A Usability Testing of Chinese Character Writing System for Foreign Learners

Manlai You and Yu-Jie Xu

123 University Road, Section 3, Douliu, Yunlin 64002, Taiwan
{youm,g9930807}@ yuntech.edu.tw

Abstract. Currently, the study of Chinese has become increasingly popular in the world. However, not every non-native Chinese speaker learning Chinese can have formal guidance from qualified instructors. Xi-Zi-e-Bi-Tong (習字e筆通) is one of the systems for writing Chinese characters and is used by the Ministry of Education's E-innovation School and E-bag Experimental Teaching Program in Taiwan. It was developed for native Chinese speaking elementary school students. However, foreign learners come from a variety of cultural backgrounds and ages, so this study looks at the efficacy of this system for these types of students. As a case study, this research performed a usability testing with this system in order to identify what typical usability problems may exist in off-the-shelf products for foreign learners. The usability testing is with thinking aloud, in order to avoid the frustration of participants during tasks, the combination of coaching method to provide help appropriately. The subjects for this research were six foreign students, they came from different cultural backgrounds and all were unfamiliar with Chinese. It was hoped that testing this level of learners would make it easier to ascertain the usability problems of the system. Each was given six tasks associated with system manipulation that was related to the research purpose, and the tasks were designed in accordance with the instructions. When they had completed all the tasks, in order to measure the satisfaction of the system, they were asked to immediately fill out a questionnaire on user interaction satisfaction (QUIS). The problems they encountered in the test can be categorized, in accordance to the interactive design principles and concepts, into: mental models, visibility, feedback, and control. This study can be used as a reference for the redesigning of a program to teach the writing of Chinese characters.

Keywords: Chinese Characters, Interactive Interface, Chinese Learning, usability testing.

1 Introduction

Currently, the study of Chinese has become increasingly popular in the world. However, not every non-native Chinese speaker learning Chinese can have formal guidance from qualified instructors. Thanks to the rise of learning technology, this deficiency can now be partially resolved by well-designed interactive interfaces. The role of Chinese characters is very important in Chinese history. But it is a very

A. Marcus (Ed.): DUXU/HCII 2013, Part II, LNCS 8013, pp. 149–157, 2013.

arduous task to learn how to write Chinese characters for foreign learners. Xi-Zi-e-Bi-Tong (習字e筆通) is one of the systems for writing Chinese characters and is used by the counseling program of Taiwan's E-innovation School and by the E-bag Experimental Teaching program. It was developed for native Chinese speaking elementary school students. Since foreign learners come from a variety of cultural backgrounds and ages, this study looks at the efficacy of this system for these types of students.

Chinese characters are not only communication symbols but also a form of art in China—calligraphy. Presently, though, most modern people are likely to use computers to write Chinese characters rather and writing them by hand. However, according to some research, writing Chinese characters by hand is helpful (Li, 2009; Tan et al., 2005). The reason is that the motor function of the hand is connected with the brain and research shows this is cognitively helpful for beginning learners (Mangen & Velay, 2010; Tan et al., 2005). With the current trend for digital learning have come many products for Chinese learning—most are found on the internet. A lot of detail should be considerate of Chinese characters writing learning system design, those include interactive design between the user and the system not only using feel about it but also the system target - learning writing character.

Usability testing can help researchers to understand the real conditions of the product that is being used, and it can also discern how easy the product is to use and how effective it is in teaching the target language. Rogers, Sharp & Preece (2011) have pointed out that the goals of usability include effectiveness, efficiency, safety, utility, learnability and memorability. The principle of design mix from several fundamental knowledge of theories, experience and common sense. It is important for researchers to understand how users of this interactive technology feel about the process and what they are seeing. The design principles that are most used include visibility, feedback, limitation, mapping, mental models, consistency and affordance (Rogers, Sharp & Preece 2011; Norman, 1988).

This study evaluates the typical work that users learn with the system, and both the thinking aloud method and the coaching method are selected. Thinking aloud demands that the user speak aloud what they are thinking and what they want to do when they operate the task. Thus, by the recordings and notes, we can analyze the operation and the voice records after the usability testing, which can help the researcher understand the user's problems and their thoughts (Nielsen, 1993). The coaching method guides the user through problems during the processing of doing tasks (Mack & Burdett, 1992).

2 Method

This study includes two parts, first is the task design for the Chinese character writing system. Second is the usability testing with thinking aloud and coaching, and after the experiment is finished, participants will fill out a QUIS Questionnaire to assess the user interface. The Questionnaire for User Interaction Satisfaction (QUIS) is a tool developed by a multi-disciplinary team of researchers in the Human-Computer Interaction Lab (HCIL) at the University of Maryland at College Park.

