

# ISO Standards and Enterprise Software: A Case Study Using SUMI and SUS in an International Sale

Joyce Ohgi and Jeff Sauro

Oracle USA, 300 Oracle Parkway MS3op343, Redwood Shores, California 94065 USA  
joyce.etsuko.ohgi@oracle.com

**Abstract.** Usability standards play an important role in Europe, where companies are legally obligated to ensure that the workplace meets software ergonomic standards. Therefore, usability can be a critical factor in a company's decision about which enterprise software solution to implement. In this paper, we provide a real-world case study where meeting ISO standards and providing usability data to a customer heavily influenced a competitive sales deal.

**Keywords:** ISO, SUMI, SUS, international usability testing, enterprise software, user experience, software procurement.

## 1 Introduction

A prospective European customer included usability metrics as a requirement in its request for tender (RFT) document. The customer wanted to evaluate Oracle E-Business Suite 12 (EBS 12) applications along with other enterprise solutions. The customer wanted a clear and objective measurement of user perceptions regarding the usability of the software and how usability affected a user's ability to accomplish daily work. The customer requested Software Usability Measurement Inventory (SUMI) scores from all 10 of the enterprise applications that it was considering. The customer selected SUMI [1] because SUMI is referenced in ISO 9126 [2] and ISO 9241 [3].

The Oracle sales consultant working on the RFT reached out to the Oracle Applications User Experience team to provide the requested usability data. At the time, the Oracle user experience team had System Usability Scale (SUS) data for seven of the 10 applications under consideration and a SUMI score for one application, and could satisfactorily provide the correlation between SUS and SUMI data to meet the RFT requirement. Oracle lacked data for two of the applications, however, and needed to collect SUMI data for these two applications in less than two weeks to meet the RFT deadline.

The Oracle sales consultant identified an existing customer (located in the same country as the prospective customer) who was willing to provide qualified end users to participate in a study to collect the necessary data. Oracle flew a usability engineer to this customer site to conduct the study and collect SUMI data on the two remaining Oracle EBS R12 applications and to create a report in time to meet the RFT deadline.

## 2 Procedure

### 2.1 Materials

To collect the usability data, we needed to rapidly create, validate, and translate all of the following key materials:

- **User profile.** We needed to identify and validate the appropriate user profile for the applications being tested. The user profile ensured that we were conducting the study with the appropriate, qualified users. Oracle's existing repository of user profiles enabled us to quickly select the relevant user profile and obtain the presales team's validation.
- **Hosted test environment.** We needed a stable environment in which to run the test sessions, ensuring a consistent experience across all 10 users. Oracle's internal technical support team set up the environment for the test.
- **Task list of key product flows.** We needed a task list of the key product flows in the two enterprise applications being tested. We used two primary criteria to determine the key tasks: (1) The tasks needed to reflect the most common actions that a typical end user would accomplish within the applications, and (2) the average test session length needed to be around 1.5 hours, which is fairly representative for the usability tests conducted at Oracle. The task list that we created included six key tasks, with scenarios and sample data that participants needed to complete each of these tasks. Participants used this task list during the test sessions.
- **Test plan.** We needed a test plan for the usability engineer. The test plan that we created included: the task list; the expected starting point, click path, and ending point for each task; and the sample data that needed to be seeded in the test environment for each task.
- **Introductory screener and script.** To ensure that each participant met the criteria listed in the user profile, we created a document with a short list of questions asking about the participants' backgrounds and job roles. We also included a written introduction to the study in the users' native language. Although all of the users spoke both their language and English, we wanted to be certain that the users received the instructions in their native language, as the usability engineer spoke only English.
- **SUMI and SUS scores.** Although the RFT required only SUMI data for Oracle EBS Release 12 applications, we decided to collect both SUMI and SUS scores. Depending on the results, we knew that having both scores could give our statistical analyst further data to illustrate the correlation between SUMI and SUS. In addition, we opted to collect SUMI and SUS scores for both Oracle EBS 11 release 11.5.10 and Oracle EBS 12 applications. Again, depending on the results, we knew that these scores could show increased user satisfaction with the newer Oracle EBS 12 applications over the Oracle EBS 11 release 11.5.10 applications, demonstrating an improvement in usability over time with Oracle applications.

