

Contents

Module_10 – Usability Testing

10.1 - Introduction and Background	2
10.2 - Purpose and Importance of Usability	2
10.3 - Scope and Environment of Usability	5
10.4 - Different Methods of Usability Testing	9
10.5 - Environment in Usability Testing.	10
10.6 - Extended Resources	12
10.7 - References	13
10.8 - PowerPoint Presentation–Refer to D2L	

10.1 - Introduction and Background

Usability testing was created to maximize user satisfaction. Bad design will hinder use, comfort, and gratification, lowering the chances of a user becoming a returning customer. Usability testing can greatly improve the overall design.

What makes a well-designed product? Ease, simplicity, and eye-catching visuals all come to mind. No one wants to fumble around for hours trying to do a simple task. A great example would be the large amount of iPhone users. Why does a company that started out making computers now have one of the most popular smart phones? Design. Nevertheless, it is not simple to create good usability testing. One must understand the purpose of the product, the user, the methods of testing, and one must know how to analyze this information. After all of that one must still come to a conclusion. Was this experiment successful? Should we repeat these steps? Should we change any variables? All of these things must be questioned to complete a well-done usability test.

10.2 – Purpose and Importance of Usability

Before discussing the details, importance, and examples of usability testing in computers, let's first talk about what exactly the term means and its historical background. Usability can be defined 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” (Barnum 2011). With this formal definition of usability, usability testing can properly be defined as an activity that focuses on observing users working with a product, performing typical daily tasks and determining whether the product needs to be improved. The primary goal of usability testing is to improve the product to fit the users' needs and to also keep from having the same problems happen in future developments. Usability is an attribute in every product—even those that are not part of the computer field.

The practice of usability testing started around the early 1980s and was known to be very costly and time consuming. Only usability experts who had extensive learning in cognitive science, experimental psychology, or human factor engineering could conduct these tests. These experiments would have consisted of 30 to 50 people to be considered accurate. This went on for many years until the early 1990s, when Jakob Nielsen and Tom Landauer discovered a new experimental method for usability studies. They found that testing five users would still yield accurate data. While it may take a few more tests and revisions to get the project flawless for the users, it will ultimately still save a massive amount of time and money. Nielsen called this approach the “discount” usability test. In his paper on the introduction to usability he states that “it is important to test users individually and let them solve any problems on their own. If you help them or direct their attention to any particular part of the screen, you have contaminated the test results” (Nielsen 2012). Another famous quote from Nielsen is “your best guess is not good enough” which is why usability testing is essential to development. The knowledge a developer gains from conducting a usability test is very crucial to the design and development process.

There can be certain times when it is appropriate to conduct large studies and other times when small studies are more suited. To determine when it is more useful and appropriate to focus on a large study group is determined by specific product factors. Larger study groups are mostly used when focusing on a summative usability testing. Summative testing is performed “after the product is finished, with a goal of establishing a baseline of metrics or validating that the product meets requirements” (Barnum 2011). A larger study group is needed for a summative test because they produce metrics, charts, success rates, average time on task, search results, etc. With this type of return data, developers can determine what to possibly change in the future or decide that the product met all of the requirements. In many cases these tests are given while the product is already in use.

A smaller study group focuses on a completely different set of product factors and testing instances. Smaller study groups are mostly used when focusing on a formative testing. Formative testing is conducted “while the product is in development, with a goal of diagnosing and fixing problems; repeated during development” (Barnum 2011). Most development teams focus on smaller or formative testing because it provides them with a list of findings to analyze and fix throughout the development process. Later, they will test again to see if they had fixed the previous findings. Formative studies have also been useful in showing what users like and don’t like about the design or function of the product. The process of conducting many small studies can be defined as iterative testing: “The essential advantage of iterative testing is that you can learn from users, make changes based on what you learned, then test again” (Barnum 2011).

Now that the definition, history, and the difference in small/large usability testing’s have been discussed we can now talk about the people who are in these tests and why they are so important. The people who are in these usability tests must be members of the group who will be using the product when completed. This is a major requirement, because they are the ones who are going to be using the product, after all. It would be pointless to conduct a usability test on those who will not be using the product or on those who developed it. There can be quality assurance or expert reviewers who test the product and their suggestions should be taken under consideration as well, but these tests won’t deliver the same kind of results: “If the participants are more experienced than actual users, you may miss problems that will cause the product to fail in the marketplace. If the participants are less experienced than actual users, you may be led to make changes that aren’t improvements for the real users” (Dumas 1999). This statement shows how important it is to have multiple users in a test group and to use different users in the each usability test to get accurate results.

