

# Operating Systems Usability: A Comparative Study

**URBANO B. PATAYON**

<http://orcid.org/0000-0002-1295-2151>

[patayonurbano233@gmail.com](mailto:patayonurbano233@gmail.com)

Jose Rizal Memorial State University-Tampilisan Campus  
Znac, Tampilisan, Zamboanga del Norte

**NERICO L. MINGOC**

[nerico.mingoc@g.msuiit.edu.ph](mailto:nerico.mingoc@g.msuiit.edu.ph)

Mindanao State University- Tawi-Tawi  
Tawi-Tawi, Philippines

Originality: 100% • Grammar Check: 100% • Plagiarism: 0%



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

---

## ABSTRACT

As usability testing becomes more popular and widely recognized, operating systems' users are still relying on reviews that are based on the price, standard feature and satisfaction survey as to which product will be patronized. Measuring usability requires assessment on three product attributes or factors namely: effectiveness, efficiency, and user satisfaction. There are thirty-seven (37) respondents used in the study. Each respondent is required to perform the given task in each version of the Windows operating system. Time to complete the task and behavioral manifestations were recorded. Based on the data gathered and analyzed, results show that Windows 10 has the most number of the task with the highest completion rate in comparison with two operating systems in the study. Regarding efficiency, Windows 8 has the highest average task completion time. As to user satisfaction, the majority of the respondents were frustrated in the different task under Windows 8 while most of them are delighted in the tasks

under Windows 7. Regarding engagement and boredom, the result reveals that users are engaged at the same time felt bored on tasks under Windows 10.

**Keywords** — Usability, effectiveness, efficiency, user satisfaction, operating system, Philippines

## INTRODUCTION

The market is saturated with competing brands claiming to be superior to others. Product reviews serve as a tool for end users/customer to select which product is best fitted for their needs. Some companies see this as a factor and baseline for researching and developing products with user-oriented methods instead of technology-oriented methods (Holm, 2006).

ISO 9241-11 defines usability as 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. It is the ease of use and learnability of a human-made object such as a tool, software application, website, book, machine, process, vehicle, or anything a human interacts with (Brödner & Adler, 1995). In human-computer interaction and computer science, usability focuses the elegance and clarity with which the cooperation with a computer program or a website is designed (Nielsen & Levy, 1994). Usability becomes an essential factor since it will help companies and organizations in achieving its own goals because its primary concern is the productivity of the user (Mifsud, 2011). Usability is also important in website development because according to Jakob Nielsen (1994), “Studies of user behavior on the Web find a low tolerance for intricate designs or slow sites.”

According to ISO 9241 part 11, usability is consists of three aspects: effectiveness, efficiency, and satisfaction. Effectiveness is the accuracy and completeness with which users achieve specific goals (Frøkjær, Hertzum, & Hornbaek, 2000). Efficiency indicators include input rate, mental effort, usage patterns, communication effort, learning measures and time controlled (Hornbaek, 2006). On the other hand, efficiency is the relation between the accuracy and completeness with which users achieve specific goals (Baily, 1993) and the resources expended in performing them (Bevan, 1995). Its indicators are preference, ease-of-use, and attitude (Hornbaek, 2006). The third and the last usability aspect as stipulated in ISO 9241 is satisfaction. It is the users’ comfort with and positive attitudes towards the use of the system. Users’ satisfaction can be measured by attitude matrix and rating scales (Hornbaek, 2006). These

three are considered independent aspects of usability, and for usability testing of computer systems having complex tasks, measures of efficiency, effectiveness, and user satisfaction must be included (Frøkjær, Hertzum, & Hornbaek, 2000).

