# Three Case Studies on Methods of Working with Older People on the Design of New Technologies

Zaidatol Haslinda Abdullah SANI[1], Helen PETRIE, David SWALLOW and Andrew LEWIS
*Human Computer Interaction Research Group, Department of Computer Science, University of York, United Kingdom*

**Abstract.** The proportion of older adults in the population is rapidly increasing and the proportion of younger adults to care for them is decreasing. Part of the solution to support older adults in living independently is to provide them with appropriate assistive technologies. To develop technologies that are effective for older adults we need methodologies that are appropriate for working with this user group. Yet there is little systematic research on how to work with older adults and how to adapt methods already used with younger adults. This paper reports on three case studies which investigated the use focus groups, expert evaluations and user evaluations with older adults. In the case of focus groups, the size of the focus group was investigated; for expert evaluations, an existing set of heuristics for evaluating apps for older adults was investigated; for user evaluations, a low-fidelity prototype design was evaluated using think-aloud protocols.

**Keywords.** research methods, older adults, focus group, expert evaluation, user evaluation, technology, mobile technology

## 1. Introduction

Currently the number of older adults in the world is approximately 901 million [1]. The United Nations (UN) reports that the number of older adults is expected to double to more than 2 billion by 2050 and also predicts that the number of older adults is projected to exceed the number of children aged 15 years or younger for the first time in 2047 [1]. In the United Kingdom (UK), the Office of National Statistics [2] predicts that population of older adults will increase from 10.8 million in 2013 to 17.6 million in 2037. This increase in the older population and change in balance between younger and older people will place many stresses on society. There will be fewer people of working age to care for older people, so older people will need to be more independent. However, this fits with the fact that older people generally prefer to live independently in their own homes for as long as they can [3]. Part of the solution to these stresses is to provide older people with technologies to support them in being independent. While this is a considerable challenge for the current cohorts of older people, who grew up

[1] Corresponding Author, Human Computer Interaction Research Group, Department of Computer Science, University of York, Deramore Lane, Heslington East, York YO10 5GH, United Kingdom; E-Mail: zas508@york.ac.uk.

before the advent of personal computers and smartphones, for coming cohorts of older people, this will seem quite natural.

However, designing technologies that really work for older people brings its own challenges. We can apply user-centred and participatory design methodologies which use a range of specific methods such as interviews, focus groups, and task-based evaluations. All these methods are well-established for working with younger people, but a number of studies [e.g. 4, 5] have found that these methods are not necessarily so suitable for use with older adults.

The aim of this paper is to investigate the suitability of three methods for working with older people when developing technologies to support their independence: focus groups, which are often used to elicit initial requirements for new technologies; collaborative heuristic evaluation (CHE), a variation of heuristic evaluation developed by Petrie and Buykx [6] that can be used by experts to evaluate early prototypes of new technologies; and think-aloud protocols [7] which requires users to perform tasks and verbally express their thoughts about a technology. The three case studies were all used in the context of developing a suite of smartphone and tablet computer apps for older people to support them in maintaining good nutrition and intake of liquids, both significant problems for older people [8].

## 2. Case Study 1: Focus Groups

### *2.1. Introduction*

Focus groups are a very common method to discuss requirements and ideas for new technologies with users [9 - 11]. For younger participants, the recommended number of participants in a focus group ranges from 3 to 12 participants. However, there are mixed views in the literature about the number of participants for focus groups with older adults [e.g. 5, 12 - 15]. Inglis et al. [14] reported that focus groups larger than three older adults are hard to manage. They reported that the characteristics of the older adults, for example poor hearing, and thus difficulty in following the discussion, influenced the session negatively. Hawthorne et al. [13] had four to six participants aged more than 80 years per focus group. They reported the participants had various age-related health problems and required constant care during the discussion.

Lines and Hone [5] reported that having 12 older adults (aged more than 65 years) per focus group session contributed to having the participants to discuss unrelated topics and side discussions among themselves. This led to difficulties for the moderator in managing the sessions. Similar findings were reported by Lyons et al. [15] with groups of two to ten participants (aged more than 65 years with a mean age of 75.2 years). Brondani et al. [12] found that having five to nine participants (aged between 64 and 93 years) per focus group contributed to having participants who made few attempts to dominate the discussion and few serious disagreements. The same issues were also reported by Lyons et al. [15].

Thus one of the aims of Case Study 1 was to explore the effects of the number of older adult participants in focus groups. We also provide reflections on how best to conduct focus groups with older adults.

