Usability Evaluation Methods of Smartphone Interfaces for Older Users

Ingvild Engelsen
Department of Design
NTNU, Norwegian University of Science and Technology

ABSTRACT

The global older population is growing rapidly. Aging involves varied degrees of physical and psychological changes, resulting in older adults being a diverse and complex user group. Different needs and limitations require different design implications, making it extra important to truly understand this user group in order to design 'elderly friendly' interfaces.

By means of literature search, this paper aspires to increase understanding of how aging influence usability and how usability can be evaluated and tested with regards to older users. This paper highlights age-related impairments and accompanying usability problems encountered by older adults when using smartphones. Additionally, three usability evaluation methods are presented and discussed; Heuristic Evaluation, Cognitive Walkthrough and Think Aloud. It is argued that empathy is essential in order to design usable interfaces for older users. In addition, it is shown that future research is necessary to adapt usability testing with older users without solely relying on cognitive capacities. The paper concludes that no single method is superior of others, and a combination of expert- and user-based methods throughout the design process provides best usability results.

KEYWORDS: older users, usability evaluation, heuristic evaluation, cognitive walkthrough, think aloud

1. INTRODUCTION

As people grow older their abilities change. This process of change includes a decline over time in the cognitive, physical and sensory functions, and each of these will decline at different rates relative to one another. This pattern of capabilities varies widely between individuals, and as people grow older, this variability increases. In addition, any given individual's capabilities vary in the short term. This can be due to, for example, temporary decrease in, or loss of function due to a variety of causes including illness, blood sugar levels and state of arousal (Gregor, Newell & Zajicek, 2002).

In 2017, there are an estimated 962 million people aged 60 or over in the world, comprising 13% of the global population. The population aged 60 or above is growing at a rate of about 3% per year. Currently, Europe has the greatest percentage of population aged 60 or over (25%). Rapid aging will occur in other parts of the world as well, so that by 2050 all regions of the world except Africa will have nearly a quarter or more of their populations at ages 60 and above (Paez & Del Río, 2019)

The smartphone is a valuable technological device for the elderly, aged 60 years and above, as it enables them to communicate fully with their family and friends, making them feel less isolated, more secure and independent. Despite the necessity for smartphones to be designed to better accommodate elderly users, the user interface (UI) on current available smartphones do not appear to be optimized for the elderly. Failure to design ‘elderly friendly’ interfaces may lead to reluctance to the use of smartphones by the elderly, while a properly designed UI that respects the elderly’s needs can tackle this issue (Salman, Ahmad & Sulaiman, 2018).

In order to design smartphone interfaces suited for elderly users, it is essential to know the qualities, limitations, needs and struggles of the user group. Thus, it is important to evaluate the interfaces with the proper criteria and test the usability on actual representatives of the user group. In this paper, usability and usability evaluation methods with regards to older users are presented and discussed.

2. DESIGNING FOR OLDER USERS

2.1 Older user group

‘Older adults’ encompass an incredibly diverse group of users, and, even small subsets of this group, tend to have a greater diversity of functionality than is found in groups of younger people (Gregor et al., 2002). This is in part because of the increased likelihood of illness or age-related impairment, in part also because of the wide age-range that the term ‘older’ conventionally covers, e.g. anyone over 60 or 65, and in some cases, anyone over 50. While aging is associated with an increased risk for factors such as tremors, memory difficulties and mobility problems, this simply widens the range of functionality and makes the group more diverse (Dickinson, Arnott & Prior, 2007).

2.2 Physical and psychological changes

Some of the degenerative effects of aging include diminished vision, varying degrees of hearing loss, psychomotor impairments, as well as reduced attention, memory and learning abilities (Zaphiris, Ghiawadwala & Mughal, 2005).

Vision is the most common physiological change associated with aging. After the age of 55 years, many seniors will experience vision changes, including presbyopia, a loss in near vision, reduced field of vision. Older adults will also experience a decline in contrast sensitivity as well as reduced color sensitivity, particularly in the blue and green range (Zaphiris et al., 2005).

Hearing, overall declines with age; and research has shown that 20% of people between 45-54 years have a form of hearing impairment, which rises to 75% for persons between 75-79 years of age. Psychomotor abilities also decline with age. In older adults, response times increase with more complex motor tasks (Zaphiris et al., 2005).