As usability testing becomes more popular and widely recognized (Holm, 2006), operating systems' users are still relying on reviews that are based on the price, standard feature and satisfaction survey as to which product will be patronized. Some companies and online websites such as GlobalStats and Stack Overflow are doing customer satisfaction survey and market share review as to which operating system mostly preferred by the buyers and end users. As to the perspective of Erik Frøkjær, Morten Hertzum, and Kasper Hornbaek (2000), measuring only a subset such as the satisfaction of the three usability aspects is an insufficient indicator of overall usability. Also, anchoring to this perspective, the researchers are encouraged to conduct the study on usability comparing the four commonly used Windows operating system; the Windows 10, Windows 8 and Windows 7.

## OBJECTIVES OF THE STUDY

The researchers aim to achieve the following, (1) to determine the hardware specification of the computers to be used; (2) to evaluate Windows 7, 8 and 10 operating system versions based on the three aspect of usability; and (3) to identify the differences of Windows 7, 8, and 10 operating system versions in terms of three aspect of usability.

## METHODOLOGY

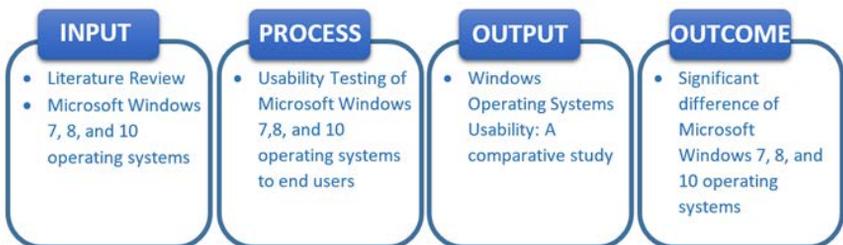


Figure 1. The research design of the Study

To start the preliminary phase of this study, reviewing related studies were made. As observed, Microsoft is the most used operating systems for desktop computers, especially in the Philippines. But, it was not clear as to what type or version of the Microsoft operating system is fitted for users. To differentiate Windows 7, 8, and 10 versions of operating systems, the researchers conducted usability testing to end users. These will have equal results compared to users that are highly exposed or have a high literacy to computers. As an outcome, this study will determine the significant differences between Microsoft Windows 7, 8, and 10 versions of operating systems tested by end users.

### Hardware Specifications

To determine the hardware specifications to be used. According to Microsoft, the three versions of operating systems under study will require the following minimum requirements to be used:

Table 1. Microsoft Windows 7 Operating System Version Minimum Requirements

|              |   |
|--------------|---|
| Processor    | 1 gigahertz (GHz) or faster 32-bit (x86) or 64-bit (x64) processor                                |
| Memory (RAM) | 1 gigabyte (GB) RAM (32-bit) or 2 GB RAM (64-bit)   |
| Storage      | 16 GB available hard disk space (32-bit) or 20 GB (64-bit)  |
| Graphics     | Microsoft DirectX 9 graphics device with WDDM (Windows Display Driver Model) 1.0 or higher driver |

Table 2. Microsoft Windows 8 Operating System Version Minimum Requirements

|              |  |
|--------------|--|
| Processor    | 1 gigahertz (GHz) or faster with support for PAE (Physical Address Extension), NX (No-eXecute), and SSE2 (Streaming SIMD Extensions 2) |
| Memory (RAM) | 1 gigabyte (GB) RAM (32-bit) or 2 GB RAM (64-bit)  |
| Storage      | 16 GB available hard disk space (32-bit) or 20 GB (64-bit)   |
| Graphics     | Microsoft DirectX 9 graphics device with WDDM driver   |

Table 3. Microsoft Windows 10 Operating System Version Minimum Requirements

|              |   |
|--------------|---|
| Processor    | 1 gigahertz (GHz) or faster processor or SoC (System on Chip)                           |
| Memory (RAM) | 1 gigabyte (GB) RAM (32-bit) or 2 GB RAM (64-bit)                                       |
| Storage      | 16 GB available hard disk space (32-bit) or 20 GB (64-bit)                              |
| Graphics     | DirectX 9 graphics device with WDDM (Windows Display Driver Model) 1.0 or higher driver |

### Participants

In order to get statistically significant numbers, Nielsen (1994) stipulates that a usability test can be deployed in at least twenty (20) users applicable to quantitative studies. For eye tracking, at least thirty-nine (39) users are necessary to provide a stable heat map. To assure homogeneity, respondents were selected based on predefined criteria such as the level of computer literacy. Respondents were given a structured survey questionnaire. Only those students whose responses are greater than 90% to the question if they have less exposure to the use of computer were included.

