An Indonesian Adaptation of the System Usability Scale (SUS)

Zahra Sharfina[1], Harry Budi Santoso[2]
[1,2] Faculty of Computer Science
Universitas Indonesia, Depok, Indonesia
zahrasharfina32@gmail.com[1], harrybs@cs.ui.ac.id[2]

***Abstract*—As a key aspect of product development, usability of a product is needed to be assessed by conducting usability evaluation. Within the usability evaluation methods, there are several usability assessment questionnaires; one of the most commonly used is the System Usability Scale (SUS). However, the SUS is originally developed in English and there had not been any study conducted to develop the Indonesian version of SUS. The adapted version is expected to help usability researchers and practitioners in Indonesia while conducting usability product evaluation. Therefore, this study was carried out to adapt the original SUS into the Indonesian version by using cross-cultural adaptation and reliability test. The Cronbach's Alpha of the Indonesian adaptation of SUS was 0.841, which concludes that this version is reliable to be used by usability practitioners.**

Keywords*—*usability; SUS; adaptation; reliability

I. INTRODUCTION

Usability is the 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 [1]. Because of the wide range of usability, therefore, usability needs to be assessed by usability evaluation to ensure a product meets the users' goals or not. Usability evaluation consists of iterative cycles of designing, prototyping, and evaluating [2][3]. Within the usability evaluation methods, questionnaires are commonly used and important for qualitative data collection about the characteristics, thoughts, feelings, perceptions, behaviors or attitude of users [4]. Despite being the low budget advantage, somehow questionnaires provide useful information and reflect users' opinions. One of the most commonly used questionnaires for assessing usability is the System Usability Scale (SUS) [5], known as a "quick and dirty" survey scale [6, p. 2] that allows practitioners to quickly and easily access the usability of a given product. It is an effective tool, yet inexpensive, for assessing the usability of a product, as well as a wide range of user interfaces [7].

Since the usability evaluation is a key aspect in the development of products, it is important to apply reliable questionnaires within the assessment. As for the SUS, it is originally built in English, therefore, it is necessary to adapt the SUS to other languages in order to give optimal usability measures. Recent studies have translated the SUS into Spanish, French, Dutch, Portuguese, Slovene, Persian, and Germany [7-11]. Despite the SUS is also widely used in Indonesia [12-14], there had not been any study conducted to adapt the SUS to an Indonesian version with a sufficient reliability. However, it is not enough to perform a literal translation of the original version of SUS. The items of SUS must not only be translated well linguistically, but also must be adapted culturally to maintain the content validity. Translation is merely the first stage of the adaptation process. Cultural, idiomatic, linguistic, and contextual aspects have to be considered when adapting an instrument [15]. Instrument adaptation enables the researcher to compare data from different samples and backgrounds. Therefore, it leads to greater fairness in the evaluation. By adapting an instrument, it assesses the construct based on the same theoretical and methodological perspectives [16]. An instrument adaptation also enables a greater ability to generalize and investigate differences within a diverse population [15].

Therefore, the objective of this study is to translate and adapt the original SUS into the reliable Indonesian version of SUS. This paper is organized in five parts: 1) introduction explains about the background and objective of this study; 2) relevant literature review tells the literature review result which is related to the study; 3) methods consist of relevant techniques that were used; 4) results show the result of this study; and 5) conclusions and discussions explain the main points of this study and also ideas for further research.

II. RELEVANT LITERATURE REVIEW

A. Usability

Usability is defined as user experience quality while users interact with the product or system [17]. Usability is also known as a quality attribute that assesses how easy user interfaces are to use [18]. The components of usability are divided by 5 quality components, i.e. 1) learnability, assesses how easy for users to do their tasks in their first time experience while using the system; 2) efficiency, assesses how quickly for users to do their tasks after learning how to use the system; 3) memorability, assesses how easy for users to build their capability to use the system after they do not use it for a short period of time; 4) errors, assesses how many mistakes that users do, how severe the mistakes, and how easy the users solve their problems; and 5) satisfaction, assesses how pleasant the system usage for users [18].