With age, it is generally accepted that there is a decline in the ability to process items from working memory into short term memory. Studies have also found that there is a decline in episodic memory (memory of specific events) and procedural memory (memory for how we carry out tasks) (Zaphiris et al., 2005).

2.3 Older users and smartphones

Older adults selectively make use of new technologies, but often encounter more trouble than younger adults in acquiring devices, becoming competent users and troubleshooting. The older a person is, the more likely they are to seek help with operating a new technology (Calak, 2013).

The most common usability problems older adults encounter when interacting with a smartphone application are mainly related to visual, psychomotor and cognitive limitations. The size and sharpness of visual elements, including hard colors and inappropriate contrast colors represent significant problems when vision is reduced. Regarding psychomotor limitations, fast and repetitive movements for interaction, a large number of steps required to complete a task and use of scrollbars represent inconveniences. Irrelevant graphics, animations or decoration with no explaining text tends to confuse older adults with reduced cognitive abilities. In addition, use of complex text and navigation through deep and expandable menu hierarchies are confusing for older adults (Paez & Del Río, 2019).

2.4 User Sensitive Inclusive Design

It is important to remember that there is a wide diversity of abilities among traditional users, and that all users' abilities are dynamic over time. User Sensitive Inclusive Design encourages designers to seek out diversity, and is appropriate for the process of design where older and/or disabled people are part of the potential user group (Gregor et al., 2002).

User Sensitive Inclusive Design focuses on a close relationship with the intended users, and thus follows the principles of Participatory Design, rather than traditional User-Centred Design. 'Sensitive Design' also encourages the designer to consider the whole person, not simply their physical characteristics (Newell, Gregor, Morgan, Pullin & Macaulay, 2011).

The approach underlines the extra levels of difficulty involved when the range of functionality and characteristics of the user groups can be so great that it is impossible in any meaningful way to produce a useful specification of the user group and very difficult to achieve a small representative sample of the user group (Newell et al., 2011).

3. USABILITY EVALUATION FOR OLDER USERS

3.1 Usability

Usability is by the ISO 9241-11 standard 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' (ISO, 2019). Usability is therefore highly linked to the intended user of the system/interface/product and the context in which it will be used. To design a usable system it is important to understand the intended user, their level of expertise, the time expected to be used on a system, and that the users' needs will change as they gain experience (Goodwin, 1987, s. 232).

3.2 Usability evaluation

Interactive systems are usually designed through an iterative process of initial design (prototype), evaluation, and redesign (Jaspers, 2009). Consequently, usability evaluation is a fundamental step this design process . The goal is to assess the degree to which the system is effective, efficient and satisfactory for the intended user (Bastien, 2010).

Usability evaluation techniques can be classified into two groups; expert-based and user-based. In expert-based methods; also known as inspection methods, experts inspect the UI and predict problems users would have while interacting with the interface. The outcome of such studies can be a formal report that highlights problems identified or recommendations for changes. User-based methods; also known as testing methods, find and identify usability problems through observing people who are representing users while using/commenting on the system interface. The aim of user-based tests is to assess the degree to which the system supports the intended users in their workflow (Salman et al., 2018).

This paper will further examine the three most used usability evaluation methods; Heuristic Evaluation, Cognitive Walkthrough and the Think Aloud method (Jaspers, 2009).

3.3 Heuristic Evaluation

Heuristic Evaluation is one of the most widely used expert-based evaluation methods (Salman et al., 2018). The interface is evaluated against a list of recognized usability principles, *the* *heuristics*, by a group of evaluators. Each evaluator goes through the interface in depth, conducts a list of flaws and errors corresponding to each heuristic, and estimates the severity of each problem. The results are compared and summarized in a report that describes the usability flaws of the system in relevance to the heuristics (Jaspers, 2009).