### Data Gathering Procedure

For the researcher to gather data, questionnaire and observation method was used. The instrument was administered personally by the researchers after seeking approval from the dean of the College of Agriculture and Technology of Jose Rizal Memorial State University-Tampilisan Campus through the duly signed communication letter. Data gathering procedures used in the study were designed to solicit information for usability and three of its measures namely: effectiveness, efficiency, and satisfaction (Hornbæk, 2006). The researchers had organized twenty (20) tasks to be performed by the respondents. These tasks are commonly found in the three operating systems under this study. The respondents used computer stations based on their availability.

*Effectiveness.* In measuring effectiveness, a task completion rate was used by the researchers. Task completion rate is a usability key performance indicator. It is a number or percentage of tasks that users completed (Sismeiro & Bucklin, 2004).

*Efficiency.* Efficiency defines as resources expended about the accuracy and completeness with which users achieve goals (ISO, 1998). Efficiency in this study was measured using task completion time (Frekjm et al., 2000). Task Completion Time is the average of individual task time for a single attempt (Desai et al., 2008).

*Satisfaction.* According to Frøkjær, Hertzum, and Hornbmk (2000), satisfaction is the users' comfort with and positive attitudes towards the use of the system. Users' satisfaction can be measured by attitude rating scales such as SUMI (Kulkarni, Padmanabham, Sagare, & Maheshwari, 2013) and IBM Computer Usability Satisfaction Questionnaires (Lewis, 1995). In the study, the researchers used a behavioral/psychological matrix to record the attitude of the respondents towards each given task. There were four attitudes were observed and recorded in the study, these are the frustration, delight, engage, and boredom. For the researchers to identify the different attitudes, observation in the facial expression (Farnsworth, 2016) and gesture (Castellano, Kessous, & Caridakis, 2008) of the respondents in every task were done. The behavioral observation was performed without the respondent's awareness to eliminate biases.

## **Data Analysis**

To derive comparisons from the observation and responses of the participants, and arrive at the correct analysis and interpretation of data, the researchers used the following statistical tools:

*Effectiveness.* To evaluate effectiveness, a task completion rate was used. Task completion rate is calculated by dividing the number of assigned employees who completed tasks successfully by the total number of assigned employees (Sismeiro & Bucklin, 2004). According to Sauro (2011), the good completion rate is 78% since it is above the quartile which is 75%. To identify the significant difference in task completion rate between the three operating systems, analysis of variance (ANOVA) was used. ANOVA was used by the researchers based on the following assumptions: a) Data is interval or ratio in scale and normally or approximately normally distributed (Reston, 2004). b) Variances are homogeneous across treatments/groups (Reston, 2004).

*Efficiency.* Efficiency is measured using task completion time (Frekjmjr et al., 2000). Task Completion Time is the sum of individual task time for a single attempt (Desai et al., 2008). In the study, the researchers set specific time allotted for each task to be accomplished. Respondents were not informed regarding the time allotted for each task to avoid being pressured which affects behavior/attitude. Time allotment was used to identify if the user is efficient in all the tasks or not (Hornbeak, 2006). To quantify the significant difference in task completion time between the three operating systems, the researchers used the analysis of variance (ANOVA). ANOVA was used by the researchers based on the assumptions mentioned.

*Satisfaction.* To measure the degree of user satisfaction, frequency count and simple percentage were used. To quantify the significant difference in task completion time between the three versions of operating systems used in the study, the researchers utilized the analysis of variance (ANOVA).