2.1 Task Design

Xi-Zi-e-Bi-Tong was selected as a case study. This system is composed of lesson selections, writing character learning, writing character exercises and writing character tests (Fig. 1). Users can select the default content, which provides both strokes and radical indexing for selection to help users input by themself on the lesson selection. The writing character learning part provides writing demonstrations, and users can follow the demonstration to practice writing. The system evaluates the writing results and provides the right stroke order and neat writing, and then gives a score. The writing character exercise is similar to ones found in a traditional Chinese writing book, and users can continue to practice writing Chinese characters, and each writing results will be graded and recorded. The writing character test will quiz learners on the learning content in order to confirm the learner's acquisition outcomes. The usability of writing character learning section is evaluated in this study.

The tasks were designed in accordance with the instructions of writing character learning, and each task contains operation actions for several function keys. They are described as follows:

1. Please play the pronunciation of each character, and find out if it is a polyphone or not. This mission evaluates how the vocabulary table, pronunciation playing and polyphone detecting ability are used.
2. Please play the writing demonstration of the steps of the strokes for each character, and then speak out the last stroke. This task evaluates the stroke tracing function for the writing demonstration that shows the steps of the character's strokes.
3. Please play the writing demonstration showing the continuous steps for writing each character, and then pause after the fourth stroke, and then again continue. This task evaluates the play and the pause functions of the system's writing demonstration.
4. Please write each character once with and without using the grid individually, and then evaluate and speak out the results. This task evaluates the functions for the grid, the evaluation and the erasing.
5. Turn on the character tracing function, and then select tracing with each stroke and write each character once individually. Then, evaluate the results and speak them aloud. This task evaluates how the user performs the functions for the stroke tracing steps, evaluation and turn off.
6. Turn on the character tracing function, then select tracing for whole characters and write each character once individually. Then, evaluate the results and speak them aloud. This task evaluates how the user performs the functions for the continuous strike tracing option, evaluation and turn off.

Participants were asked to sequentially perform pre-set tasks in this study to evaluate Xi-Zi-e-Bi-Tong's usability testing with the think aloud and coaching options. The entire process was designed to help the researchers understand the errors users made while operating the interface. After the experiment, the QUIS questionnaire with a 5-point scale was used to evaluate whether the users found the system satisfactory.

Fig. 1. Writing character learning of Xi-Zi-e-Bi-Tong

Table 1. The basic information of subjects

	S1	S2	S3	S4	S5
Age	8	8	8	9	8
Sex	F	M	F	F	F
Similar systems experience	No	No	No	No	No
Computer experience	No	Yes	Yes	Yes	Yes

2.2 Subjects

This study focused on usability testing of the interface design for a Chinese characters writing system. Presently, most systems are designed for beginning students learning Chinese characters. The subjects for this research were six foreign students studying basic Chinese at National Taichung University of Education's Language Center, and each of them performed six pre-designed tasks. The number of subjects is admittedly small. However, our findings using usability testing showed that 75% of problems could be ascertained (Rogers, Sharp & Preece, 2011; Nielsen, 1993). The beginners came from different cultural backgrounds and all were unfamiliar with Chinese. It was hoped that testing this level of students would make it easier to ascertain the usability problems with the system.Before the start of the experiment, participants finished reading the notice of informed consent and then were given oral instructions and finally gave their signatures.

2.3 Device and Process

Lenovo X220t Tablet PC with stylus served as the test tool, and it provided a similar feeling to using a real pen and paper, so that the eyesight of participants could look the pen tip in the process of writing, and avoid the problems of using a traditional handwriting Tablet which splits the visual and writing space leading to visual gap

problems. During the experiment, a webcam was set up above the tablet PC, the recording of the screen and the area for the user's hands.

The experiment was described and explained prior to starting the experiment and the subject then gave their consent to participate in the experiments. An observer sat on the side of the participants to observe and record observations, and also provided help when the participants had difficulties. Participants explored the system for about 10 minutes, and then they started the tasks which took 30 minutes. When they had completed all the tasks, in order to measure the satisfaction of the system, they were asked to immediately fill out a questionnaire on user interaction satisfaction (QUIS). This questionnaire was developed to assess the user's responses with the user interface by the Computer Interaction Laboratory at the University of Maryland. This was then followed by unstructured interviews, which were used in order to understand the subjective feelings of the participants about the system in actual use.

3 Result and Discussion

3.1 Performance Evaluation

The performance evaluation methods for this study included the following: 1) the time to complete all the tasks, 2) the time to perform a task and its error rate, and 3) a recording of the scores of the writing characters tasks. The calculation of time to complete the task continued until it was completed, and when the participants wanted to tap a function but failed, then this counted as one error, and the same error would be calculated cumulatively.