Jakob Nielsen's (Nielsen, 1994) ten general principles of interaction design are widely used heuristics for user interface design (Watkins, Kules, Yuan, & Xie, 2014);

1. Visibility of system status
2. Match between system and the real world
3. User control and freedom
4. Consistency and standards
5. Error prevention
6. Recognition rather than recall
7. Flexibility and efficiency of use
8. Aesthetic and minimalist design
9. Help users recognize, diagnose, and recover from errors
10. Help and documentation

3.3.1 Heuristic Evaluation and older users

Heuristics guidelines specific for older adults have been developed for web (Chisnell & Redish, 2005; Zaphiris et al., 2005) iPad (Watkins et al., 2014) and smartphone (Calak, 2013; Kneale, Mikles, Choi, Thompson & Demiris, 2017; Silva, Holden & Jordan, 2015).

In the domain of mobile health applications (mHealth), an aging barriers framework, MOLD-US has been developed. The MOLD-US framework provides a visual overview of the possible impact of each aging barrier category, related diminishing age dependent abilities and medical conditions to the usability of mHealth in older adults (Wildenbos, Peute & Jaspers, 2018). The framework helps with classification of the usability problems encountered, but does not provide recommendations on (re)designing apps for older users based on the usability issues (Wildenbos, Jaspers, Schijven & Dusseljee-Peute, 2019).

Another study used SMASH, a set of 12 smartphone usability heuristics, to evaluate a smartphone user interface for older adults. SMASH is built on an earlier version of validated usability heuristics for Touchscreen-based Mobile Devices (TMD), and is experimentally validated (Inostroza, Rusu, Roncagliolo, Rusu & Collazos, 2016). The study was complemented by testing with eight elderly participants to check to what extent the usability problems identified by the experts would be perceived as real problems by the elderly. 79,17% of the usability problems encountered by elderly participants were identified through Heuristic Evaluation using SMASH (Salman et al., 2018).

To include the diversity of the older population, Chisnell and Redish (2005) suggest a new methodology to Heuristic Evaluation. Creating personas, not only based on age, but ability, amplitude and attitude, and actively using them when evaluating the interfaces, makes the evaluation more user-oriented (Chisnell & Redish, 2005). The evaluator first records observations using the interface as the persona, before fitting these observations into the heuristic framework. Kneale et al. (2017) further developed this methodology by introducing scenarios as a companion for the personas to guide reviewers' assessment.

3.4 Cognitive Walkthrough

The Cognitive Walkthrough is an expert-based evaluation method. The method is usually used in the early stage of the design process and focuses on evaluating the learnability of the system by exploration. Contrary to the Heuristic Evaluation, the Cognitive Walkthrough is explicitly guided by the user's tasks and in that way highly structured. A usability expert analyses the cognitive processes that is required to accomplish typical tasks supported by the interface. For each task, the evaluator analyses the four steps needed for the user to accomplish the task (Jaspers, 2009).

1) Set a goal to be accomplished
2) Inspect the available actions on the user screen
3) Select one of the available actions
4) Perform the action and evaluate the system's

When analyzing these four steps, the evaluator tries to answer four questions (Jaspers, 2009):

1) Will the user try to achieve the correct effect
2) Will the user notice that the correct action is available
3) Will the user associate the correct action with the desired effect
4) Will the user notice that progress is being made toward accomplishment of his goal

The specific task is found to be without usability problems if all four questions receives a positive answer (Jaspers, 2009).

3.4.1 Cognitive Walkthrough and older users

The Empathic Walkthrough (EW) is a method derived from Cognitive Walkthrough, adapted to maximize empathy with the end user and context. Through focus on user stories and personas, the designer is forced to imagine the use of a conceived product or system within a different context and set of user concerns. This empathy-driven retelling of the concept promotes the identification of obstacles or critical concerns (Gray, Yilmaz, Daly, Seifert & Gonzalez, 2015).
The Barrier Walkthrough (BW) method is an analytical technique based on a combination of Heuristic Evaluation and Cognitive Walkthrough (Lunn, Yesilada & Harper, 2009); the Heuristic Walkthrough (Sears, 1997). In the BW method, an evaluator considers a number of predefined possible barriers which are interpretations and extensions of well known accessibility principles. Scenarios composed of types of users, settings, goals and possible tasks are identified, and the degree of severity of the barrier are defined. Barriers encountered by older users of the Web have been specified, and can be used in conjunction with the BW method to assess the impact a Web page has on older users and their ability to interact with the content effectively (Lunn et al., 2009).