### *2.2. Method*

#### *2.2.1. Design*

Two sizes of focus group were conducted: small focus groups with three to four older adults, and larger focus groups with six to seven older adults. Two of the focus groups (one small, one larger) were about the possible use of technology for older adults in relation to nutrition, hydration, and exercise. The other two focus groups (one small, one larger) were about the possible use of technology for older adults in relation to mobility and wellbeing [16, 17]. Both focus groups for each topic were facilitated by the same person, but a different person facilitated each topic.

#### *2.2.2. Participants*

There were nine participants for the focus groups. The inclusion criteria was to be 65 years or over and living independently, either alone or with a partner. The demographics of the participants, in terms of gender compositions, age, living arrangements, educational level and current employment status are summarized in Table 1.

**Table 1.** Compositions of the focus groups

| **Group** | **Composition** |
|---|---|
| Co-Motion (Small) | 4 participants: 3 women, 1 man<br>57 – 86 years |
| Co-Motion (Large) | 7 participants: 3 women, 4 man<br>56 – 82 years |
| Nutrition (Small) | 3 participants: 2 women, 1 man<br>65 – 70 years |
| Nutrition (Large) | 6 participants: 3 women, 3 man<br>65 – 80 years |

#### *2.2.3. Procedure*

At the beginning of each group, the moderator for the group introduced the objectives and procedures for the discussion. Participants completed an informed consent form and a demographic questionnaire. The moderator then took the group through the discussion of the themes, bringing the discussion back to the topic as needed. At the end of the discussion, the moderator debriefed and summarised the key points mentioned by the participants during the discussion. The discussions were audio-recorded for later detailed analysis. The length of the discussions ranged from 54 minutes to 109 minutes, with a mean of 85 minutes.

### *2.3. Results*

Table 2 shows the total number of contributions made in each focus group and the number of contributions per participant and per 10 minutes of discussion time. As can be seen, there are very big differences between the groups, which is not surprising,

given individual differences between people in their willingness to contribute to a group discussion, their interest and knowledge on a topic and the dynamics of the group. It is perhaps not surprising that the larger focus groups produce less contributions per participant, as there are more people competing to add something to the discussion. But it is interesting that for both topics and overall, the number of contributions per 10 minutes of discussion is substantially higher in the larger groups than in the smaller groups. So it does not seem that the larger groups of older adults has any more difficulty generating contributions than the smaller groups.

**Table 2.** Total number of contributions, contributions per person and per time for the four focus groups

| | **Small Focus Group** | **Large Focus Group** |
|---|---|---|
| **Co-Motion** | Participants: 4<br>Total time: 80 min<br>Total contributions: 92<br>23.0 contributions per participant<br>11.5 contributions per 10 minutes | Participants: 7<br>Total time: 54 min<br>Total contributions: 100<br>14.3 contributions per participant<br>18.5 contributions per 10 minutes |
| **Nutrition** | Participants: 3<br>Total time: 109 min<br>Total contributions: 188<br>62.0 contributions per participant<br>17.25 contributions per 10 minutes | Participants: 6<br>Total time: 97 min<br>Total contributions: 218<br>36.3 contributions per participant<br>22.5 contributions per 10 minutes |
| **Overall** | 42.5 contributions per participant<br>14.3 contributions per 10 minutes | 25.3 contributions per participant<br>20.5 contributions per 10 minutes |

In terms of conducting the focus groups with both large and small groups, the moderators had minimum difficulties. Some participants were clearly more forthcoming than others and the moderators attempted to ensure that each participant had an opportunity to speak and share their thoughts, drawing more reticent participants into the discussion when necessary. The moderators also made sure that the discussions kept reasonably close to on the topic of the discussion. Interestingly, with these groups there was little diverging from the topic by the participants, and very few examples of participants having side discussions. We found that the participants in all groups were quite firm with their thoughts and ideas. There were some participants who tended to try dominate the discussions and further analysis of the data will investigate the group dynamics and the number of different distinct themes that emerged from groups of different sizes.

## 3. Case Study 2: Collaborative Heuristic Evaluation (CHE)

### *3.1. Introduction*

When developing new technologies, be they for younger or older users, one does not want to ask potential users to give opinions about new systems until any obvious usability problems have been eliminated. Yet in developing a new system, one can get

very close to a design and thing which seem clear to the developers are not necessarily clear to the users. So a common method is to initially conduct an evaluation with experts, eliminate the problems the experts think users might have, and only then start evaluations with users.

The most well-known expert evaluation method is heuristic evaluation (HE), developed by Nielsen and Molich in the early 1990s [18, 19]. In an HE, three to five experts are asked to work through a system, looking for problems that users might have, using a set of heuristics. Heuristics are short, easy to remember principles of good interface design. Nielsen and Molich developed a set of 10 heuristics based on their work of evaluating many interactive systems. After the experts have worked through the system, they come together and discuss all the potential problems they have found and come up with an agreed list of problems, and rate them on a four point scale for severity (catastrophic problem, major problem, minor problem, cosmetic problem). Since Nielsen and Molich's original work, many different sets of heuristics for different types of systems have been developed. For example, there are heuristics for evaluating mobile computing applications [20], for interactive websites [21] and for smartphone applications for older adults [22].

