words and phrases on the different prototypes were randomized. A researcher or facilitator briefed the participants and then observed their behavior during the testing. A research assistant recorded response times and errors made. Follow-up questions were asked post-testing to examine the participants' performance.

Results

Total mean of response times

The Kaiti typeface performed best among the elderly participants based on the overall mean response times (Table 2). However, there were subtle distinctions among the three typefaces based on the total aggregate search time in the legibility tests. The participants' mean response times were 70.89 seconds for Songti, 70.65 seconds for Heiti and 69.72 seconds for Kaiti to complete all of the tasks. The 14-point Kaiti characters had no advantage over their Songti and Heiti counterparts, therefore, but the 15.5-point Kaiti characters provided comparatively better legibility and accessibility, with a mean response time of 59.54 seconds to complete all of the tests.

Table 2. The mean and standard deviations of response times for the tests

	Songti 14 pt		Heiti 14 pt		Kaiti 14 pt		Kaiti 15.5 pt	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Reading test	57.32	21.64	53.60	17.77	58.33	19.84	45.31	15.50
Word recognition	8.18	7.14	11.93	9.25	7.24	4.41	10.03	8.32
Search a phrase	5.38	3.64	5.13	4.94	4.15	2.89	4.21	3.19
Total	70.89	25.62	70.65	23.72	69.72	22.57	59.54	21.68

N = 80, time in seconds

Age and legibility

To investigate whether any of the typefaces studied were more legible to a particular age group, we divided the 80 participants into 3 groups based on their ages: 23 in the 60- to 69-year-old group, 40 in the 70- to 79-year-old group, and 17 participants in the 80- to 91-year-old group.

It is not surprising that age affected to the response times. The older the participants were, the more time they took to complete the tests (Table 3). The results for the 14-point Heiti characters, for example, indicated that the aggregate mean response time was 58.88 seconds for the youngest group (60 to 69 years old), 72.07 seconds for the middle group (70 to 79 years old), and 83.22 seconds for the oldest group (80 to 91 years old). This pattern was also observed in the results relating to the other two typefaces. For example, with the 15.5-point Kaiti characters, the aggregate response times were 52.77, 58.58, and 70.96 seconds for the youngest, middle and oldest groups, respectively.

Table 3. Total response times (in seconds) for all tests in different age groups

	Songti 14 pt			Heiti 14 pt		
	Mean	Coefficient	Standard error	Mean	Coefficient	Standard error
60-69	46.45	Control group		58.88	Control group	
70-79	75.86	19.41	6.33**	72.07	13.19	5.85*
80-91	78.71	22.26	7.73**	83.22	24.34	7.15**
N = 80	R-squared = 0.13			R-squared = 0.13		

*p < 0.05
**p < 0.01

	Kaiti 14 pt			Kaiti 15.5 pt		
	Mean	Coefficient	Standard error	Mean	Coefficient	Standard error
60-69	61.55	Control group		52.77	Control group	
70-79	70.07	8.52	5.73	58.58	5.8	5.49
80-91	79.96	18.41	7.01**	70.96	18.19	6.7**
	R-squared = 0.08			R-squared = 0.08		

* p < 0.05
**p < 0.01

The standard error results showed that the test outcomes of the Songti and Heiti characters were statistically significant in the middle group (70 to 79 years old), whereas those of the 14- and 15.5-point Kaiti characters were insignificant (p = 0.05). This meant that for this group only the mean response times relating to the Songti and Heiti characters could be compared. For the Songti characters, the middle group's mean time was 75.86 seconds and that for the Heiti characters was 72.07 seconds. The coefficient values further revealed that the participants took 13.19 seconds more than the youngest (control) group when testing the Heiti characters and 19.41 seconds more when testing the Songti characters. This indicates that, in the middle group, the Heiti characters performed better than the Songti characters.

Among the oldest group (80 to 91 years old), all outcomes were statistically significant meaning that for this group the response times relating to all four typefaces could be compared. The participants' mean response times were 78.71 seconds for Songti, 83.22 seconds for Heiti, and 79.96 and 70.96 seconds for 14- and 15.5-point Kaiti. Regarding the coefficients, the participants took an additional 22.26 seconds for Songti, 24.34 seconds for Heiti, and 18.41 and

18.19 seconds for 14- and 15.5-points Kaiti. These results first of all indicate that different from those of the middle group, between Songti and Heiti, Songti came out better. More importantly, however, the results indicate that although the prototype using the 15.5-point Kaiti generated the shortest response time, the difference in response time between the 14- and 15.5-point Kaiti was particularly minimal.