3.5 The Think Aloud method

The Think Aloud method is a user-based testing method, and one of the main techniques used by usability professionals when conducting usability studies (Olmsted-Hawala & Bergstrom, 2012). The aim is to assemble insight and information from human's cognitive behavior while performing tasks. The Think Aloud method provides valuable information on cognitive processes and how humans solve problems (Jaspers, 2009).

The two most common Think Aloud protocols are Concurrent Think Aloud (CTA) and Retrospective Think Aloud (RTA). In CTA, the participant is encouraged to 'think out loud' while working on a task, while in RTA the participant talks only *after* the session is completed, typically while watching a video replay of the session. According to Olmsted-Hawala et al. (2012), CTA is optimal when the goal is to obtain a real sense of the users' experience with the interface, while RTA is a better choice when the user's insight of the issues with the interface is of interest (Olmsted-Hawala et al., 2012).

3.5.1 The Think Aloud method and older users

Older participant diversity makes it important to carefully control the experimental conditions and measures. The quality of the data obtained is affected by processing capacity, education, physical impairments and memory and this can reduce the technique's efficacy with older participants (Dickinson et al., 2007).

Research studies have demonstrated limitations of Concurrent Think Aloud with older research participants. The process of thinking aloud can interfere with the completion of the experimental task, particularly for participants with cognitive impairments, as they struggle with unfamiliar interfaces (Dickinson et al., 2007).

A study using Retrospective Think Aloud to obtain insight on older adults' email usage, learned that processing and memory difficulties meant that this approach was not effective. Older participants often found it very difficult to recall what they had done in order to complete a task and they rarely remembered processes accurately until they had repeated them several times. In addition, confusion among older beginners is often general, poorly articulated and non-specific (Dickinson et al., 2007), making it difficult to retrieve valuable information.

Wildenbos et al. (2019) used the MOLD-US framework to classify identified usability problems found in two mHealth apps for older adults using the Think Aloud usability evaluation. According to Wildenbos et al. (2019), MOLD-US provided understanding of the prominent barriers hampering older adults' usage of the apps and allowed for improved analysis of the results.

4. DISCUSSION

4.1 The importance of designing for diversity

Older adults form a diverse user group due to a great diversity of functionality. However, the diversity of abilities among traditional users are also wide, and all users' abilities are dynamic over time (Gregor et al., 2002). Physical and psychological limitations associated with aging, like reduced vision, hearing, psychomotor or cognitive abilities can also apply to other user groups. User Sensitive Inclusive Design can therefore be considered relevant for most user groups, as it aims to seek out diversity and look holistically at the users.

It is rarely possible to design a product that is truly accessible to all potential users. 'Inclusive' is therefore more achievable than 'universal' or 'for all'. Although User Sensitive Inclusive Design is said to be appropriate when designing for user groups where older and/or disabled people are part of the user group (Newell et al., 2011), it can also be applied advantageously when designing for other user groups.

4.2 When to use which method?

Depending on time, budget, resources, user group, desired outcome and where in the system development life cycle (SDLC), different usability evaluation methods have their advantages and limitations.

Heuristic Evaluation is an effective usability evaluation method with high cost-benefit ratio. When time and resources are limited, skilled experts can give results of high quality without needing to involve the users (Jaspers, 2009). The Heuristic Evaluation identifies a larger number of usability problems at a lower cost, compared to the Cognitive Walkthrough and Think Aloud method (Jeffries, Miller, Wharton & Uyeda, 1991). However, the Heuristic Evaluation is an unstructured approach, only guided by the heuristics. As heuristics are generally defined and can be interpreted differently by the evaluators, the overlap of usability problems found by different evaluators tend to be low. Accordingly, the value of the Heuristic Evaluation method is highly dependent on the skill-level of the evaluator (Jaspers, 2009).