## RESULTS AND DISCUSSION

### Hardware Specification

To eliminate hardware bias in the study, the researcher used the same computer specification where the three operating systems were installed. Specifically, the following were the details of the three computer unit being used:

Table 4. Hardware Specification Used In Usability Testing

|              |                                  |
|--------------|----------------------------------|
| Processor    | 3.70 gigahertz (GHz)             |
| Memory (RAM) | 4 gigabyte (GB)                  |
| Storage      | 500 GB available hard disk space |
| Graphics     | Intel graphics device            |

### Effectiveness

The Task Completion Rate of Windows 7, Windows 8, and Windows 10 as shown in Figure 1 shows that Windows 10 has the highest average task completion rate of 75.33% while Windows 8 has the lowest average completion rate of 60.41%. The result implies that Windows 10 has the most numbered of the completed task while Windows 8 has the least. Further, both operating systems are rated not good based on 78% standard completion rate.

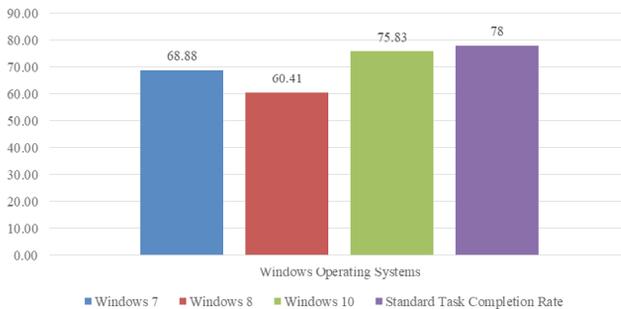


Figure 1. Task Completion Rate of Windows 7, Windows 8, and Windows 10

Table 5 presents the analysis of variance (ANOVA) to determine the difference between Windows 7, Windows 8 and Windows 10 in terms of Task Completion Rate. The table reflects the mean of Windows 10 (75.83), Windows 8(60.41), and Windows 7(68.88). The table further shows that computed p-value at 0.05 alpha is 0.0023 which is interpreted as significant. The result implies that there is a significant difference in the Task Completion Rate of the three versions operating systems namely; Windows 7, Windows 8, and Windows 10.

Table 5. Analysis Of Variance to Determine the Difference between Windows 7, Windows 8, and Windows 10 In Terms Of Effectiveness

| Operating System Versions | Mean (Task Completion Rate (%)) | df | p-value @ 0.05 alpha | Interpretation |
|---------------------------|---------------------------------|----|----------------------|----------------|
| Win. 10                   | 75.83 <sup>A</sup>              |    |                      |                |
| Win. 8                    | 60.41 <sup>AB</sup>             | 2  | 0.0023               | Significant    |
| Win. 7                    | 68.88 <sup>B</sup>              |    |                      |                |

### Efficiency

The Average Task Completion Time of Windows 7, Windows 8, and Windows 10 as shown in figure 2 shows that Windows 8 has the highest task completion time of 28.53 minutes while Windows 10 has the lowest average completion time of 20.24 minutes. This implies that Windows 8 has the longest while Windows 10 has the shortest average time of completion in all the tasks. Both operating systems are efficient since the average task completion of the three versions of the operating system is below the allotted time.

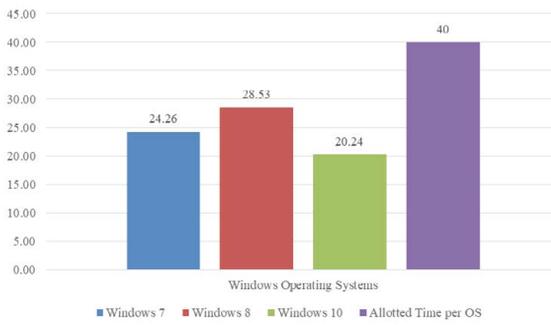


Figure 2. Average Task Completion Time of Windows 7, Windows 8, and Windows 10

Table 6 presents the analysis of variance (ANOVA) to determine the difference between Windows 7, Windows 8 and Windows 10 in terms of Task Completion Time. The table reflects the mean of Windows 10 (20.24), Windows 8(28.53), and Windows 7(24.26). The table further shows that computed p-value at 0.05 alpha is 0.0267 which is interpreted as significant. The result implies that there is a significant difference in the Task Completion Time of the three operating system versions namely; Windows 7, Windows 8, and Windows 10.