Analysis of experiment's results showed that the error times for the control tasks (Task1 – Task 3) were higher than they were for the writing tasks (Task 4 – Task 6). The results of one-way ANOVA showed that there were no significant differences between the two groups. The writing method is divided into non-tracing (Grid, and Non-Grid) and tracing (Tracing Stroke, and Tracing Whole Characters) groups, and it was ascertained that there was a significant difference between the two group with ANOVA, and the scores of written results with tracing were better than with non-tracing. Because tracing is a writing guidelines in the process for participants. The task processing performances are affected by the previous tasks, and the Latin grid was not introduced to avoid any relative affects. The reason for this is that each task was designed to be an independent event. However, possibly trial and error could have had some impact following the pre-learning tasks.

3.2 Questionnaire Analysis

The QUIS questionnaire was used by this study. It is a modified user interactive satisfaction questionnaire from the Human-Computer Interaction Lab (HCIL) at the University of Maryland at College Park. The study extracted past experiences, overall user reactions, screen, learning, on-line tutorials and multimedia from the original QUIS questionnaire. The researchers changed some items on the questionnaire to

Table 2. Performance evaluation

	Task 1		Task 2		Task 3		Task 4		Task 5		Task 6	
	CT*	NE**	CT*	NE**	CT*	NE**	CT*	NE**	CT*	NE**	CT*	NE**
S1	6'19"	27	6'21"	33	2'16"	0	3'30'	7	1'59"	7	1"11"	0
S2	6'10"	5	9"54"	3	6'30"	5	5'50"	2	3'24"	6	2'14"	0
S3	1'53"	3	2'53"	0	2'59"	1	8'36"	0	5'09"	12	3'09"	4
S4	1'37"	2	3'45"	0	2'29"	0	3'53"	2	2'18"	0	1'48"	1
S5	2'30"	0	3'35"	1	5'20"	3	9'18"	1	3'53"	0	3'58"	0
S6	3'25"	10	3'07"	2	2'41"	1	4'12"	5	2'50"	1	3'19"	13
Mean		7.8		6.5		1.7		2.8		4.3		3.0
SD		9.99		13.03		1.97		2.64		4.84		5.14

* CT (Completion time) **NE (The number of errors)

Table 3. Writing score

	Grid	Non-Grid	Tracing stroke	Tracing whole character
S1	84.3	86.5	96.5	97.3
S2	81.5	84.8	98.3	98.3
S3	81.5	61.8	99.0	98.8
S4	83.0	93.3	98.3	96.5
S5	84.0	86.8	97.3	98.8
S6	74.5	76.3	94.0	98.0
Mean	81.4	81.6	97.2	98.0
SD	3.61	11.12	1.81	0.91

meet the specific needs of this research. For example, the nine-point scale was revised to five-point scale. The participants showed positive satisfaction with this system, so just low average grades are described for reference.

The system flexibility relative assessment, with an average score of 3.5; adequacy of the amount of information on the screen, with an average score of 3.7; through attempts to explore the new features, the average score was 3.7; information succinct and pointed out that the focus, the average score was 3.5. The average scores for these items are low on the questionnaire. Possibly this indicates that participants felt that the learning system still presented a certain degree of difficulty for them.

The participants' feedback comments as follows: 1) participants expected the system to be able to provide more learning information, for example, the pinyin, words, sentences, and so on to help them remember the Chinese character information; 2) the direction of the strokes for writing Chinese characters should not only provide a writing demonstration, but should also provide practice writing; 3) the speed of the writing demonstration should be adjustable; and 4) English explanations should be provided.

3.3 Analysis of the Interviews

After the participants finish the QUIS, they are immediately asked three questions. The first is the general feeling about the system's usability. Most gave a positive response. They felt the system was able to tutor them to write Chinese characters. Moreover, the scoring system seemed to motivate them to practice in order to achieve better results.

The second is which kind of practice ways were helpful to learn writing Chinese characters, and all participants agreed that the tracing stroke by stroke was the best for learning to write. One of the participants hoped that a future system could combine tracing stroke by stroke with whole character tracing. The participant felt this would be better.

The third are suggestions for future systems, and most of the participants expected different language interfaces and explanations, as well as more information that could help them remember how to write the Chinese characters.

3.4 The Description of Usability

Through the video record analysis of the system's operation, the researcher discovered and explored several usability problems. These problems as well as suggestions for their improvement are described below.

