Measuring the Usability of Safety Signs: A Use of System Usability Scale (SUS)

Annie W. Y. Ng, Honour W. C. Lo, and Alan H. S. Chan

***Abstract*—System usability scale (SUS) was initially developed as a tool for measuring the subjective usability of industrial systems evaluation. The SUS has now been used for various evaluations such as voicemail system, virtual keyboard, paper ballots, mobile phone, and website. Usability is a critical parameter for considering the effectiveness of safety signs. However, the application of SUS for usability assessment of safety signs or other graphical symbols had never been seen. This study was undertaken to investigate how the SUS could be used for determining the usability of safety signs. Three hundred and ninety-eight participants (17 – 58 years old) were asked to complete a self-administered questionnaire on safety sign comprehension first and then SUS which was modified in accordance to the unique contexts and functions of safety signs. The results showed that the SUS offers a valid and robust measurement for practitioners to measure the usability of safety signs. Besides, the SUS provides valuable benchmark information about the comprehension performance of safety signs. Moreover, the SUS is a point estimate measure of usability which can be used to positively supplement a usability testing and evaluation for safety signs quickly and easily. Overall, this study would serve as a valuable reference on assessing the usability of safety signs and other graphical symbols with the SUS instrument.**

***Index Terms*—safety signs, usability, system usability scale (SUS), comprehension**

I. Introduction

THERE are many possible safety precaution measures that can be taken to attempt to reduce accidents and injuries in workplaces. Provision of safety signs is one of the significant measures in risk reduction. Safety signs are intended to identify and warn against specific hazards without the use of words [1]. They may represent a hazard, a hazardous situation, and a result of not avoiding a hazard. They also may describe safety precautions, advise users of the evasive actions to take, or provide other directions to eliminate or reduce hazards. Safety signs should provide good communication to users [2] as a failure to convey warning information effectively can lead to injury or death [3]. A recent study by Caldwell [4] indicated that 'if safety signs are not readily identifiable then their communicative value and hence usability are suspect'.

According to the International Organization for Standardization, usability is defined as 'extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use' [5]. Based on this definition, Jordan [6] identified five distinct usability components, namely; guessability, learnability, experienced user performance, system potential, and re-usability that focus on user performance with a product and apply in relation to a particular task in a specific context of use. Guessability is a measure of the cost (e.g. time on task, errors made) to the user of a product when performing a task for the first time. The learnability component is concerned with the cost to the user in achieving some competent level of performance on a task with a product, while the experienced user performance is the relatively stable performance level that an experienced product user reaches. The system potential component represents the theoretical optimal performance obtainable with a product with respect to a task, and the re-usability refers to the performance level achieved when a user returns to a task with a product after an extended period of non-use. These five components are utilized as the major objective measures of usability for safety signs. The guessability and learnability of some safety signs had been reported [7] – [9] and the results showed that many signs were not successfully guessed and were, therefore, unable to convey the intended safety messages.

To administer an easily implemented overall assessment of usability, Brooke [10] has developed a system usability scale (SUS) for industrial systems evaluation. The SUS instrument is composed of ten statements that are scored on a five-point Likert scale of strength of agreement with 1 indicating strongly disagree and 5 representing strongly agree. The statements cover a variety of aspects of system usability such as the need for support, training, and complexity, and thus have a high level of face validity for measuring usability of a system. The SUS instrument is generally used after the respondent has had an opportunity to use the system being evaluated, but before any debriefing or discussion takes place. The score of each statement contribution ranges from 0 to 4. The statements presented with positive and negative items are arranged alternatively. For statements 1, 3, 5, 7, and 9 the score contribution is the scale position minus 1, whereas for statements 2, 4, 6, 8, and 10 the contribution is 5 minus the scale position. Multiplying the sum of the scores from all statements by 2.5 leads to the overall value of system

Manuscript received November 17, 2010. The work described in this paper was supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 110508).

Annie W. Y. Ng is with the School of Design, The Hong Kong Polytechnic University.

Honour W. C. Lo is with the HKCC/CPCE, Hong Kong Polytechnic University.

Alan H. S. Chan is with the Department of Manufacturing Engineering and Engineering Management, City University of Hong Kong (corresponding author. phone: +852 27888439; fax: +852 2788 8423; e-mail: alan.chan@cityu.edu.hk).