For a product to be successful during the usability testing, developers must first understand those for whom they are writing the product. A developer should sit down with the users and try to understand the tasks they intend to perform. A developer cannot write a program for someone else without knowing what exactly they are going to use it for and why: “People connect usability with productivity, because no one gets paid for time spent just sitting at a computer. They get paid for processing invoices, or for noticing and resolving alarms on the computer network, or for analyzing samples in a laboratory machine” (Dumas 1999). If the product is consistent, efficient, and easy to use then users will appreciate the product and be able to get the most possible usage out of it. Developers design to make people’s lives easier, and usability testing is designed to make the developers lives easier by helping detect problems and errors that may be in the product. Without usability testing, there would be thousands of unhappy customers throughout the world.

Usability testing is extremely important in today’s web development and program design. If a user finds a website or program difficult or confusing to use, they will leave and find a better version of what they are looking for. As an example, “if the user perceives that the online bill-paying feature offered by her bank

is not worth the effort to set up and use, then she will continue to write checks, put stamps on envelopes, and mail in her payments” (Barnum 2011). Our goal is to make sure the product fits the users’ needs not what the developer thinks is best for the product. If the product cannot fulfill the goals of the user what is the purpose of creating it? For successful testing the developer needs to look at five quality components that make up a great product:

- **Learnability:** How easy is it for the users to understand the functions and abilities of the product?
- **Efficiency:** Does the product have accurate information and quick response time?
- **Memorability:** If the user accessed the product months after the implementation date, could they remember the steps to use the product?
- **Errors:** Does the user find any errors while using this product? How severe are the errors?
- **Satisfaction:** Are the users happy with the design and purpose of the product?

With these five quality components, usability testing can be more important than the price or performance of a development, because users being able to use the interface and understand it can save the company time and money on support lines. The greater the results of the usability testing, the more of an impact the product can have on the users. It is often shown that if the users are more productive with their work, they will purchase more products in the future. Users that are happy with a company’s product can also promote that company’s work by spreading the word to other interested customers. If instead they were not than satisfied, they will be hesitant to buy any future products.

In some usability testing groups, testers assign user representatives, which help the users stay connected throughout the testing process. The representative can also help the developers stay on track with the testing and make sure they get all of the necessary information to make sure the test is successful. A user representative would focus on defining the user’s jobs and tasks, keep the team focused on the users’ needs, and give daily input on how to correct or modify a design idea. The person who is usually appointed as the user representative is more experienced and knowledgeable with computers than the other users. They will most likely know more about the job or task the development is based on as well. An example of a user representative may be a manager, assistant manager, or any type of supervisory position in the work place.

Sellers who are in charge of creating the product benefit from usability testing by:

- Selling more of the product
- Selling other products
- Enhancing the company’s reputation
- Saving money on internal products
- Reducing support costs
- Reducing training costs
- Reducing the needs for updates and maintenance releases
- Making documentation and training easier to develop

(Dumas 1999)

Usability is very important to customers and the sellers of the product. Online today, everyone almost always goes straight to the product reviews before buying a product. If a development team does not conduct a usability test, then they are jeopardizing their name and potential customers. Displaying that care was taken

in the development of a product can make a difference in sales as well. For instance, “the manufacturer of a software-driven laboratory machine found from customer surveys that their professional-quality documentation was a major factor in customers choosing their machine over its competitors. It wasn’t only that the documentation made the machine easier to use. Customers felt that they could trust machines from a company that cared enough to put the time and effort into good documentation” (Dumas 1999).

In the benefits listed above, Dumas stated that saving money on internal products is a huge factor in saving a company money. One should consider usability testing as equal or double the equivalent cost: no one should spend money on development and then turn around and have to spend even more money on support lines helping users struggle with an unproductive product. If a program stops or interferes with users’ productivity, the developer will be in deep trouble. This is why it is always important to develop with care, whether it be for an internal or external product. There is nothing more troublesome than having to contact a local help desk, to have to put in an issue for the development team and then to just sit and wait for them to fix the problem. (Dumas 1999)

10.3 - Scope and Environment of Usability

Testing the usability of a user interface is key to providing an efficient system, and understanding the scope of the software is critical when developing a testing method. Scope can be defined as the purpose and goals of the software. This section will explain how understanding what software is supposed to do and the crowd of people this software is intended for effects the approach of how developers test their software.