**Informatie deelnemer**

Naam: \_\_\_\_\_

Functie: \_\_\_\_\_

Hoe lang ben je werkzaam in je huidige rol?

\_\_\_ minder dan 1 jaar  
 \_\_\_ 1 - 3 jaar  
 \_\_\_ 3 - 5 jaar  
 \_\_\_ 5 - 10 jaar  
 \_\_\_ 10 jaar of meer

Hoe lang maak je gebruik van Oracle software?

\_\_\_ minder dan 1 jaar  
 \_\_\_ 1 - 3 jaar  
 \_\_\_ 3 - 5 jaar  
 \_\_\_ 5 - 10 jaar  
 \_\_\_ 10 jaar of meer

Welke Oracle software gebruik je op dit moment?

\_\_\_ Debiteuren  
 \_\_\_ Cash Management  
 \_\_\_ Anders: \_\_\_\_\_

Welke van de volgende taken behoort tot je functie:

\_\_\_ Klantgegevens up to date houden  
 \_\_\_ Facturen aanmaken  
 \_\_\_ Facturen aanpassen en betalingen verwerken  
 \_\_\_ Handmatig betalingen invoeren en toewijzen  
 \_\_\_ Anders: \_\_\_\_\_

**1. Introductie**  
 Bedankt dat de klant

**2. Instructies voor de**  
 Vandaag gaan we kij  
 expert in jouw functie  
 deze applicaties zul  
 starten zou ik willen  
 duidelijk is. Daarna g

Voor de duidelijkheid  
 misschien niet volled  
 dat je iets anders ka

Ik wil graag dat je ee

• **Vraag om hulp**  
 Wanneer je een t  
 Daarna zullen we  
 Als je het gevoel  
 Als je op enig md  
 aan.

• **Wees eerlijk**  
 We zijn echt op z  
 belangrijk dat we  
 kunnen verbeterd  
 feedback.

**3. We testen niet "jou**  
 Tot slot nog een bel  
 Besef alsblijft dat v  
 gebruikersinterface d  
 moeten maken.  
 Dus als je problemen  
**Nog vragen?**

**System Usab**

1. Ik denk dat ik dit sys  
 regelmatig wil gebruike

2. Ik vond het systeem

3. Ik vond het systeem  
 gebruiken

4. Ik denk dat ik onder  
 van een technisch pers  
 te kunnen gebruiken

5. Ik vond dat de versc  
 dit systeem erg goed g

6. Ik vond dat er leveel  
 in het systeem zaten

7. Ik kan me voorstelle  
 mensen zeer snel leren  
 gebruiken.

8. Ik vond het systeem  
 gebruik

9. Ik voelde me erg ver  
 systeem

10. Ik moest erg veel leren voordat ik aan  
 de gang kon gaan met dit systeem

1. Dit computerpro  
 2. Ik zou dit comp  
 3. De gebruiker we  
 verwacht  
 4. Het computersp  
 5. Het leren omga  
 6. Met dit compute  
 7. Ik vind het wer  
 8. De help-informa  
 9. Als het compute  
 10. Het duurt te lan  
 11. Ik vraag me som  
 12. Het werken met  
 13. De wijze waaro  
 14. Het geeft mij ee  
 gebruik  
 15. De documentati  
 16. Het kijkt alsof h  
 17. Het werken met  
 18. Er is nooit geno  
 19. Ik heb het gevo  
 20. Ik gebruik het liefst die voorzieningen van het computerprogramma die ik het best ken.

1     2     3     4     5

Fig. 1. Sample of test materials translated into a user’s native language

**2.2 Method**

From an existing Oracle customer located in the same region as the prospective customer, we identified 10 users who fit the identified user profile. These users were also familiar with Oracle EBS 11 release 11.5.10 applications.

The usability engineer flew out to the customer site to conduct the sessions in person. Each test session involved one participant and was limited to 1.5 hours. Each participant was greeted by the usability engineer and was oriented to the test procedure in his or her native language. The usability engineer explained how the evaluation would proceed. The participant read each task, attempted each task, and indicated when he or she was finished with each task. The user was given six tasks to complete with the Oracle EBS 12 applications. The user worked to the best of his or her ability to complete these tasks. At the end of a session, each user completed SUS and SUMI surveys (also translated into his or her native language) on the Oracle EBS 12 applications.