Results of the reading test

In the reading test, 80 participants were asked to read 72 pharmaceutical Chinese characters pre-set in a non-logical, random and meaningless arrangement aloud. The 15.5-point Kaiti characters produced the shortest response times for all three age groups (Table 4), and Heiti produced the shortest times among the three 14-point typefaces with mean response times of 45.03, 55.21, and 61.3 seconds for the youngest, middle, and oldest groups, respectively. Based on the participants' feedback, Heiti's font style presented a monolinear weight in the characters' horizontal and vertical strokes, which created an optically darker effect that enabled the elderly to read more effectively when scanning the texts.

Table 4. Response times for reading test in different age groups

	Songti 14 pt			Heiti 14 pt		
	Mean	Coefficient	Standard error	Mean	Coefficient	Standard error
60-69	47.23	Control group		45.03	Control group	
70-79	60.9	13.68	5.17*	55.21	10.17	4.44*
80-91	62.59	15.37	6.68*	61.39	16.35	5.42**
N = 80	R-squared = 0.09			R-squared = 0.09		

	Kaiti 14 pt			Kaiti 15.5 pt		
	Mean	Coefficient	Standard error	Mean	Coefficient	Standard error
60-69	51.61	Control group		40.37	Control group	
70-79	59.11	7.5	5.09	45.54	5.17	3.98
80-91	65.58	13.96	6.22*	51.45	11.08	4.86*
	R-squared = 0.06			R-squared = 0.06		

* $p < 0.05$
** $p < 0.01$

Although between the three 14-point typefaces the response times for Heiti were the shortest in the reading test, Heiti produced the highest number of reading errors (163 errors, see Table 5).[3] The implication is that the lack of contrast in the Heiti style hinders character recognition, particularly when two similarly structured characters are juxtaposed. The percentage of reading errors in Heiti was 32.1% (35 out of 109 total reading errors) in the youngest group, 32.3% (85 out of 263) in the middle group, and 37.1% (43 out of 116) in the oldest group. The total percentage of reading errors for Heiti was 41.4% (163 out of 488 total reading errors), followed by Songti (24.4%; 119 out of 488). The Kaiti style in both font sizes (14 and 15.5 points) scored well and had the lowest frequency of reading mistakes among the three age groups. The 14- and 15.5-point Kaiti had error rates of 22.9% (112 out of 488) and 19.3% (94 out of 488), respectively. In summary, although the Kaiti style did not perform best in the reading test, it did have the lowest frequency of reading errors among the elderly.

Table 5. Number of errors in the reading test

Age	Songti 14 pt	Heiti 14 pt	Kaiti 14 pt	Kaiti 15.5 pt	Total errors in age group
60-69	23	35	25	26	109
70-79	69	85	61	48	263
80-91	27	43	26	20	116
Total errors in each typeface	119	163	112	94	488

Results of the word recognition test

In the searching a phrase test, Heiti had the longest response times among the three typefaces and across the three age groups: 8.92, 12.97, and 13.54 seconds in the youngest, middle, and oldest groups, respectively (Table 6). This implies that the Heiti was the most difficult for the elderly to read. This finding differed from the first test, in which Heiti performed quite well, suggesting that it allowed the elderly to read the texts easily. However, in the second test, the participants had to pay close attention to a character's features and find it among the surrounding text. The Heiti style's monolinear strokes and lower contrast may have resulted in the participants taking more time to recognize the character accurately. The Songti style's high thick–thin contrast between vertical (thicker) and horizontal (thinner) strokes resulted in the participants having

difficulty seeing the thinner horizontal strokes against the white background and under dim lighting. Thus, Songti had the second-shortest response times.