ISBN: 978-988-19251-2-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

usability, viz. the SUS score, which ranges from 0 to 100. The higher the SUS score, the higher is the perceived usability.

The SUS instrument has been used for a variety of evaluations, for example on the calculator [11], virtual keyboard [12], paper ballots [13], e-government services [14], mobile phone [15], website [16], CD player [17], voicemail system [18], training game [19], in-vehicle user-interfaces [20], and multimodal remote control [21]. Regarding the context of the item to be evaluated, the SUS statements were slightly modified. For assessing the usability of products like phones and modems, Bangor et al. [22] indicated that approximately 90% of the SUS data presented in their study was collected using a slightly modified form of the original SUS. The modifications include replacing the word 'cumbersome' with a more recognizable synonym 'awkward' in statement 8 and changing the word 'system' to the word 'product' throughout the survey. For measuring the perceived usability of financial information websites [23], the SUS was adapted by replacing the word 'system' in every question with 'website'. The original phrase 'the support of a technical person' was revised as 'Tech Support', while another phrase 'to learn a lot of things before I could get going with this system' was amended as 'to learn a lot about this website before I could effectively use it'.

Usability is a critical factor for considering the effectiveness of safety signs. However, the application of SUS for overall assessment of usability on safety signs or other graphical symbols had never been seen. Here, the purpose of this study was to investigate whether the SUS can be useful for determining the usability of safety signs. The relationship between SUS scores and objective measures on safety sign comprehension performance was also examined. Demographic factors such as gender, education background, and work nature of participants were also considered in the analysis. The findings of this study would provide a valid and robust SUS instrument for designers and safety practitioners to measure the usability of safety signs quickly and easily. The study would also serve as a reference on evaluating other graphical symbols with the SUS instrument.

II. METHOD

A. Participants

Three hundred and ninety-eight Hong Kong Chinese students (332 males and 66 females), aged between 17 and 58 years old, participated in the study. 15% of the participants were from the higher certificate programme, 33% were from the diploma programme, 32% from the higher diploma programme, and the remaining 20% from the undergraduate programme. More than half of the participants (57%) did not have any work experience.

B. Survey Instrument

Safety Signs

Forty safety signs were randomly selected for testing from the British Standards [24] – [26]. Table I shows the safety signs used in this experiment. Eleven of them were mandatory signs that specify a specific course of action is to be taken (S2, S4, S9, S15, S18, S21, S32, S35, S37, S39, and S40), 7 were prohibition signs that indicate certain behavior is prohibited (S1, S6, S7, S11, S14, S29, and S36), 17 were warning signs that give warning of a hazard (S3, S8, S10, S12, S13, S16, S17, S20, S23, S24, S25, S26, S28, S31, S33, S34, and S38), and 5 were safe condition signs that provide information about safe conditions (S5, S19, S22, S27, and S30).

TABLE I
SAFETY SIGNS USED IN THE EXPERIMENT

	Sign	Meaning		Sign	Meaning
S1		No smoking	S21		Use adjustable guard
S2		Wash hands	S22		Emergency stop push-button
S3		Caution, risk of fire	S23		Caution, corrosive substance
S4		Foot protection must be worn	S24		Caution, overhead hazard
S5		Emergency eye wash	S25		Caution, risk of ionizing radiation
S6		Do not operate	S26		Caution, risk of electric shock
S7		Not drinking water (that is, do not drink)	S27		Emergency shower
S8		Caution, non-ionising radiation	S28		Caution, overhead load
S9		Wear face shield	S29		Smoking and naked flames prohibited
S10		Caution, industrial trucks	S30		First aid
S11		Do not use ladder	S31		Caution, slippery surface
S12		Caution, strong magnetic field	S32		Use guard
S13		Caution, trip hazard	S33		Caution, laser beam
S14		Pedestrians prohibited	S34		Caution, biological hazard
S15		Eye protection must be worn	S35		Respiratory protection must be worn
S16		Caution, risk of explosion	S36		Do not extinguish with water
S17		Caution, toxic hazard	S37		Keep locked
S18		Head protection must be worn	S38		Caution, risk of fire, oxidizing materials
S19		Emergency telephone	S39		Hearing protection must be worn
S20		Caution, limited overhead height	S40		Hand protection must be worn