If time or budget is limited, and therefore the focus is on identifying the more severe usability problems, the Cognitive Walkthrough might be a better choice. The Cognitive Walkthrough detects around one third of the usability flaws identified by a Heuristic Evaluation, but again proportionally more severe than specific and low priority problems (Jeffries, 1991). Compared to the Heuristic Evaluation, the Cognitive Walkthrough is a more structured approach with a stronger focus on learnability. Accordingly, high levels of detailed task and user background is required to obtain valuable results. Whereas the Heuristic Evaluation do not appear to provide enough structure, the Cognitive Walkthrough might provide too much (Jaspers, 2009).

The combination of Heuristic Evaluation and Cognitive Walkthrough, the Heuristic Walkthrough, provides more structure than Heuristic Evaluation but less than Cognitive Walkthrough, resulting in an effective task-oriented evaluation technique. A study indicates that Heuristic Walkthrough are more thorough than Cognitive Walkthrough and more valid than Heuristic Evaluation. In other words, Heuristic Walkthroughs resulted in finding more problems than Cognitive Walkthroughs and fewer false positives than Heuristic Evaluations (Sears, 1997).

Equivalent to the Cognitive Walkthrough, the Think Aloud method has proven to identify about one third of the usability problems detected by a Heuristic Evaluation (Jeffries, 1991). However, the Think Aloud identifies significantly more problems of a severe and recurring degree than the Cognitive Walkthrough. Accordingly, the Think Aloud is superior in detecting the underlying causes of the usability problems (Jaspers, 2009), but at a higher cost than the Heuristic Evaluation and the Cognitive Walkthrough. Similar to the Heuristic Evaluation, the Think Aloud method is dependent on the skills of the evaluator. The results are also highly affected by the representativeness of the user group and the task selection (Jaspers, 2009). Considering the diversity of older adults as a user group, this can be challenging.

Usability evaluation at each phase of the SDLC are critical to ensure that the product will actually be usable (Fernandez, Insfran & Abrahão, 2011). Expert-based usability evaluation methods are often used to test prototypes early in the SDLC, due to the ability to detect usability flaws at a low cost (Jaspers, 2009). Detecting usability problems early results in improved quality and reduced use of resources (Fernandez et al., 2011). However, expert-based methods should not hold as a substitute for user testing. Experts seldom have enough knowledge and understanding of the user to detect all usability problems (Jaspers, 2009). Especially when the user group consists of a great diversity of functionality and needs, like older adults, the importance of testing on real users can not be emphasized enough. No single usability evaluation method addresses all the existing usability problems (Fernandez et al., 2011). When used in combination throughout the SDLC, expert-based and user-based methods can complement and support each other, allowing for superior and more user-friendly results.

4.3 Does empathy facilitate usability?

Usability is deeply linked to the intended user and the context of use. It is therefore vital to understand the user and their difficulties. Consequently, usability evaluation methods need to be adapted to the user in order to obtain relevant and important information.

Complex and diverse user groups, like older adults, require additional focus on understanding the intended users. The wide use of personas and scenarios in the expert-based evaluation methods used for older adults might be an attempt to obtain this valuable insight.

Creating authentic personas and scenarios can support the evaluator to empathize with the users while conducting the evaluations. Personas and user scenarios are essential parts of both the Empathic Walkthrough (Gray et al., 2015) and The Barrier Walkthrough (Lunn et al., 2009). As shown earlier, the Heuristic Evaluation can also benefit from integrating personas (Chisnell & Redish, 2005) and scenarios (Kneale et al., 2017) when evaluating interfaces for older users. With this task-oriented focus, the method resemble more a Heuristic Walkthrough. Regardless of method definitions, the ambition of including personas and user scenarios is the same; to truly understand the users and their needs.

4.4 Are domain specific heuristics necessary?

While general heuristics are suitable to evaluate most user interfaces, there is still a need to establish heuristics for specific domains to ensure that the specific usability issues are identified. Due to lack of standard reporting formats, guideline organization and validation procedures (Nurgalieva, Laconich, Baez, Casati & Marchese, 2017), it is difficult to compare different studies. According to Hermawati and Lawson (2016), it is inconclusive whether domain specific heuristics perform better or worse than general ones. This is due to lack of validation quality and clarity on how to assess the effectiveness of heuristics for specific domains. Out of 70 studies, only 19 studies could be used to provide indication on the effectiveness of domain specific heuristics. Furthermore, over 80% of the studies used similar heuristics as Nielsen's (Hermawati & Lawson, 2016), making the divergence between the heuristics limited.