In contrast, the Kaiti style's strong serif features that resemble traditional Chinese calligraphy made its characters quick and easy to identify. Surprisingly, however, the response time for the 15.5-point Kaiti characters was longer than that for their 14-point counterparts in the character recognition test. The text presented in 15.5- and 14-point Kaiti characters had mean response times of 9.16 and 6.68 seconds, respectively, in the youngest age group. The 15.5-point Kaiti had also longer response times than the 14-point Kaiti in the middle and oldest age groups.

Table 6. Response times for word recognition test in different age groups

	Songti 14 pt			Heiti 14 pt		
	Mean	Coefficient	Standard error	Mean	Coefficient	Standard error
60-69	5.51	Control group		8.92	Control group	
70-79	8.693	3.42	1.83	12.97	4.05	2.39
80-91	10.02	4.51	2.24*	13.54	4.61	2.92
N = 80	R-squared = 0.06			R-squared = 0.04		

	Kaiti 14 pt			Kaiti 15.5 pt		
	Mean	Coefficient	Standard error	Mean	Coefficient	Standard error
60-69	6.68	Control group		9.16	Control group	
70-79	6.94	0.26	1.15	8.96	-0.2	2.14
80-91	8.73	2.05	1.41	13.72	4.56	2.62
	R-squared = 0.03			R-squared = 0.05		

* $p < 0.05$
** $p < 0.01$

Results of the searching a phrase test

The findings of the searching a phrase test showed that the Kaiti typeface had the best response times for all three age groups (Table 7), particularly the 14-point Kaiti characters: 3.26, 4.02, and 5.65 seconds in the youngest, middle, and oldest groups, respectively. Comparatively, the Heiti style had the longest response time (8.3 seconds) in the oldest group and the Songti style had the middle response time among the three typefaces studied. The 15.5-point Kaiti

characters had shorter response times than the 14-point Kaiti characters in the first test, but longer times in the second test and mild differences in the third test. As mentioned, the participants preferred the Kaiti typeface for its higher legibility due to its similarity to traditional Chinese calligraphy. Moreover, most elderly patients have been reading the HA's medicine labels in their current format for years, and most of them have encountered common chronic illnesses such as high blood pressure, cholesterol, and diabetes. Such users take medicines daily and go to HA hospitals for regular check-ups and prescription refills. Due to this familiarity with the label design configuration and information organization, which was almost unchanged in the third test, the varied font sizes had little effect on the participants' response times. Thus, the 14- and 15.5-point Kaiti characters revealed no major differences in response times and performed equally well in terms of character recognition.

Table 7. Response times for searching a phrase test in different age groups

	Songti 14 pt			Heiti 14 pt		
	Mean	Coefficient	Standard error	Mean	Coefficient	Standard error
60-69	3.71	Control group		4.93	Control group	
70-79	6.03	2.31	0.92*	3.9	-1.03	1.22
80-91	6.1	2.39	1.13*	8.3	3.37	1.50*
N = 80	R-squared = 0.08			R-squared = 0.12		

	Kaiti 14 pt			Kaiti 15.5 pt		
	Mean	Coefficient	Standard error	Mean	Coefficient	Standard error
60-69	3.26	Control group		3.24	Control group	
70-79	4.02	0.76	0.73	4.08	0.84	0.81
80-91	5.65	2.4	0.9**	5.79	2.55	0.99*
	R-squared = 0.09			R-squared = 0.08		

* $p < 0.05$

** $p < 0.01$

Discussion

There is much room for improvement. Due to the lack of any previous research and legibility tests of medicine labels in Chinese society, the methods of Western legibility tests and screen-based Chinese legibility tests were used as references for the main framework of this

research. However, a number of other factors affected the test results and should be considered in future research:

- **Avoid high familiarity**. The participants with chronic illnesses make frequent trips to the HA hospitals for check-ups and prescription refills. As a result, they are very familiar with the labels issued by the HA and know where to look for information, even with altered typefaces and font sizes as long as the information organization and visual hierarchy have not been changed.

- **Real-life prototypes**. The medicine label prototypes in this study were printed on A4-size sheets, which did not reflect the real-world experience gained when administering daily medicine. It would be preferable for further tests to use real medicine labels on varied containers and bags.

- **Real-life testing.** All of the user tests in this study were conducted in centers for the elderly, and thus did not truly reflect real-life situations that are likely to involve factors such as dim lighting and lack of supervision. Thus, it is necessary to conduct real-life tests to generate more accurate results.