Table 6. Analysis of Variance to Determine the Difference between Windows 7, Windows 8, and Windows 10 In Terms Of Efficiency

| Operating System Versions | Mean (Task Completion Time(min.)) | df | p-value @ 0.05 alpha | Interpretation |
|---------------------------|-----------------------------------|----|----------------------|----------------|
| Win. 10                   | 20.24 <sup>A</sup>                |    |                      |                |
| Win. 8                    | 28.53 <sup>AB</sup>               | 2  | 0.0267               | Significant    |
| Win. 7                    | 24.26 <sup>B</sup>                |    |                      |                |

**Satisfaction**

The user satisfaction percentage of end-user towards tasks in Windows 7, Windows 8, and Windows 10. In terms of delight, Windows 7 has the highest percentage of 34.59% while Windows 8 has the lowest percentage of 31.73% among the other two operating systems. In terms of engagement, Windows 10 has the highest percentage having 37.02% while Windows 8 has the lowest percentage of 31.40%. In terms of boredom towards every task, Windows 10 has the highest percentage of 37.84% while Windows 8 has the lowest percentage of 27.03. In terms of frustration, Windows 8 is the highest with a percentage of 36.25% while Windows 10 has the lowest percentage of 30.00%.

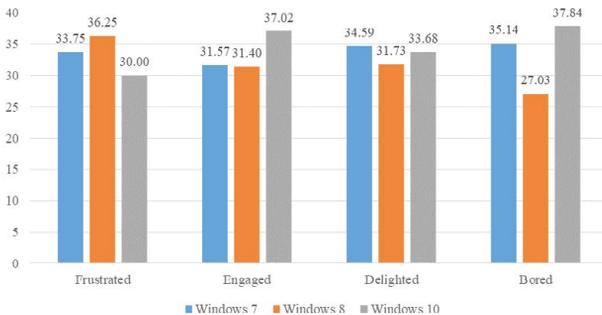


Figure 3. User Satisfaction Percentage

Table 7 presents the analysis of variance (ANOVA) to determine the difference between Windows 7, Windows 8 and Windows 10 in terms of user satisfaction. The table reflects the mean of Windows 10 (2.55), Windows 8(2.40), and Windows 7(2.50). The table further shows that computed p-value at 0.05 alpha is 0.4687 which is interpreted as not significant. The result implies that there is a no significant difference in terms of user satisfaction of the three versions operating systems but Windows 8 has the lowest user satisfaction mean.

Table 7. Analysis of variance to determine the difference between Windows 7, Windows 8, and Windows 10 in terms of user satisfaction.

| Operating System Versions | Mean (User Satisfaction) | df | p-value at 0.05 alpha | Interpretation  |
|---------------------------|--------------------------|----|-----------------------|-----------------|
| Win. 10                   | 2.55 <sup>A</sup>        |    |                       |                 |
| Win. 8                    | 2.40 <sup>A</sup>        | 2  | 0.4687                | Not Significant |
| Win. 7                    | 2.50 <sup>A</sup>        |    |                       |                 |

## CONCLUSIONS

Measuring usability requires assessment on three product attributes or factors namely: effectiveness, efficiency, and user satisfaction.

Based on the data gathered and analyzed, results show that Windows 10 has the most number of the task with the highest completion rate in comparison with the other two operating systems in the study. Regarding efficiency, Windows 8 has the highest average task completion time. This means that given tasks require a longer time to be accomplished in Windows 8. As to user satisfaction, the majority of the respondents were frustrated in the different task under Windows 8 while most of them are delighted in the tasks under Windows 7. Regarding engagement and boredom, the result reveals that users are engaged at the same time felt bored on tasks under Windows 10.