In today’s world, technology has become so advanced that people turn to technology to make their everyday lives easier. For instance, the article “Robot services for elderly with cognitive impairment: Testing usability of graphical user interfaces” talks about the development of socially assistive robotics for elderly. This new technology is intended to assist elderly people with everyday tasks. In order to develop such an advanced system, the intended users need to be taken into consideration. Since the scope of the software is to help elderly people with daily tasks using robotics, one must take in account the needs and knowledge these elderly people may have. Factors such as age or disease-related changes in perceptual or cognitive abilities and familiarity with computer technologies influence technology use that must be considered when designing interfaces for these users (C. Granata, G. Legouverneur, J.-S. Vidal, P. Bidaud, and A.-S. Riguid, 2013). With these factors in mind, an experiment was constructed in order to test the user interface of such a system.

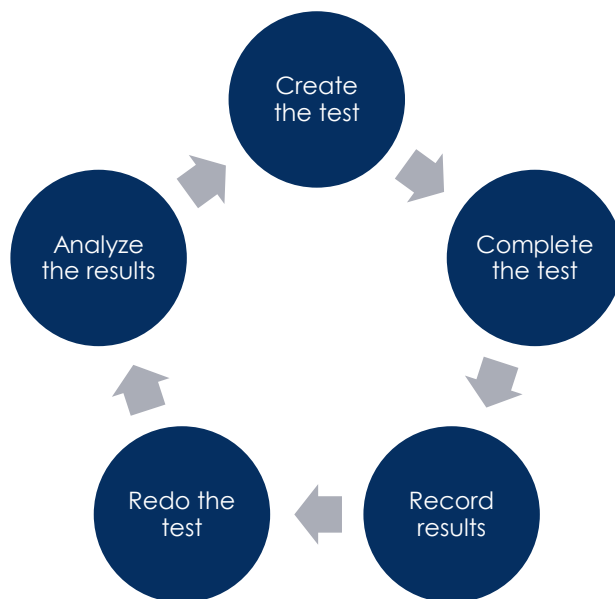
For the experiment, a total of 22 elderly people were selected to test the needs of elderly people in a robotics assisted interface. Eleven of the elderly persons with mild cognitive impairment and a group of eleven cognitively healthy elderly individuals took part in this study. Participants of the study were given tasks to do that involved making an agenda and a shopping list. A structured questionnaire was used to collect socio-demographic data. Then, a test moderator explained the general characteristics of the robot, the evaluation procedure and conducted the tests. Data then were analyzed with execution time and number of errors as measures. (C. Granata, G. Legouverneur, J.-S. Vidal, P. Bidaud, and A.-S. Riguid, 2013)

The results showed that the group with cognitive impairments had more errors than the cognitively healthy group. Results also showed that younger participants and those who were more experienced with technology were faster at completing tasks. This study helps determine what factors must be taken into consideration when developing software. This was accomplished by understanding the scope of the software and testing according to those scopes. By determining the purpose and goals of the software, one can develop tests that are specific to that scope, which enables the best results when developing a user interface.

Testing has now been completed. According to Carol M. Barnum (2019), “Making sense of it all can be overwhelming, but not if you have a plan.” You’re sure that you have done enough testing. What is your next step? Well now, hopefully, you have more than enough data and records. This data should have been gathered from participants, and/or researchers. It is time to look it over and make logic of it. It’s a good rule to look at the pros and the cons. What was the worst outcome of the experiment? What does it mean or why did it occur? There are a million ways to analyze this data. That is something you don’t want to overlook. Find patterns, statistics, or any general relation. This can help to find outliers. Why was this one case different from the others? Can the case be recreated? Analyze is one of the most important parts. What does data mean to you when it is illegible?

First what has happened over the course of this test? To acquire the results wanted, knowledge must be acquired on what can be done. Typically, results are not always expected. Now what does that mean and what should be done about it? We should take this step by step and not jump the gun. This is important. After everyone has looked through the testing and gathered opinions, list what was beneficial and what was not. Collect the top results. This includes the most beneficial, the most detrimental, and the most unexpected. According to Carol M. Barnum (2019), “It is quite powerful to see the top findings come to light in this informal way.” Remember to stay organized. Nothing is worse than losing data or not being able to find the right data at the right time.