Conclusion

Based on the results of the 2010 SSY survey, the 16-point Kaiti typeface was recommended to increase the legibility of medicine labels for elderly users. However, this recommendation was not based on evidence from testing various typefaces and font sizes among elderly patients. This study therefore pursued increased legibility by comparing different Chinese typefaces and font sizes in a user-centered approach, examining which performs best. The combination of the reading, word recognition, and search a phrase tests, allowed a more elaborate understanding of the legibility of the selected typefaces and font sizes as presented in the medicine label prototypes.

The overall results (the total mean response times) indicated that the 14-point Songti, Heiti, and Kaiti styles performed similarly well, while the 15.5 Kaiti style had a significantly better overall performance. However, when taking a closer look, none of the typefaces dominated in all three tests. The san serif of the Heiti style, for instance, performed well in the reading test but poorly in the word recognition and searching a phrase tests. Heiti also produced the highest number of reading errors among the three typefaces. Compared with Kaiti and Heiti, further, Songti's performance was middling in all of the tests. Even though specifically the 15.5-point Kaiti presented a significantly shorter mean response time, compared with the 14-point Kaiti it did not

perform better, despite its size advantage, in the word recognition and searching a phrase tests. It also caused more errors in the reading test. This indicates that simply increasing the font size does not necessarily improve legibility among elderly users.

In summary, although the results of this study are aligned with the SSY's typeface recommendation, it appeared that different typefaces perform well in different tests while looking at font size alone would disregard other typographical factors. Compared with the SSY's survey, in this research participants were more accurately presented with label prototypes based on real-life examples and allows therefore a more user-focused outcome. Finally, the test has confirmed that age influences response times and reading errors, suggesting that legibility of medicine labels should be taken seriously, especially with regards to the older members of the society.

References

Beckman, A. G. K., Parker, M. G., & Thorslund, M. (2005). Can elderly people take their medicine? *Patient Education and Counseling*, *59*, 186–191.

Beier, S. (2012). *Reading letters: Designing for legibility.* Amsterdam: BIS Publishers.

Cai, D. et al. (2001). The legibility threshold of Chinese characters in three-type styles, *International Journal of Industrial Ergonomics 27*, 9–17.

Carliner, Saul et al., eds. (2006) *Information and document design*. Amsterdam: John Benjamins Publishing.

Census and Statistics Department. (2006). *Hong Kong 2006 population by-census thematic report : Older persons*. Retrieved from http://www.censtatd.gov.hk/freedownload.jsp?file=publication/stat_report/population/B11200532006XXXXB0100.pdf&title=Hong+Kong+2006+Population+By-census+Thematic+Report+%3a+Older+Persons&issue=-&lang=1&c=1

Census and Statistics Department. (2006). *Cognitive function of older persons*. Retrieved from http://www.censtatd.gov.hk/freedownload.jsp?file=publication/stat_report/social_data/B11302282006XXXXB0100.pdf&title=Thematic+Household+Survey+Report+-+Report+No.28&issue=-&lang=1&c=1

Census and Statistics Department. (2010). *Hong Kong population projections 2010-2039*. Retrieved from http://www.censtatd.gov.hk/freedownload.jsp?file=publication/stat_report/

population/B1120015042010XXXXB0100.pdf&title=Hong+Kong+Population+Projections+2010-2039&issue=-&lang=1&c=1

Department of Health, Hospital Authority and Social Welfare Department. (2007). *The residential care homes (elderly persons) ordinance: An operational manual on drug management*. Retrieved from http://www.info.gov.hk/elderly/publications/drug/drug_manual.pdf

Department of Justice. (2016). *Pharmacy and poisons regulations*, Chapter 138A, P.19. Retrieved on 25 January 2016, http://www.legislation.gov.hk/blis_pdf.nsf/6799165D2FEE3FA94825755E0033E532/98C7D17989BE0FD1482575EE0043A3C8/$FILE/CAP_138A_e_b5.pdf

Dowse. R., & Ehlers, M. (2005). Medicine labels incorporating pictograms: Do they influence understanding and adherence? *Patient Education and Counseling 58*, 63–70.