**3 Satisfaction Measures Overview**

**3.1 SUMI**

SUMI is an inventory questionnaire specifically designed to measure user satisfaction with products. SUMI data has been collected from hundreds of thousands of

participants. This data enables researchers to compare industry norms for satisfaction with usability on five main factors: efficiency, affect, helpfulness, control, and learning. In addition, there is a global factor that can be used when a single, global user-derived value is needed. Oracle typically uses this SUMI questionnaire as part of its usability methodology on software going through the usability labs in a post-release form.

## 3.2 SUS

The SUS questionnaire is another satisfaction measurement tool for computer systems [4]. SUS was developed as part of the usability engineering program in integrated office systems development at Digital Equipment Co. Ltd., Reading, United Kingdom. Oracle typically uses the SUS questionnaire as part of its usability methodology on software going through the usability labs in a post-release form.

## 3.3 SUMI vs. SUS Equivalency Scores

SUMI normative scores have been standardized against a proprietary dataset of other products. The SUMI median score of 50 represents an average score, meaning roughly half the products in the dataset have higher perceived usability and half have lower perceived usability. The product names and dates of data collection are not shared.

The SUS questionnaire is open-source with similar psychometric properties as the SUMI (valid and reliable). There is not a proprietary set of SUS products. However, because the SUS has been in use for more than 20 years, there are publicly available sets of data to generate a similar benchmark as the SUMI [6]. In addition, it has been shown that the SUS and SUMI have a strong correlation ( $r > 0.8$ )—a correlation that suggests that the majority (> 64 percent) of variation in SUMI scores can be accounted for by SUS scores [5].

The strong correlation between SUMI and SUS scores enabled us to generate a SUMI score that we expected to correspond closely to a SUS score. To generate our SUMI equivalency score from the SUS standardized questionnaire, we used the publicly available mean and standard deviations from [6], [7], and [8]. This data set represented more than 200 products and applications from dozens of companies and domains, rivaling the diversity and size of the SUMI database.

We took the raw SUS scores and converted them to a standardized score by subtracting each score from the mean and dividing this result by the standard deviation. The resulting normal scores (z-scores) were then converted to percentages similar to the SUMI.

To confirm the relationship between SUS and SUMI scores, we administered both questionnaires to users following a usability test and found the relationship to hold ( $r > .5$ ), further validating the SUS equivalency scores.

## 3.4 Improved Usability in Oracle EBS 12

Much of our data came from Oracle EBS 11 releases 11.5.9 and 11.5.10. There has been substantial improvement to many features and functions in Oracle EBS 12, as well as focused improvement on usability. To understand how much usability improved, we gathered data on both the older (11.5.10) and newer releases (12) of

two applications. We found both SUMI and SUS scores to be significantly better for Oracle EBS 12 ( $p < .05$ ), with an approximate improvement in usability of 10 percent (the average improvements seen between SUS and SUMI questionnaires).

### 3.5 Sample Size

While a larger sample size is better, it has been shown that sample sizes of around 10 provide a tolerable amount of variability in the estimate of the central tendency of the scores. The number 10 is not a hard cut-off; rather, it is an approximate point in the same way that the actual variability in the mean score is a function of both the sample size and variability between respondents scores [9]. Therefore, sample sizes of 10 +/- 2 will likely not have a material difference in the stability of the mean.

We collected SUS and SUMI survey results on the customer's existing Oracle EBS 11 release 11.5.10 applications, as well as SUS and SUMI survey results for Oracle EBS 12 applications. The results gave us user feedback on Oracle EBS 11 release 11.5.10 and Oracle EBS 12 applications. We collected both SUMI and SUS data to confirm the correlation between the two scores. After all 10 sessions were completed, we analyzed the data and provided the necessary findings report on the two applications to be included in the RFT. In addition, we provided SUMI equivalency scores for the SUS data that we had for the other seven applications. To generate the SUMI equivalency scores from the SUS data, we used the publicly available mean and standard deviations from [6], [7], and [8].