Now that the data has been analyzed. What will happen next? What does all this mean? Making sense of the results is critical:



(Naquan Smith, n.d.)

Sometimes you will have to redo the test and reanalyze. This increases accuracy!

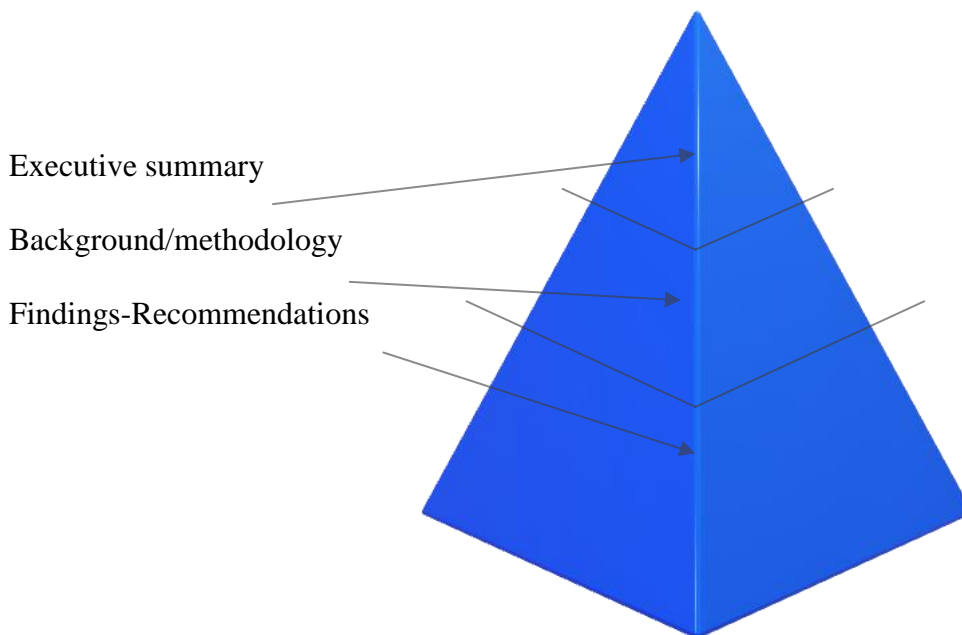
After the analysis is done, the next step is to report the findings. How will the report be displayed? Would a written report be better or an oral report? Will the presentation be formal or informal? Graphs are also great for presentation and overall visuals. Now the pitch must also be kept in mind. In what way will this be delivered?

A report is similar to a paper. First, let's structure the write up. Carol Barnum (2011) stated, "Although you have many media to choose from to present your study results most study results are written up in some sort of report." Who is the audience? Talking to co-workers is very different from a presentation to friends. Why was this test done? The purpose also helps with the focus of the report. Will the report start a bunch of questions or will the report be recorded as supporting information? It is essential to be prepared for any follow up questions.

Lastly, consider the format of the report. There are a few traditional styles. Carol Barnum (2011) wrote, "The organizational structure of a report for a mixed audience begins with an executive summary." This is a pyramid format that starts with the executive summary. Followed by the background/methodology, and concluding with the findings, recommendations, and next steps, the overall format is ultimately the designer's choice. Some may consider an inverted pyramid for informed readers. This starts with the top findings, proceeds to the recommendations, and ends with the background. A report for well-informed readers starts with recommendations, but then goes to top finding before ending with the background information.

In conclusion, no matter what report is created, some formal structure is needed. Carefully planning will lead to a great report. Don't forget that graphs are always a good way to present data in a more visually pleasing way. Data accuracy is also important. No one wants to be presented with false information. This could hurt a company's integrity and reputation. With all of this information in mind, anyone will be well equipped and prepared for their next report.

Pyramid created by Naquan Smith (n.d.):



Readers need more information than a generic summary. Tables that showcase trial and error, user behaviors, and use of features show the findings quickly and effectively. Carol Barnum (2011) wrote, "Illustrating the findings allows you to show as well as tell your readers what the issues are." Charts and graphs, quotes, and video clips also illustrate the idea. Using screen shots really gives a picture to unfamiliar viewers, especially when illustrating something that has changed. The top visual experience is a video clip.