Editorial. (2008). Dying from drugs--at home, *The Lancet 372*, 419.

Fisher, P. & Sless, D. (1990). Information design methods and productivity in the insurance industry, *Information Design Journal 6*(2), 103–129.

Frascara, J., & Ruecker, S. (2007). Medical communications and information design, *Information Design Journal 15*(1), 44–63.

Hong Kong Medical Association. (2007). *Good dispensing practice manual (2nd edition)*. Retrieved on 25 January 2016 http://www.hkma.org/download/others/Good%20Dispensing%202007.pdf

International Institute for Information Design. (2016). Retrieved on 25 January 2016, http://www.iiid.net/Information.aspx

Medicines and Healthcare products Regulatory Agency. (2014). *Best practice guidance on labeling and packaging of medicines*. London: MHRA. Retrieved on 25 January 2016, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/474366/Best_practice_guidance_labelling_and_packaging_of_medicines.pdf

Rogers, D., Shulman A., Sless, D., & Beach, R. (1995). *Designing better medicine labelling report to PHARM*, Canberra: Communication Research Institute of Australia.

Shulman, A. & Sless, D. (1992). Product labeling regulation: Can it lead to good information design? *Designing information for people*. Canberra: Communication Research Press, 63–97.

Sless, D., & Wiseman, R. (1994). *Writing about medicines for people: Usability guidelines and glossary for consumer product information*. Canberra: Department of Health and Human Services.

Sless, D. (2004). Designing public documents. *Information Design Journal 12*(1), 24–35.

Spencer, H. (1983). *Pioneers of modern typography.* Revised edition. Cambridge: M.I.T. Press.

Twyman, M. (1979). 'Schema for the study of graphic language' in Kolers, Paul A. et al., eds. *Processing of visual language*, volume 1. New York: Plenum Press, 117–150.

Van der Waarde, K. (2006). Visual information about medicines for patients. *Designing effective communications: creating contexts for clarity and meaning*, New York, NY: Allworth Press, 38–50.

Walker, S., & Barratt, M. (2008). Information design. Retrieved on 16 February 2008, http://www.designcouncil.org.uk/en/About-Design/Design-Disciplines/Information-Design-by-Sue-Walker-and-Mark-Barratt/

Webb, J., Davis, T.C., Bernadella, P., Clayman, M. L., Parker, R. M., Adler, D., & Wolf, M. S. (2008). Patient-centered approach for improving prescription drug warning labels, *Patient Education and Counseling 72*, 223–449.

World Health Organization. (2007). *Global age-friendly cities guide*. Retrieved on 25 January 2016. http://www.who.int/ageing/publications/Global_age_friendly_cities_Guide_English.pdf

World Health Organization*. Ageing and life course*. Retrieved on 4 February 2011. http://www.who.int/ageing/en/

文湜. (15 July 1997). 〈五成長者有食錯藥問題〉《松柏之聲》
Retrieved from http://www.thevoice.org.hk/v0239/025.htm

傳媒報導. (June 2008). 〈大包小袋藥物易溝亂重複服用有危險〉《松柏之聲》Retrieved from http://www.thevoice.org.hk/electrical/2008/200806/extrapage_jun2.pdf

明報專訊. (30 December 2010). 〈團體促統一藥袋標貼 減長者食錯藥風險〉《明報》Retrieved from http://hk.news.yahoo.com/article/101229/4/lyyu.html

Notes

[1] Sik Sik Yuen is a renowned religious charity organization that provides a wide range of medical, educational, and community services for the elderly. The survey was conducted in seven districts of Hong Kong during the period from November to December 2010, including Yuen Long, Sham Shui Po, Tuen Mun, and others. The survey collected data from 928 respondents, 569 of which were elderly and 359 who were non-elderly.

[2] The standard label is 4 inches wide by 2 inches high for easy printing and use in all of the Hospital Authority's institutions. Its design accommodates internal use and patients' information. The label space is programmed to display a maximum of three lines of text for precautions and one line for dosages and method of administration. The label width can accommodate a maximum of 18 Chinese characters, including punctuation.

[3] Reading errors were recorded when the participants engaged in any of the following: character misrecognition, word skipping, word mispronunciation, or wrong association of another phrase.