## 4 Key Success Factors

### 4.1 Cross-Functional Team

Assembling a cross-functional team of a sales consultant, technical support engineer, usability engineer, and statistical analyst was crucial to the success of this project. Working together, we were able to quickly create and validate the materials needed, set up the test environment, run the study, and provide the data needed by the RFT deadline. Here are descriptions of each team member's responsibilities:

- **Sales consultant.** The project started with the sales consultant identifying usability as a differentiating factor in the prospective sale and reaching out to Oracle's Applications User Experience team for assistance. The sales consultant, located in the same country as the prospective customer, served a number of functions. She provided the rest of the team with the history between Oracle and the prospective customer, identified the specific type of data needed for the RFT, reached out to the existing customer with users matching the user profile, helped seed the data into the test environment, and served as a document translator.
- **Technical support engineer.** The technical support engineer set up the test environment in the users' native language, helped seed the sample data, and ensured the environment's stability for the duration of the usability study.
- **Usability engineer.** The usability engineer identified the necessary user profile appropriate for the two applications that lacked SUMI data and obtained the sales

consultant's validation. The usability engineer also created the task list, test plan, screener, and introductory script. Once on site at the customer location, the usability engineer moderated the 10 sessions, collected the SUMI and SUS scores, and sent the data to the statistical analyst.

- **Statistical analyst.** Before the usability test, the statistical analyst assisted in acquiring the SUMI and SUS collection documents in the users' native language. After the data was collected, the statistical analyst provided the data in the format required by the RFT and included analysis to support the correlation between SUS and SUMI data.

## 4.2 User Experience Group as a Valid Partner to Sales

Another key factor was demonstrating to the sales team and the prospective customer that the Oracle Applications User Experience team possessed the agility and expertise to meet its needs in the 2-week time frame provided. By providing timely, real-world support for an important prospective sale, we dispelled any misconceptions that usability work is slow, inflexible, or applicable only in a lab.

In addition, by engaging with the existing customer while collecting the needed usability data on site, the usability engineer fostered a lasting relationship with the customer and additional contacts in that region, resulting in additional activities for the user experience team, such as usability test sessions organized for the region's Oracle user group conference.

## 4.3 Usability Data as a Sales Differentiator

Although a number of other dimensions were included in the RFT, we later discovered that the key competitor in this sale was not able to provide any usability data for its enterprise applications. For the usability section of the RFT, the competitor received no points, thus significantly impacting its overall score on the RFT. Oracle ended up winning the contract.

## 5 Conclusions

As countries continue to adopt regulations that require companies to provide a workplace environment that complies with ISO standards, usability metrics will increasingly become a factor for companies in choosing which enterprise software solution to adopt. Enterprise software companies will benefit from (1) making it standard practice to use commonly accepted methods like administering the SUMI and SUS questionnaires to collect data on their key products and (2) being prepared to quickly assemble agile teams to collect additional data should the need arise.

## References

1. Kirakowski, J., Corbett, M.: SUMI: the Software Usability Measurement Inventory. *Brit. J. Ed. Technol.* 23, 101–102 (1994)
2. International Organization for Standardization: ISO 9126-2. *Software engineering — Product quality — Part 2: Internal Metrics* (2005)

3. International Organization for Standardization: ISO 9241-110. Dialogue principles: Guidance on usability (2006)
4. Brooke, J.: SUS: a quick and dirty usability scale. In: Jordan, P., Thomas, B., Weerdmeester, B., McClelland, I. (eds.) *Usability Evaluation in Industry*, pp. 189–194. Taylor and Francis, London (1996)
5. Hoyler, A.: *Methods for Evaluating User Interfaces*. Doctoral Research, University of Sussex, Brighton (1993)
6. Bangor, A., Kortum, P., Miller, J.: An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction* 24, 574–594 (2008)
7. Lewis, J., Sauro, J.: The Factor Structure of the System Usability Scale. In: *Human Computer Interaction International Conference*, pp. 94–103. Springer, Heidelberg (2009)
8. Tullis, T., Albert, B.: *Measuring the User Experience: Collecting, Analyzing and Presenting Usability Metrics*. Morgan Kaufman, Burlington (2008)
9. Tullis, T., Stetson, J.: Comparison of Questionnaires for Assessing Website Usability. In: *Usability Professionals Association (UPA)*, Minneapolis, vol. 2004, pp. 7–11 (2004)