Pronunciation problems:

1. The left and right arrow buttons for polyphones caused confusion with the steps for writing strokes. Fig. 2 (labels 1 vs. 2) (s1, s2, and s4) shows that the buttons violate both the visibility and response principles. The buttons and responses should be located next to each other with precise labels explaining their functions.
2. The left and right arrow buttons are always present, and it does not matter whether or not the vocabulary word is a polyphone or not (Fig. 2, label 1). Even when the user pushes both buttons to repeat the action, it only responds verbally when the character is a real polyphone (s1, s2, s3, s4, and s6). This violates the response principle. We suggest canceling the polyphone or presenting the polyphone button when it is a real polyphone.

There are problems in the writing demonstration too. Participants confused the button for the writing demonstration function with the one for continuous steps for the strokes (Fig. 2, label 2 vs. 3) (s1, s2, and s4). This violates both the visibility and mapping principles and the labels for the function should be more precise, or possibly an icon could be provided for the button which would be more intuitive.

There are several problems in the writing practice:

1. The hierarchy relation in the tracing function (tracing with each stroke and tracing with a whole character) (Fig. 2, label 6) has a problem, in other words, participants must turn on the tracing function first and then select tracing with each stroke or tracing with whole character presentation, but they cannot select the needed function directly. This can cause some confusion (s1, s2, s3, and s6) which violates

the visibility principle. We suggest canceling the tracing function hierarchy setup and letting the users select which function they need directly.

2. There is confusion when there is no response to the request to change the tracing or grid options (s1, s2, s3, and s6). This is caused by the masking that occurs when users are writing characters in a certain area (see Fig. 2, labels 4 vs. 5 and 4 vs. 6). It is a problem with the response design. We suggest having a separate function setup area which apart from the main area. It would give immediate responses when the users prompt the function setup.
3. The presenting of unclear evaluation results (see Fig. 2, label 7) (s1, and s5) violates the visibility principle. One suggestion is to provide a detailed explore button, or present the evaluation results in a more visible way.
4. During the tracing practice, sometimes an incorrect operation will occur because of a touch or broken strokes. The system cannot accurately judge of the stroke is complete or not, and thus, it shows the next tracing stroke resulting in user confusion. This is the program problem due to the visibility and feedback.

Further problems include (a) function button sometimes misses being touched or misses the correct location (s2, s5, and s6), which violates the visibility principle. One suggestion is to resize the button to a more appropriate size, and arrange proper gaps between buttons. (b) after miss touching the button during writing practice (Fig. 2, label 8) the user cannot go back to the writing lesson where the entry writing practice began (s1, and s4). This violates the visibility principle. (c) Participants try to push the button for the writing character exercises and writing character tests (Fig. 2, label 8) (s1, s2, s4, and s5) when they find the function, and this violates the response principle. For both of the above problems, we suggest locating buttons nearby each other when the functions are close to decrease users accidentally touching the wrong one.

Fig. 2. The usability problems of writing character learning

4 Conclusion

This study used the Xi-Zi-e-Bi-Tong for its research case, which was to detect the foreign learner typical learning barriers when they are using the Chinese characters writing system. Although participants could interact with system, still some barriers existed. After analyzing the results, we hope to propose some suggestions for improving or redesigning the system.

Acknowledgements. We would like to thank the National Science Council of the Republic of China (Taiwan) for financial support of this research under contract numbers NSC 101-2410-H-224 -019 -MY2.

References

1. Li, C.Y.: The Cognitive and Neural Basis for Learning to Read Chinese. Journal of Basic Education 18(2), 63–85 (2009) (in Chinese)
2. Mack, R.L., Burdett, J.M.: When novices elicit knowledge: Question-asking in designing, evaluating and learning to use software. In: The Psychology of Expertise: Cognitive Research and Empirical AI, pp. 245–268. Springer, New York (1992)
3. Mangen, A., Velay, J.L.: Digitizing Literacy: Reflections on the Haptics of Writing. In: Zadeh, M.H. (ed.) Advances in Haptics. InTech (2010) ISBN: 978-953-307-093-3
4. Nielsen, J.: Usability Engineering, United Kingdom edition publish by Academic Press Ltd. (1993)
5. Norman, D.A.: The psychology of everyday things. Basic Books, New York (1988)
6. Preece, J., Rogers, Y., Sharp, H.: Interaction Design: Beyond Human-Computer Interaction. Wiley, New York (2002)
7. Tan, L.H., Spinks, J.A., Eden, G.F., Perfetti, C.A., Siok, W.T.: Reading depends on writing. PNAS 102(24), 8781–8785 (2005) (in Chinese)
8. The Questionnaire for User Interaction Satisfaction, `http://lap.umd.edu/quis/`