On the other hand, Salman et al. (2018) showed that it might not be necessary to have specific heuristics for older adults. The MOLD-US framework can help the evaluator recognize and organize the usability problems encountered. Although the framework is developed for mobile health applications, it can be modified to support other domains.

4.5 Is the Think Aloud method suitable for older adults?

The Think Aloud method relies heavily on the cognitive capacities of participants, such as communication, attention and speed of comprehension, exactly the cognitive capacities that decline with aging. These cognitive skills, especially attention, are deeply solicited by the Think Aloud method, hindering people with cognitive limitations in retaining sufficient attention for using the app under evaluation (Wildenbos et al., 2019).

Additionally, the combination of inexperience with technology and experimental techniques can place considerable stress on processing capacity. Age-related reduction in processing capacity can exacerbate this stress. These factors can reduce the efficacy of participant self-reporting (Dickinson et al., 2007). Moreover, usability issues provoked by physical ability barriers are not easily identified by the Think Aloud method (Wildenbos et al., 2019).

It is important to again emphasize the diversity of older participants. The concern is that data are not elicited from the less articulate or from those who find the combination of tasks difficult in terms of processing capacity (Dickinson et al., 2007).

An implication of this is the possibility that usability evaluation approaches may need adjustments to prevent reporter bias and become better suited for testing with the older adult population. To tackle these issues, research into suitable usability evaluation methods adapted to older users' characteristics is needed. Such methods could ultimately provide unbiased sight on usability issues older users may experience while interacting with technology (Wildenbos et al., 2019). Possible modification aspects might be how feedback is acquired from the older users and in which environments and settings.

5. CONCLUSIONS

The purpose of this paper has been to understand how aging influences usability and how usability can be evaluated and tested with regards to older users. This paper argues for the importance of designing for diversity and looking holistically at the user. The diversity in functionality among older users require this mentality to improve usability for older adults. In this regard, User Sensitive Inclusive Design is presented as a suitable approach. Further, this paper argues that empathy is essential in order to design usable interfaces for older users. Including authentic personas and user scenarios when evaluating the usability of an interface, increase the evaluators' understanding of the needs and limitations of older adults, resulting in more accurate and reliable findings.

A variety of factors influence which usability evaluation methods to decide on, and all methods have advantages and limitations. This paper concludes that no single method is superior of others, and a combination of expert- and user-based methods throughout the design process provides best usability results. It is inconclusive whether domain specific heuristics are necessary or not. Regardless, good results using the MOLD-US framework suggests that a classification framework might be valuable when evaluating usability for older users. The demand for developing a classification framework applicable to other domains are present.

Usability testing of interfaces with older adults is a field that needs further research and exploration, including the Think Aloud method. Considering that the Think Aloud method relies heavily on cognitive capacities, such as attention and memory, it is evident that it is not suitable for older users. Further research on how to obtain insight from older users, without relying on cognitive capacities, could avoid inaccurate assumptions and further improve smartphone usability for older users.

REFERENCES

Bastien, J. C. (2010). Usability testing: a review of some methodological and technical aspects of the method. *International journal of medical informatics*, *79*(4), e18-e23.

Calak, P. (2013). Smartphone evaluation heuristics for older adults (Doctoral dissertation).

Chisnell, D., & Redish, J. (2005). *Designing web sites for older adults: Expert review of usability for older adults at 50 web sites* (Vol. 1, pp. 1-60). San Francisco: AARP.

Dickinson, A., Arnott, J., & Prior, S. (2007). Methods for human–computer interaction research with older people. *Behaviour & Information Technology*, *26*(4), 343-352.

Fernandez, A., Insfran, E., & Abrahão, S. (2011). Usability evaluation methods for the web: A systematic mapping study. *Information and software Technology*, *53*(8), 789-817.

Goodwin, N. C. (1987). Functionality and usability. *Communications of the ACM*, *30*(3), 229-234.

Gray, C. M., Yilmaz, S., Daly, S. R., Seifert, C. M., & Gonzalez, R. (2015). Idea generation through empathy: Reimagining the 'cognitive walkthrough'.