Questionnaire

A self-administered questionnaire was designed to gather information about participants' safety sign comprehension performance, usability scores on safety signs, and demographic information. For effective communication, the

ISBN: 978-988-19251-2-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Chinese version of the questionnaire was used for testing. The *first* part of the questionnaire was about safety sign comprehension. The safety signs were displayed in color and in squares of 2 cm x 2 cm without boundary. Respondents were asked to complete a set of multiple-choice questions for the signs for evaluating their understanding of safety signs. Four-option multiple-choice questions were designed as they could greatly diminish the risk of guessing by participants [27]. The selection choices consist of one correct answer and three plausible distractors. Participants were awarded 1 mark for a correct answer and 0 mark for an incorrect answer.

The *second* part of the questionnaire was about the subjective usability of safety signs. In this study, the subjective usability of safety signs was collected using a slightly modified form of the original SUS. The original SUS statements from Brooke [10] and the modified statements used in this experiment are shown in Table II. The word 'system' was replaced with the word 'safety signs' throughout the survey. The phrase 'to use' in safety sign system can better be interpreted as 'pay attention to' and 'to understand'. Statement 1 was thus revised as "I think that I would like to pay attention to safety signs at workplace and other public places". Statements 3, 7, and 8 were modified as "I thought that the safety signs conveyed clear meaning", "I think that most people can understand the safety signs after having suitable training", and "I found the safety signs very difficult to understand', respectively. Regarding statement 2, the original term 'unnecessarily complex' in safety sign system was somewhat ambiguous and revised as "I found that the safety signs conveyed simple meanings but could not be interpreted correctly without text explanation". A system consists of various functions which would be well integrated or could lead to inconsistency. However, individual safety sign belongs to a particular category (e.g. mandatory, prohibition, warning, and safe condition) and transmits a single meaning only and thus the original statement 5 was amended as "I found that the safety signs were well classified". Statement 6 was also modified as "I thought that there was too much confusion in interpreting the safety signs". Within a safety sign system, the provision of supplemental explanation would help prospective users to be able to comprehend the safety signs effectively. Thus, statement 4 was revised as "Without any supplementary text, I need somebody to explain the meaning of safety signs and statement 10 was updated as "I needed to have detail explanation before I could get going with the safety signs".

The *third* part of the questionnaire asked about participants' demographic information. Participants were asked to complete several closed-ended questions on age group, gender, education, and work experience.

C. *Pilot Testing and Assessing Validity*

To assess how well a questionnaire measures what it is intended to measure, the validity of a questionnaire is usually assessed by individuals with expertise in some aspect of the subject under study [28]. Before distribution of the questionnaire to the participants, its face and content validity were assessed by eight academics and professionals who were knowledgeable with usability measurement and to comment and suggest changes for the questionnaire.

D. *Survey Administration*

Participants were briefed on the objectives of the survey and given verbal instructions at the beginning of the test. The participants were asked to complete the safety sign comprehension, then the system usability scale, and finally the demographic information. The survey was undertaken during normal lecture hour to ensure participation rate and to allow participants to have enough attention to complete the questionnaire. Four hundred questionnaires were distributed and 398 questionnaires were completed and returned.

TABLE II
THE ORIGINAL SUS STATEMENTS (BROOKE [10]) AND THE MODIFIED STATEMENTS USED IN THIS EXPERIMENT

	Original SUS statements	Modified SUS statements	Mean	SD	Absolute
1	I think that I would like to use this system frequently	I thought that I would like to pay attention to safety signs at workplace and other public places	2.68	1.148	2.68
2	I found the system unnecessarily complex	I felt that the safety signs conveyed simple meanings but could not be interpreted correctly without text explanation.	2.43	1.035	3.57
3	I thought the system was easy to use	I thought that the safety signs conveyed clear meaning.	1.68	0.900	1.68
4	I think that I would need the support of a technical person to be able to use this system	Without any supplementary text, I need somebody to explain the meaning of safety signs.	2.37	1.044	3.63
5	I found that the various functions in this system were well integrated	I found that the various types of safety signs were well classified	1.93	0.906	1.93
6	I thought that there was too much inconsistency in this system	I thought that there was too much confusion in interpreting the safety signs.	2.47	0.998	3.53
7	I would imaging that most people would learn to use this system very quickly	I thought that most people can understand the safety signs after having suitable training.	2.90	1.056	2.90
8	I found the system very cumbersome to use	I found the safety signs very difficult to understand	1.68	1.109	4.32
9	I felt very confident using the system	I felt very confident using the safety signs	2.48	1.047	2.48
10	I needed to learn a lot of things before I could get going with this system	I needed to have detail explanation before I could get going with the safety signs.	2.02	1.063	3.98