With your observation reports you want to also tie in post results. Survey responses create helpful data. You can present the median, mean, or mode for the total responses. You can also give single responses to show specific examples. Carol Barnum (2011) wrote, “In addition to the qualitative responses you may have used as quotes in your report, you can also collate and visually display the results of qualitative responses you get from other methods, such as the product reaction cards that were used.”

A concise summary is a great start with any audience, whether formal or informal. This helps everyone understand the plans and focus. Many people start and stop at the summary; this is fine if the underlying information is made easily accessible, or as Carol Barnum (2011) put it, “that’s OK because you have served their needs and respected their time.”

After the intro is done, then the details start. There are multiple ways to show your results. Some examples:

- Using graphs
- Showcasing with screen shots, quotes, and clips
- Sorting information
- Using surveys
- Reaction cards
- Recommendations

Now the dilemma associated with presenting. There are countless people who can vouch and speak about anxiety associated with presenting. Breaking the stages down helps:

1. Researching your topic. Knowledge about your subject is the most important aspect.
2. Preparation for the presentation is important. Details and resources that will be used in the presentation may be backed by images, clips, or any other supporting information.
3. Assembling the visuals. From slide shows to pamphlets there are countless options.
4. Rehearsing the presentation. Going over components of the presentation will make the presentation more appealing and well done.
5. Presenting the presentation.

In the end, remember the presenter is in control of what the audience perceives.

10.4 – Different Methods of Usability Testing

When testing programs, there are different approaches that can be taken. Two commonly known methods are the conventional method and the remote method. In the remote method, the subjects and the evaluator are not in the same location and may not even be conducting the experiment at the same time. Two subcategories under the remote method are the forum and diary method. Each has advantages and disadvantages.

In order to understand this experiment, one must understand how each of the two methods work. The forum method gives users several tasks; at the end of the tasks, the user makes notes about the problems that they had while completing the tasks. The user reports the problems in the forum and discusses those problems with other users (S B Yudhoatmojo and S F Sutendi, 2017). But before the user posts the problem, they must

first check if any other user experienced the same problem. If there a post with the same problem experienced, the user must instead respond to post. The diary method, on the other hand, is performed only by the user and across several days. On the first day, the user is given several tasks; if the user finished them, then the user must report the problems encountered and the duration for completing each task (S B Yudhoatmojo and S F Sutendi, 2017). On the days that follow, additional tasks will be given to the user, and the same process is followed.

To test the efficiency of the two methods, an in-house application was used. It is a student peer-reviewed application which is designed for assessing group work contributions (S B Yudhoatmojo and S F Sutendi, 2017). With the forum method, the study lasted one day. For the diary method, the users wrote notes on Google Docs over the course of five days. Two groups of people, divided by students and lecturers, used the forum method approach or the diary method approach to complete their given tasks.

After evaluating the results, it was shown that the number of usability problems decreased in the diary method. The results show that the diary method is faster than the forum method in terms of the duration of working on the test scenario. However, the forum method demonstrated that finding problems with the usability becomes easier, as a user can see similar problem or problems others encountered. In conclusion, the results show how both methods have advantages and disadvantages from the time it takes to perform the test to the benefits of how the user interacts with an application and becomes more familiar with it (S B Yudhoatmojo and S F Sutendi, 2017).

10.5 - Environment in Usability Testing

When testing the usability for a user interface, it is important to take into account the environment. When deciding to select a method of experiment for usability review, one must take in account the group of people needed for evaluation. In most cases, the testers should be the people using that software on a day-to-day basis. In the article, Evaluation of User-Interface Alerts Displays for Clinical Decision Support Systems for Sepsis, a group of nurses evaluate the display of novel user interfaces for sepsis alerts of CDS systems specific to staff nurses (Devida Long, Muge Capan, Susan Masciloli, Danielle Weldon, Ryan Arnold, and Kristen Miller, 2018). This report illustrates why environment is important when testing usability in a user interface.