Not only having quality components, but usability also has divided in three principles, i.e. learnability, flexibility, and

 978-1-5090-4629-4/16/$31.00

robustness [2]. Learnability is related to how easy for a new user to learn the interface of the system. Meanwhile, flexibility is related to how many ways users can do to interact with the system. The last is robustness which is related to how well the support can solve users' problems while using the system.

B. Usability Evaluation

Usability evaluation focuses on how well users can learn and use a product to accomplish their goals [17]. Usability evaluation is also related to users' satisfaction to utilization process of those product. There are twelve specific steps to conduct the usability evaluation [2][3][19]: 1) specify the goal of usability evaluation; 2) determine user interface aspects that will be evaluated; 3) identify user targets; 4) choose the usability metric that will be used; 5) choose the evaluation method; 6) choose the tasks that will be conducted; 7) plan the experiment; 8) collect the usability evaluation data; 9) analyze and interpret the usability evaluation data; 10) criticize the user interface to give improvement recommendations; 11) repeat the process if needed; and 12) convey the results.

Results from usability evaluation can be used to predict the success of a product after it is released to the intended market. Usability evaluation also gives advantage to compare similar products and provide feedbacks about the usability context of the products [20]. There are various methods in usability evaluation, i.e. usability testing, first click testing, eye tracking, contextual interview, remote testing, mobile device testing, focus group, heuristic evaluation, and expert review [21]. Usability evaluation methods are classified as empirical and non-empirical methods. In the empirical methods, users are being observed while interacting with products or being asked to explain their perspectives of the usability of a product. On the other hand, the non-empirical methods involve the experts' judgment about the usability of a product [22][23].

C. System Usability Scale (SUS)

The SUS was developed in 1986 by John Brooke [24] and its value is to provide a single reference score for participants' view of a product's usability. The original SUS is shown below in Table I.

TABLE I. The Original SUS

No.	Original Item
1	I think that I would like to use this system.
2	I found the system unnecessarily complex.
3	I thought the system was easy to use.
4	I think that I would need the support of a technical person to be able to use this system.
5	I found the various functions in the system were well integrated.
6	I thought there was too much inconsistency in this system.
7	I would imagine that most people would learn to use this system very quickly.
8	I found the system very cumbersome to use.
9	I felt very confident using the system.
10	I needed to learn a lot of things before I could get going with this system.

This instrument also can be used to assess the usability of a wide range of products [25]. Recent study shows that it can be divided into two sub-scales of usability and learnability: usable (item 1, 2, 3, 5, 6, 7, 8, and 9) and learnable (item 4 and 10) [24]. The SUS itself consists of 10 items, the odd numbers are for positive items and the even numbers for the otherwise. The respondents of SUS are asked to rate the usability of a product on a 5-point scale numbered from 1 (strongly disagree) to 5 (strongly agree). For positive items, the score contribution is the scale position minus 1 and for the negative items, the score contribution is 5 minus the scale position. The overall SUS score is the result of the sum of item score contributions multiply by 2.5, range from 0 to 100 [6]. A product is considered having good usability if the overall SUS score is equal or above 68 [26].

D. Cross-Cultural Adaptation

The adaptation of instrument is based on international established guidelines of cross-cultural adaptation [27] in order to ensure the quality of translation result and the consistency of the meaning between this version and the original. Steps of the cross-cultural adaptation conducted in the current study are explained below:

1. Initial translation from original version into target language version was carried out and named as *forward translation*. Bilingual translators whose mother tongue is the target language produced the two independent translations. Their rationale for their choices was also summarized in the report.
2. The two translators and an observer met to synthesize the results of the translations. A synthesis of these translations was first conducted, each of the issues were addressed and resolved.
3. A translator then translated the questionnaire back into the original language. This stage is called as *back translation*.
4. Expert committee reviewed the translation results. The expert committee's role was to consolidate all the versions of the questionnaire and develop what would be considered the pre-final version of the questionnaire for field testing.
5. The final stage of adaptation process was the pretest. Each subject completed the questionnaire, and was interviewed to explain about what they thought was meant by each questionnaire item and the chosen response. Both the meaning of the items and responses were explored.