III. RESULTS

A. *Comprehension Score*

In this study, comprehension score refers to the accuracy level for understanding the meaning of a safety sign. The overall mean and standard deviation of comprehension score for all 40 safety signs were 67.54% and 23.47%, respectively. The sign with minimum comprehension score (2.26%) was 'caution, non-ionising radiation' (S8). The signs with maximum score (95.73%) were 'pedestrians prohibited'

ISBN: 978-988-19251-2-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

(S14) and 'hand protection must be worn' (S40).

The International Organization for Standardization (ISO) and the American National Institute (ANSI) recommend that symbols must reach a criterion of at least 67% or 85% correct, respectively, in a comprehension test to be considered acceptable. In this study, there were 12 safety signs reaching both the ISO and ANSI criteria, viz. 'no smoking' (S1, 92.96%), 'emergency eye wash' (S5, 86.18%), 'no drinking water' (S7, 87.19%), 'wear face shield' (S9, 91.46%), 'pedestrians prohibited' (S14, 95.73%), 'eye protection must be worn' (S15, 87.94%), 'caution, risk of explosion' (S16, 88.94%), 'caution, limited overhead height' (S20, 89.45%), 'do not extinguish with water' (S36, 91.96%), 'no naked flame' (S38, 89.95%), 'hearing protection must be worn' (S39, 94.72%), and 'hand protection must be worn' (S40, 95.73%). Another 10 safety signs achieved the lower criteria of ISO only, namely 'wash hands' (S2, 79.40%), 'caution, industrial truck' (S10, 67.59%), 'caution, trip hazard' (S13, 81.91%), 'emergency telephone' (S19, 67.84%), 'emergency stop push-button' (S22, 80.90%), 'caution, corrosive substance' (S23, 76.88%), 'caution, risk of ionizing radiation' (S25, 82.91%), 'emergency shower' (S27, 77.64%), and 'first aid' (S30, 80.90%), 'keep locked' (S37, 80.90%).

Comprehension scores for the remaining 18 safety signs did not meet either the ISO or ANSI criteria. More than half of the participants could identify the correct meaning for 'caution, risk of fire' (S3), 'foot protection must be worn' (S4), 'do not operate' (S6), 'caution, toxic hazard' (S17), 'head protection must be worn' (S18), 'use adjustable guard' (S21), 'caution, overhead load' (S28), 'caution, slippery surface' (S31), 'use guard' (S32), and 'caution, biological hazard' (S34). Whereas, less than half of the participants could recognize the correct meaning for these eight safety signs viz. 'caution, non-ionising radiation' (S8), 'do not use ladder' (S11), 'caution, strong magnetic field' (S12), 'caution, overhead hazard' (S24), 'caution, risk of electric shock' (S26), 'smoke and naked flames prohibited' (S29), 'caution, laser beam' (S33), and 'respiratory protection must be worn' (S35).

B. *Comprehension Performance*

The comprehension performance here denotes the performance level of a participant in the sign comprehension task. The overall mean comprehension performance for all participants was 67.54%, with standard deviation of 10.98%. The minimum comprehension performance was 7.5% and the maximum comprehension performance was 92.50%. To determine if there were any participants with comprehension performance very different from others, the box plot of comprehension performance for all participants was prepared (Fig. 1). The bottom and top lines of the box correspond to the first quartile and third quartile, respectively, and the horizontal line within represents the median. The vertical lines, whiskers, are drawn from the edges of the box to the largest and smallest values that are outside the box but within 1.5 box lengths. Outliers are values more than 1.5 box lengths away from the box and would have been flagged with small circles in the box plot. In this study, 13 participants were assessed as outliers below the box, indicating that their comprehension performance were much lower than other participants. The comprehension performance of these 13 participants lied between 7.5% and 42.5%.