Further, examined data had shown that there is a significant difference in terms of task completion rate and task completion time of Windows 7, Windows 8, and Windows 10. Regarding user satisfaction, results show no significant difference between the behavior of the respondents toward each task per operating system but Windows 8 has the lowest satisfaction rate among OS under the study.

Given the findings, this study recommends to perform usability testing on other features and highly technical matters of Microsoft Operating systems, compare other operating systems aside from Microsoft Windows operating

system, preferably open source operating systems, and conduct usability testing on other usability contributors that have different factors of usability aside from those based on ISO 9241.

## TRANSLATIONAL RESEARCH

In the broader aspect, the result of the study helps end users in selecting which operating system is more user-friendly. In the school setting, the study will help the MIS Officer to identify which operating system will be used to enhance productivity.

## LITERATURE CITED

- Bevan, N. (1995). Measuring usability as quality of use. *Software Quality Journal*, 4(2), 115-130. Retrieved from <https://doi.org/10.1007/BF00402715>
- Brödner, P. (1995). Adler, Paul S.(ed.): Technology and the future of work, New York: Oxford University Press, 1992; and Adler, Paul S. and Winograd, Terry A.(eds.): Usability: Turning Technologies into Tools, New York: Oxford University Press 1992. *International Journal of Human Factors in Manufacturing*, 5(2), 227-230. Retrieved from <https://doi.org/10.1002/hfm.4530050209>
- Castellano, G., Kessous, L., & Caridakis, G. (2008). Emotion recognition through multiple modalities: face, body gesture, speech. In *Affect and emotion in human-computer interaction* (pp. 92-103). Springer, Berlin, Heidelberg. Retrieved from [https://doi.org/10.1007/978-3-540-85099-1\\_8](https://doi.org/10.1007/978-3-540-85099-1_8)
- Farnsworth, B. (2016, December 6). Facial Action Coding System (FACS) - A Visual Guidebook. Retrieved from <https://imotions.com/blog/facial-action-coding-system/>
- Frøkjær, E., Hertzum, M., & Hornbæk, K. (2000, April). Measuring usability: are effectiveness, efficiency, and satisfaction really correlated?. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 345-352). Retrieved from [doi>10.1145/332040.332455](https://doi.org/10.1145/332040.332455)

- Holm, O. (2006). Integrated marketing communication: from tactics to strategy. *Corporate Communications: An International Journal*, 11(1), 23-33. Retrieved from <https://doi.org/10.1108/13563280610643525>
- Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International journal of human-computer studies*, 64(2), 79-102. Retrieved from <https://doi.org/10.1016/j.ijhcs.2005.06.002>
- Kulkarni, R., Padmanabham, P., Sagare, V., & Maheshwari, V. (2013, August). Usability evaluation of PS using SUMI (software usability measurement inventory). In *2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 1270-1273). IEEE. Retrieved from <https://bit.ly/2XvTOD1>
- Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, 7(1), 57-78. Retrieved from <https://doi.org/10.1080/10447319509526110>
- Mifsud, J. (2011). An extensive guide to web form usability. *Retrieved*, 3(09), 2014. Retrieved from <https://bit.ly/2SBZvLK>
- Nielsen, J., & Levy, J. (1994). Measuring usability: preference vs. performance. *Communications of the ACM*, 37(4), 66-76. Retrieved from <https://bit.ly/2EolssB>
- Sauro, J. (2011). What is a good task-completion rate. *MeasuringU. Available online (last accessed November 2016) at: http://www.measuringu.com/blog/task-completion.php*. Retrieved from <https://bit.ly/2EnugiE>
- Sismeiro, C., & Bucklin, R. E. (2004). Modeling purchase behavior at an e-commerce web site: A task-completion approach. *Journal of marketing research*, 41(3), 306-323. Retrieved from <https://doi.org/10.1509/jmkr.41.3.306.35985>