E. Reliability Test

An instrument is considered not valid unless it is reliable. Reliability is the ability of an instrument to give consistent measure [28]. Reliability test is a consistency test for survey or other measurements. Reliability coefficient is the statistic result which determines whether a test or survey is reliable or not. The coefficient represents the correlation between survey's variables [29]. Cronbach's Alpha is one of the reliability test techniques that only needs a single test to give unique estimation for the reliability of the given test. It was developed to provide a measure of the internal consistency of

 978-1-5090-4629-4/16/$31.00

a test or scale. Internal consistency describes the extent to ensure all items in a test measure the same concept or construct and it should be determined before a test is used for research to ensure the validity [30]. Range of Cronbach's Alpha reliability coefficient is from 0 to 1. The closer Cronbach's Alpha coefficient to 1, means the internal consistency of items being evaluated is wider. An instrument is considered reliable if the Cronbach's Alpha is equal or above 0.7 [31]. Thus, the value of Cronbach's Alpha is increased if items in a test are correlated to each other [30].

III. METHOD

The adaptation was based on cross-cultural adaptation method as mentioned in the relevant literature review. Firstly, forward translation process was conducted by translating the original SUS with two translators. One of the translators was be aware of the concepts being examined in the SUS being translated and the other translator was neither be aware nor informed of the concepts being quantified. In this study, the first translator was the author and the other was the certified professional translator. The translators each produce a report of the translation. The gaps between two translations were synthesized into one translation after being discussed by both translators.

Another translator then translated the forward translation result of SUS back into the original language. This process is to validate the translated version reflects the same item content as the original version. The translator is a native speaker in both language, English and Indonesian, who was not be aware nor be informed of the concepts explored. The back translation result then was evaluated by the expert to achieve the cross-cultural equivalence. The expert reviewed the translation and reached a consensus on any discrepancy between the original and translated version. The pre-final version of Indonesian SUS was pretested to target users by face validity.

Face validity is a subjective assessment to measure an instrument is already appropriate with the respondents' understanding [32]. Thus, face validity was conducted to 10 respondents from technical and non-technical backgrounds who were asked to give scale of their understanding towards the items of translated SUS. The scale has 5 point, numbered from 1 which indicates for "strongly not understand" and 5 indicates for "strongly understand". The respondents were also asked to give their rationales for their choices. Face validity result was analyzed by calculating the scale average and mapping the respondents' rationales to the final version of Indonesian SUS with a guidance from the expert.

Furthermore, to make sure the Indonesian version of SUS is reliable and can be used across different population, a reliability test was conducted by online questionnaire. The respondents of the test were 108 students in the Faculty of Computer Science, Universitas Indonesia, from different batch. The context of the reliability test was to give score for usability of Student-Centered Learning of this faculty by using the Indonesian version of SUS. The result of questionnaire was analyzed by Statistical Package for Social Sciences (SPSS) to get the Cronbach's Alpha score.

IV. RESULTS

The cross-cultural adaptation resulted in 10 Indonesian translated items of SUS that were considered equivalent to the corresponding items of the original SUS (See Table II).

TABLE II. The Indonesian Version of SUS

No.	Item in Indonesian
1	Saya berpikir akan menggunakan sistem ini lagi.
2	Saya merasa sistem ini rumit untuk digunakan.
3	Saya merasa sistem ini mudah untuk digunakan.
4	Saya membutuhkan bantuan dari orang lain atau teknisi dalam menggunakan sistem ini.
5	Saya merasa fitur-fitur sistem ini berjalan dengan semestinya.
6	Saya merasa ada banyak hal yang tidak konsisten (tidak serasi) pada sistem ini.
7	Saya merasa orang lain akan memahami cara menggunakan sistem ini dengan cepat.
8	Saya merasa sistem ini membingungkan.
9	Saya merasa tidak ada hambatan dalam menggunakan sistem ini.
10	Saya perlu membiasakan diri terlebih dahulu sebelum menggunakan sistem ini.

This version of SUS had also been validated with 10 respondents as mentioned before in the previous section. This instrument has the semantic equivalence prior to the same meaning with the original SUS and there is no grammatical difficulty in the translated version. In addition, Cronbach's Alpha result by involving 108 students showed that the Indonesian version of SUS was considered reliable prior to the score which was above 0.7 (See Table III). The result also showed that this instrument is still reliable if one the item is deleted because the Cronbach's Alpha score is not changed significantly and it remains consistent.