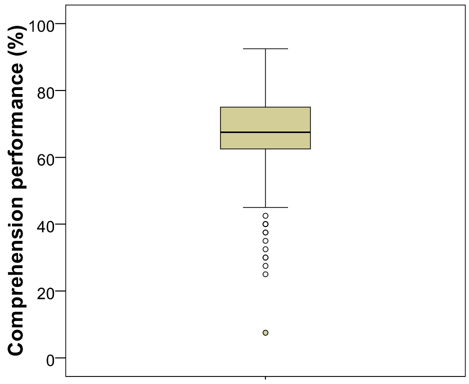

Fig. 1. Box plot of comprehension performance for all participants. 13 outliers are flagged with small circles.

C. *System Usability Scale (SUS) Score*

Responses to individual SUS statements are depicted in Table II. For each statement, the mean and standard deviation for the raw scores was presented. An absolute score that adjusts the scores of Statement 2, 4, 6, 8, and 10 was calculated, so that positive responses are associated with a larger number, like the other five statements. Statement 8 ("I found the safety signs very difficult to understand") received the most extreme of all responses, with a mean absolute score of 4.32. Statement 3("I thought that the safety signs conveyed clear meaning") had the lowest standard deviation of its raw scores, at 0.9, well below the others, which ranged from 0.906 to 1.148. The most varied responses were to Statement 1 ("I think that I would like to pay attention to safety signs at workplace and other public places"), with a standard deviation of 1.148.

Multiplying the sum of the raw scores from all statements by 2.5 led to the overall value of sign usability (called SUS score). The SUS scores on the test signs of 152 participants were below 50, 223 participants were between 50 and 70, and 23 participants were above 70. On the whole, the mean SUS score for all participants was 51.75, with standard deviation of 12.82. The minimum SUS score was 10 and the maximum score was 87.50. In accordance with the rule-of-thumb on the interpretation of SUS scores on products (Bangor et al., 2008), products with scores less than 50 should be cause for significant concern and are judged to be unacceptable, products with scores between 50 and 70 are marginal acceptable, and products with SUS scores above 70 are passable. In this study, the mean SUS score of 51.75 indicated that the safety signs were generally perceived to be marginal at best.

D. *Relationship between SUS Score and Comprehension Performance*

Table III shows the demographic factors, the number of participants' responses to each of the factors, and the comprehension performance and SUS score of the participants in each of the response categories. The major findings were: most participants (83%) were males and studied technical or engineering programme. For the 170 participants that had work experience, 76% reported that their work was technical and engineering related while 23%

ISBN: 978-988-19251-2-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

were in non-technical and engineering area. Mann-Whitney test showed that the demographic factors of gender, education background, and work nature did not have significant influence on SUS score (p's > 0.05).

To investigate the relationship between SUS score and comprehension performance, Spearman correlation was conducted and revealed that SUS score was significantly associated with comprehension performance ($r_s = 0.147$, n = 398, $p = 0.003$). The relationship between SUS score and comprehension performance was found significant in particular if participants were male ($r_s = 0.155$, n = 332, $p < 0.05$) and had technical /engineering education background ($r_s = 0.152$, n = 329, $p < 0.05$).

TABLE III
A SUMMARY OF RESPONSES FOR THE 398 PARTICIPANTS AND MEAN COMPREHENSION PERFORMANCE AND SUS SCORE FOR DIFFERENT GROUPS OF PARTICIPANTS

Demographic factor	Response	No.	Comprehension performance (%)		SUS score	
			Mean	*SD*	*Mean*	*SD*
Gender	Male	332	67.73	10.54	51.48	13.27
	Female	66	66.63	13.03	53.11	10.23
Education background	Technical/engineering	329	68.32	10.29	51.98	13.31
	Non-technical/engineering	53	63.73	14.25	50.57	9.79
	Did not report	16	-	-	-	-
Work nature	Technical/engineering	130	68.21	9.95	53.37	12.75
	Non-technical/engineering	39	63.59	14.55	51.22	14.69
	Did not report	1	-	-	-	-
Gender	Male	332	67.73	10.54	51.48	13.27

IV. DISCUSSIONS

The system usability scale (SUS) was developed by Brooke [10] as a tool for measuring the subjective usability of industrial systems evaluation. Since then, the SUS has been used for a variety of evaluations such as voicemail system, virtual keyboard, paper ballots, mobile phone, and website [12, 13, 15, 16, 18]. Sometimes, regarding to the context of the item to be evaluated, the SUS was slightly adjusted [23]. Usability is a critical parameter for considering the effectiveness of safety signs. This study was conducted to investigate how the SUS can be used for determining the usability of safety signs. The original SUS instrument was modified in accordance to the unique contexts and functions of safety signs, and then its face and content validity were assessed with several academics who were familiar with usability measurement.