Sepsis is a life-threatening disease state characterized by organ dysfunction due to a dysregulated host response to infection and affects more than 1 million persons in the United States every year. Hospitals turn to a clinical decision support system to aid specific patients. However, users sometimes find that they lack the knowledge on how to differentiate alerts given by the system. (Devida Long, Muge Capan, Susan Masciloli, Danielle Weldon, Ryan Arnold, and Kristen Miller, 2018)

As a result of these problems, a survey was developed to further evaluate the problem and to understand the required updates needed to accommodate the staff's needs. The survey contained questions split into five domains: general perception of CDS systems; risk parameters (prioritizing information provided in an alert); alert word taxonomy (which signal words best indicated the highest level of alert severity); vital signs, context, and recommendations (which vital signs are the most valuable elements of an alert for predicting severe sepsis); and evaluation of novel interface designs (information design display and location of information). The data from the participants given from the survey allowed for the evaluation of the problems participants had with the interface. Results showed that alerts based on an established treatment protocol were presented in a pop-up format, and addressed the patient's clinical condition rather than

following regulatory guidelines (Devida Long, Muge Capan, Susan Masciloli, Danielle Weldon, Ryan Arnold, and Kristen Miller, 2018).

In conclusion, when testing for usability of an interface, it is crucial that the focus is on the users using the software. The environment of the software guides the development of the testing of usability. Accommodating specific users of the interface, rather than a random group of participants is key. In the study of the nurses, it was seen that the users had to come from a selected group of participants that fit the specific environment of a sepsis system. This is one of many effective ways of addressing bugs/issues of a user interface.

10.6 - Extended Resources

Descriptions & Links

1. Usability Inspection Methods

[http://www.idemployee.id.tue.nl/g.w.m.rauterberg/lecturenotes/0h420/nielsen\[1994\].pdf](http://www.idemployee.id.tue.nl/g.w.m.rauterberg/lecturenotes/0h420/nielsen[1994].pdf)

2. Usability Engineering Methods for Software Developers

https://www.researchgate.net/profile/Andreas_Holzinger/publication/220422205_Usability_Engineering_Methods_For_Software_Developers/links/5460e30c0cf27487b4526442/Usability-Engineering-Methods-For-Software-Developers.pdf

3. Usability Testing Example

<https://www.youtube.com/watch?v=thNZIZmMDQo>

4. Example of Usability Test with a Paper Prototype

<https://www.youtube.com/watch?v=9wQkLthHKA>

5. Usability Testing – Software Testing Tutorial

<https://www.youtube.com/watch?v=SkXDvMYQqGc>

10.7 - References

- Barnum, C. M. (2011, July 05). *Usability Testing Essentials: Ready, Set....Test!* Retrieved June 19, 2019, from <https://ieeexplore.ieee.org/abstract/document/5940190/citations#citations>
- Nielson, J. (2012, January 04). *Usability 101: Introduction to Usability*. Retrieved June 23, 2019, from <https://www.nngroup.com/articles/usability-101-introduction-to-usability>
- Mills, C., & Dye, K. (1985). *Usability Testing: User Reviews*. *Technical Communication*, 32(4), 40-44. Retrieved from <http://www.jstor.org/stable/43094561>
- Hsu, F. C. (2018). *User Satisfaction and System Environment Convenience for Interface Design of Mobile Games*. Retrieved June 23, 2019, from <http://ekolojidergisi.com/download/user-satisfaction-and-system-environment-convenience-for-interface-design-of-mobile-games-5426.pdf>
- Dumas, J. S., & Redish, J. C. (1999). *A Practical Guide To Usability Testing*. Retrieved June 24, 2019, from https://books.google.com/books?hl=en&lr=&id=4lge5k_F9EwC&oi=fnd&pg=PR9&dq=purpose of usability testing&ots=vri67He6tF&sig=ShmRB76jH6rcMAAtJsee4eTJRA
- Yudhoatmojo S. B., & Sutendi S. F. (2017). *Empirical Study on the Use of Two Remote Asynchronous Usability Testing Methods*. Retrieved June 24, 2019, from <https://iopscience.iop.org/article/10.1088/1742-6596/1193/1/012017/meta>
- Long, D., Capan, M., Masciloli, S., Weldon, D., Arnold, R., & Miller, K. (2018, August). *Evaluation of User-Interface Alerts Displays for Clinical Decision Support Systems for Sepsis*. *CriticalCareNurse*. Retrieved July 8, 2019, from <http://ccn.aacnjournals.org/content/38/4/46.short>
- Granata, C., Legouverneur, G., Vidal, J., Bidaud, P., & Riguard, A. (2013). *Robot services for elderly with cognitive impairment: Testing usability of graphical user interfaces*. *Technology and Health Care*. 21, 217-231. Retrieved on July 22, 2019, from <https://content.iospress.com/articles/technology-and-health-care/thc00718>