TABLE III. The Cronbach's Alpha Result of Indonesian Version of SUS

Reliability Statistics

Cronbach's Alpha	N of Items
,841	10

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Item01	33,70	29,481	,470	,832
Item02	34,21	26,020	,773	,803
Item03	34,04	26,111	,780	,803
Item04	34,01	30,327	,277	,850
Item05	34,05	30,194	,414	,837
Item06	34,52	29,261	,425	,836
Item07	34,62	27,154	,541	,826
Item08	34,13	26,637	,705	,810
Item09	34,32	27,922	,567	,824
Item10	35,15	27,174	,471	,836

 978-1-5090-4629-4/16/$31.00

V. CONCLUSIONS AND DISCUSSION

This study aimed to translate and adapt the original SUS into the Indonesian version. The original SUS was processed by cross-cultural adaptation and face validity to ensure that the Indonesian version of SUS is valid and can be used in different population and culture. After checking the validity, the translated items were being tested to determine if it is reliable or not. The reliability test data was analyzed to get the Cronbach's Alpha score by SPSS software. Result showed that the Indonesian version of SUS is reliable prior to the Cronbach's Alpha score was 0.841 and the score remains consistent if one of the items is deleted. Therefore, this instrument can be used by usability practitioners across different cultures for usability evaluations and research purposes.

However, this Indonesian version of SUS still needs further research in other systems, e.g. e-commerce, government system, and news portal websites, in order to optimize this instrument. When evaluating a system, instrument can be compared with the original SUS to study about its differences and to validate the contribution of the adapted version. It is also possible in the future to conduct study to further analyze the psychometric properties in the Indonesian version of SUS. It is also necessary to enhance the quality of linguistic aspects of the adapted version of the SUS.

REFERENCES

[1] ISO 9241-11. (1998). *Guidance on Usability* [Online]. Available: http://www.usabilitynet.org/tools/r_international.htm

[2] A. Dix, J. Finlay, G. Abowd, and R. Beale, *Human-computer interaction*, 3rd ed. Upper Saddle River, NJ: Prentice Hall, 2003.

[3] J. Nielsen, *Usability Engineering*. Boston, MA: Academic Press, 1993.

[4] B. Hanington and B. Martin, *Universal Methods of Design: 100 Ways to Research Complex Problems, Develop Innovative Ideas, and Design Effective Solutions*. Beverly, MA: Rockport Publishers, 2012.

[5] J. R. Lewis, *Usability Testing*. IBM Software Group, 2006.

[6] J. R. Lewis and J. Sauro, *The Factor Structure of the System Usability Scale*. IBM Software Group, 2009.

[7] A. I. Martins, A. F. Rosa, A. Queiros, A. Silva, and N. P. Rocha, "European Portuguese Validation of the System Usability Scale (SUS)", in *6th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Infoexclusion*, 2015.

[8] J. Sauro, *A Practical Guide to the System Usability Scale: Background, Benchmarks, & Best Practices*. Measuring Usability LLC, 2011.

[9] B. Blazica and J. R. Lewis. A Slovene translation of the System Usability Scale: the SUS-SI. International Journal of Human-Computer Interaction, 2015, 31(2), pp. 112-117.

[10] I. Dianar, Z. Ghanbari, M. AsghariJafarabadi. Psychometric properties of the Persian language version of the System Usability Scale. Health Promotion Perspectives, 2014, 4(1), pp. 82-89.

[11] K. Lohmann and J. Schäffer. (2013, September 18). *System Usability Scale—An Improved German Translation of the Questionnaire* [Online]. Available: https://minds.coremedia.com/2013/09/18/sus-scale-an-improved-german-translation-questionnaire/

[12] H. B. Santoso, S. Fadhilah, I. Nurrohmah, and W. Goodridge, "The Usability and User Experience Evaluation of Web-based Online Self-Monitoring Tool: Case Study Human-Computer Interaction Course", in *The 4th International Conference on User Science and Engineering*, 2016.