This study provided a better understanding of the sensitivity of the SUS to quantitative measures of performance for safety signs. Participants' SUS score was significantly correlated to their individual comprehension performance. Participants who have high SUS scores tended to perform better on objective measures and vice versa. It can be said with confidence that the SUS instrument was providing valuable benchmark information about the comprehension performance of safety signs. Besides, the SUS scores would not be affected by the respondents' characteristics of gender, education background, and work nature.

Hence, based on the above findings, the SUS can be used to positively supplement a usability testing and evaluation for safety signs. This can be simply done so by asking prospective users to complete the SUS with respect to the corresponding safety signs. Practitioners then estimate the SUS score by multiplying the sum of raw scores from all SUS statements by 2.5. In accordance with the Bangor et al. [22]'s rule-of-thumb on the interpretation of SUS score, the SUS score could give practitioners with sufficient evidence that safety signs are or are not sufficiently usable in its current form.

The SUS provides a single score that estimates the overall usability of safety signs. It can be said to be 'a point estimate measure of usability'. However, the generalizability of SUS for other graphical symbols is not fully known. Future research might consider to repeat the same study with other graphical symbols like road signs, medical signs, computer icons, and system icons. Such findings would help to generate a SUS instrument for measuring the usability of various kinds of graphical symbols. On the whole, the results of this study indicated that SUS will help further the ability of practitioners to measure the usability of safety signs.

V. CONCLUSIONS

This study successfully investigated how the SUS can be used for determining the usability of safety signs. Regarding to the contexts and functions of safety signs, the original SUS instrument was modified accordingly. It was found that the modified SUS provided a valid and robust measurement for practitioners to measure the usability of safety signs. The modified SUS was revealed to be statistically significant to quantitative measures of performance for safety signs. Prospective users who have high SUS scores tended to perform better on comprehension performance. Also, the SUS provides a single score that estimates the overall usability of safety signs which can be used to positively supplement a usability testing and evaluation for safety signs quickly and easily. In general, the study would be a reference for measuring the usability of safety signs and other graphical symbols with SUS instrument.

ACKNOWLEDGMENT

The authors would like to thank Tony Lam for his help in the data collection process of the experiment.

REFERENCES

[1] R. Ebens, "Designing effective machine safety signs," *Occupational Hazards*, pp. 29-30, July 2007.
[2] S. Arphorn, N. Augsornpeug, S. Srisorrachatr, and V. Pruktharathikul, "Comprehension of safety signs for construction workers: comparison of existing and newly designed sign," *Journal of Human Ergology*, vol. 32, pp. 87-94, 2003.
[3] M. Lesch, "Comprehension and memory for warning symbols: age related differences and impact of training," *Journal of Safety Research*, vol. 34, pp. 495-505, 2003.
[4] J. Caldwell, "Safety icons and usability: a Peircean reanalysis," in *Proceedings of the IEEE International Professional Communication Conference*, 2009, pp. 1-8.
[5] *Ergonomics Requirements for Office Work with Visual Display Terminals (VDTs) Part 11: Guidance on Usability*, ISO 9241-11, 1998.

[6] P. W. Jordan, *An Introduction to Usability*. London: Taylor & Francis, 1998.

[7] A. H. S. Chan, S. H. Han, A. W. Y., Ng, and W. Park, “Hong Kong Chinese and Korean comprehension of American security safety symbols,” *International Journal of Industrial Ergonomics*, vol. 39, no. 5, pp. 835-850, 2009.

[8] A. H. S. Chan and A. W. Y. Ng, “Investigation of guessability of industrial safety signs: effects of prospective-user factors and cognitive sign features,” *International Journal of Industrial Ergonomics*, vol. 40, no. 6, pp. 689-697, 2010.