Gregor, P., Newell, A. F., & Zajicek, M. (2002, July). Designing for dynamic diversity: interfaces for older people. In *Proceedings of the fifth international ACM conference on Assistive technologies* (pp. 151-156). ACM.

Hermawati, S., & Lawson, G. (2016). Establishing usability heuristics for heuristics evaluation in a specific domain: Is there a consensus?. *Applied ergonomics*, *56*, 34-51.

Inostroza, R., Rusu, C., Roncagliolo, S., Rusu, V., & Collazos, C. A. (2016). Developing SMASH: A set of smartphone's usability heuristics. *Computer standards & interfaces*, *43*, 40-52.

International Organization for Standardization. (2019). *Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems* (ISO Standard No. 9241-210). Retrieved from https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-2:v1:en

Jaspers, M. W. (2009). A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. *International journal of medical informatics*, *78*(5), 340-353.

Jeffries, R., Miller, J. R., Wharton, C., & Uyeda, K. (1991, April). User interface evaluation in the real world: a comparison of four techniques. In *CHI* (Vol. 91, pp. 119-124).

Kneale, L., Mikles, S., Choi, Y. K., Thompson, H., & Demiris, G. (2017). Using scenarios and personas to enhance the effectiveness of heuristic usability evaluations for older adults and their care team. *Journal of biomedical informatics*, *73*, 43-50.

Lunn, D., Yesilada, Y., & Harper, S. (2009). Barriers Faced by Older Users On Static Web Pages Criteria Used In The Barrier Walkthrough Method.

Newell, A. F., Gregor, P., Morgan, M., Pullin, G., & Macaulay, C. (2011). User-sensitive inclusive design. *Universal Access in the Information Society*, *10*(3), 235-243.

Nielsen, J. (1994, 24. April). 10 Usability Heuristics for User Interface Design. Retrieved from https://www.nngroup.com/articles/ten-usability-heuristics/

Nurgalieva, L., Laconich, J. J. J., Baez, M., Casati, F., & Marchese, M. (2017). Designing for older adults: review of touchscreen design guidelines. *arXiv preprint arXiv: 1703.06317*.

Olmsted-Hawala, E., & Bergstrom, J. R. (2012). Think-aloud protocols: does age make a difference. *Proceedings of Society for Technical Communication (STC) Summit, Chicago, IL*.

Paez, L. E., & Del Río, C. Z. (2019, July). Elderly Users and Their Main Challenges Usability with Mobile Applications: A Systematic Review. In *International Conference on Human-Computer Interaction* (pp. 423-438). Springer, Cham.

Salman, H. M., Ahmad, W. F. W., & Sulaiman, S. (2018). Usability evaluation of the smartphone user interface in supporting elderly users from experts' perspective. *IEEE Access*, *6*, 22578-22591.

Sears, A. (1997). Heuristic walkthroughs: Finding the problems without the noise. *International Journal of Human-Computer Interaction*, *9*(3), 213-234.

Silva, P. A., Holden, K., & Jordan, P. (2015, January). Towards a list of heuristics to evaluate smartphone apps targeted at older adults: a study with apps that aim at promoting health and well-being. In *2015 48th Hawaii International Conference on System Sciences* (pp. 3237-3246). IEEE.

Watkins, I., Kules, B., Yuan, X., & Xie, B. (2014). Heuristic evaluation of healthy eating apps for older adults. *Journal of Consumer Health On the Internet*, *18*(2), 105-127.

Wildenbos, G. A., Peute, L., & Jaspers, M. (2018). Aging barriers influencing mobile health usability for older adults: A literature based framework (MOLD-US). *International journal of medical informatics*, *114*, 66-75.

Wildenbos, G. A., Jaspers, M. W., Schijven, M. P., & Dusseljee-Peute, L. W. (2019). Mobile health for older adult patients: Using an aging barriers framework to classify usability problems. *International journal of medical informatics*, *124*, 68-77.

Zaphiris, P., Ghiawadwala, M., & Mughal, S. (2005, April). Age-centered research-based web design guidelines. In *CHI'05 extended abstracts on Human factors in computing systems* (pp. 1897-1900). ACM