[13] A. K. Nisa. (2016). *Usability Evaluation Sistem Informasi Manajemen Rumah Sakit: Studi Kasus Aplikasi Instalasi Gawat Darurat Rumah Sakit Umum Daerah Pasar Minggu* [Online]. Available: http://lontar.cs.ui.ac.id/Lontar/opac/themes/newui/detail.jsp?id=43510

[14] K. A. Nugroho and Herianto. Pengembangan Alat Bantu Rehabilitasi Pasien Pasca Stroke Berbasis Virtual Reality. Jurnal Teknik Industri, 2016, 11(1).

[15] R. K. Hambleton. (2005). Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures. In, R. K. Hambleton, P. F. Merenda, & C. D. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural assessment* (pp. 3-38). Marwah, NJ: Lawrence Erlbaum.

[16] J. C. Borsa, B. F. Damásio, and D. R. Bandeira. (2012). Cross-cultural adaptation and validation of psychological instruments: Some considerations. *Paidéia (Ribeirão Preto), 22*(53), pp. 423-432. [Online]. Available: http://dx.doi.org/10.1590/1982-43272253201314

[17] Usability.gov. (n.d.a). *Usability Evaluation Basics* [Online]. Available: http://www.usability.gov/what-and-why/usability-evaluation.html

[18] J. Nielsen. (2012). *Usability 101: Introduction to Usability* [Online]. Available: https://www.nngroup.com/articles/usability-101-introduction-to-usability/

[19] B. Shneiderman and C. Plaisant, *Designing the User Interface: Strategies for Effective Human-Computer Interaction, 5th edition*. Reading, MA: Addison-Wesley Publ. Co, 2010.

[20] C. Baber. (2005). Evaluation in human-computer interaction. In, J. R. Wilson, N. Corlett, editors, *Evaluation of human work*. Boca Raton: Taylor & Francis.

[21] Usability.gov. (n.d.b). *Usability Evaluation Methods* [Online]. Available: http://www.usability.gov/how-to-andtools/methods/usability-evaluation/index.html

[22] P. W. Jordan. (2001). Usability and product design. In, W. Karwowski, editor, *International encyclopaedia of ergonomics and human factors*. Boca Raton: Taylor & Francis.

[23] A. Dillon. (2006). Evaluation of software usability. In, W. Karwowski, editor, *International encyclopaedia of ergonomics and human factors*. Boca Raton: Taylor & Francis.

[24] J. Brooke. (1986). *SUS: A "Quick and Dirty" Usability Scale* [Online]. Available: http://cui.unige.ch/isi/icle-wiki/_media/ipm:test-suschapt.pdf

[25] A. Bangor, P. T. Kortum, and J. T. Miller. An empirical evaluation of the System Usability Scale. International Journal of Human-Computer Interaction, 2008, 24(6), pp. 574-594.

[26] J. Sauro. (2011). *Measuring Usability with the System Usability Scale (SUS)* [Online]. Available: http://www.measuringu.com/sus.php

[27] D. E. Beaton, C. Bombardier, F. Guillemin, and M. B. Ferraz. Guidelines for the process of cross-cultural adaptation of self-report measures. SPINE, 2000, 25(24), pp. 3186-3191.

[28] J. Nunnally and L. Bernstein, Psychometric Theory. New York: McGraw-Hill Higher, Inc., 1994.

[29] C. Heffner. (2016). *Chapter 7.3 Test Validity and Reliability* [Online]. Available: http://allpsych.com/researchmethods/validityreliability/

[30] M. Tavakol and R. Dennick. Making sense of Cronbach's alpha. International Journal of Medical Education, 2011, 2, pp. 53-55.

[31] D. George and P. Mallery, *SPSS for windows step by step: a simple guide and reference. 11.0 update*, 4th ed. Boston: Allyn & Bacon, 2003.

[32] OECD. (2013). *OECD Guidelines on Measuring Subjective Well-being* [Online]. Available: http://www.oecd.org/statistics/oecd-guidelines-on-measuring-subjective-well-being-9789264191655-en.htm

 978-1-5090-4629-4/16/$31.00