[9] A. H. S. Chan and A. W. Y. Ng, “Effects of sign characteristics and training methods on safety sign training effectiveness,” *Ergonomics*, vol. 53, no. 11, pp. 1325-1346, 2010.

[10] J. Brooke, “SUS: A “quick and dirty” usability scale,” in *Usability Evaluation in Industry*, P. W. Jordan, B. Thomas, B. A. Weerdmesster, and I. L. McClelland, Eds. London: Taylor & Francis, 1996, pp. 189-194.

[11] P. Cairns, S. Wali, and H. Thimbleby, “Evaluating a novel calculator interface,” in *Proceedings of HCI 2004*, vol. 2, A. Dearden and L. Watts, Eds. Research Press, 2004, pp. 9-12.

[12] T. Költringer and T. Grechenig, “Comparing the immediate usability of Graffiti 2 and virtual keyboard,” in *Proceedings of CHI '04: CHI '04 extended abstracts on Human factors in computing systems*, 2004, pp. 1175-1178.

[13] S. P. Everett, M. D. Byrne, and K. K. Greene, “Measuring the usability of paper ballots: efficiency, effectiveness, and satisfaction,” in *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*, 2006, pp. 2547-2551.

[14] T. A. Horan, T. Abhichandani, and R. Rayalu, “Assessing user satisfaction of e-government services: development and testing of quality-in-use satisfaction with advanced traveler information systems (ATIS),” in *Proceedings of the 39th Hawaii International Conference on System Sciences*, 2006.

[15] T. Jokela, J. Koivumaa, J. Pirkola, P. Salminen, and N. Kantola, “Methods for quantitative usability requirements: a case study on the development of the user interface of a mobile phone,” *Personal Ubiquitous Computing*, vol. 10, pp. 345-355, 2006.

[16] M. Zviran, C. Glezer, and L. Avni, “User satisfaction from commercial web sites: the effect of design and use,” *Information & Management*, vol. 43, pp. 157-178, 2006.

[17] B. Hendriks and J. Hu, “Redesigning an CD player for intuitive rich interaction,” in *Proceedings of the 12th International Conference on Human-Computer Interaction*, 2007, pp. 1607-1611.

[18] S. Perugini, T. J. Anderson, and W. F. Moroney, “A study of out-of-turn interaction in menu-based, IVR, voicemail systems,” in *Proceedings of CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, 2007.

[19] T. S. Hussain, B. Roberts, C. Bowers, J. A. Cannon-Bowers, E. S. Menaker, S. L. Coleman, C. Murphy, K. Pounds, A. Koenig, R. Wainess, and J. Lee, “Designing and developing effective training games for the US Navy,” in *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference*, 2009.

[20] S. Nestler, M. Tönnis, and G. Klinker, “Common interaction schemes for in-vehicle user-interfaces,” in *Human-Computer Interaction. Ambient, Ubiquitous and Intelligent Interaction*, J. A. Jacko, Ed. Springer-Verlag Berlin Heidelberg, 2009, pp. 159-168.

[21] I. Wechsung and A. B. Naumann, “Evaluating a multimodal remote control: The interplay between user experience and usability,” *International Workshop on Quality of Multimedia Experience 2009*, 2009, pp. 19-22.

[22] A. Bangor, P. T. Kortum, and J. T. Miller, “An empirical evaluation of the system usability scale,” *International Journal of Human-Computer Interaction*, vol. 24, no. 6, pp. 574-594, 2008.

[23] T. S. Tullis and J. N. Stetson, “A comparison of questionnaires for assessing website usability,” in *Proceedings of the Usability Professional Association Conference*, 2004.

[24] *Safety Signs and Colours Part 1: Colour and Design*, BS5378-1, 1980.

[25] *Safety Signs and Colours Part 3: Specification for Additional Signs to those given in BS5378: Part 1*, BS5378-3, 1982.

[26] *Fire Safety Signs, Notices and Graphic Symbols – Specification for Fire Safety Signs*, BS5499-1, 1990

[27] T. Newby, *Training Evaluation Handbook*. Aldershot, Hants, England: Gower, 1992.

[28] M. S. Litwin, *The survey kit: how to measure survey reliability and validity*. Thousand Oaks, CA: Sage Publications, 1995.

ISBN: 978-988-19251-2-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)