Although HE has been very popular, experts actually find it rather tedious to conduct. Petrie and Buykx [6] developed a variation of HE, collaborative heuristic evaluation (CHE) in which the experts work as a group to identify potential usability problems, but rate the problems privately (to allow them to disagree about the severity of particular problems). This has been shown to an effective method for conducting an expert evaluation, and allows experts with different areas of expertise (for example, expertise in mobile applications, expertise in the needs of older adults) to work together.

The aim of Case Study 2 was to evaluate the heuristics proposed by Silva et al. [22] which were specifically developed to evaluate smartphone applications for older adults in the context of a CHE.

### *3.2. Method*

#### *3.2.1. Design*

A low-fidelity prototype of an "app' to support older adults in monitoring their liquid intake was designed for iPhone and iPad platforms using Apple's design conventions. The heuristics developed by Silva et al. [22] were also used in guiding the design. A CHE was conducted on a paper prototype of the app.

#### *3.2.2. Experts*

Four experts participated in the CHE. All experts were members of the Human Computer Interaction Research Group at the University of York. Two experts were professionals and two experts were postgraduate students in the Group. All experts had experience with conducting CHEs and with interactive systems for older adults. Three experts were Apple device users and one was an Android device user.

#### *3.2.3. Procedure*

The CHE took place over three sessions. Two sessions were to evaluate the prototype for a smartphone (an iPhone version was developed) and one session was to evaluate the prototype for a tablet computer (an iPad was developed). All experts shared and

viewed the same prototype. Each CHE session was audio-recorded for later detailed analysis. Each session lasted approximately 75 minutes. During the session, the first author guided the group through prototype while the experts identified potential problems and asked questions. One expert recorded the potential usability problems raised. The heuristics relevant to each problem were recorded. The experts privately rated the potential severity problems using a five-point scale (1 = very minor problem to 5 = very major problem).

**Table 3.** Comments on Silva et al. heuristics (* = used in iPhone evaluation, ‡ = used in iPad evaluation)

| Heuristic | Comments from experts |
|---|---|
| Focus on one task at a time instead of requiring the user to actively monitor two or more tasks, and clearly indicate the name and status of the task at all times. * ‡ | Three different points in one heuristic. One heuristic should be about focussing one task at a time; the other about clearly labelling the task, so the user knows what they are doing; and the third about providing information about the status of the task at all times. |
| Avoid the use of interaction timeouts and provide ample time to read information. | Contradictory information – if the developer is to avoid timeouts, then providing ample time is not relevant. |
| Avoid the use of animation and fast-moving objects. | |
| Leverage mental models familiar to older adults. * ‡ | Difficult to know what mental models older adults have. This is not a heuristic. |
| Reduce the demand on working memory by supporting recognition rather than recall.* ‡ | |
| Aim at creating an aesthetical user interface, by using pictures and/or graphics purposefully and adequately to minimize user interface clutter and avoid extraneous details.* ‡ | Two different points in one heuristic. One heuristic should be about using pictures and / or images to create a suitable interface; the other is about to minimize user interface clutter and avoid unnecessary details. Simpler words should be used. |
| Give specific and clear instructions and make help and documentation available. Remember that it is better to prevent an error than to recover from it.* ‡ | Two different points in one heuristic. Providing specific / clear instructions and having help documentations are two different heuristics. |
| Provide clear feedback and when presenting error messages make them simple and easy to follow.* ‡ | |
| Make sure errors messages are descriptive and use meaningful words and verbs when requiring an action.* ‡ | Partly overlaps with previous heuristic. |
| Write in a language that is simple, clear and adequate to the audience.* | |
| Avoid pull down menus. | |
| Avoid the use of scrolling.* | What is the evidence that older adults do not like scrolling? |
| Enlarge the size of user interface elements in general; targets should be at least 14mm square.* | What is the evidence for this size? |
| Keep the user interface navigation structure narrow, simple and straightforward. ‡ | |

| | |
|---|---|
| Use consistent and explicit step-by-step navigation.* ‡ | |
| Make sure that the "Back" button behaves predictably. | |
| Support user control and freedom, allowing or alternative and flexible flows of interaction. * ‡ | |
| Disable inactive user interface objects. | |
| Do not rely on colour alone to convey information. Be aware of colour blindness. | |
| Provide not only visual feedback, but also tactile and auditory. | |
| Make information accessible through different modalities. | |
| Use lower frequencies to convey auditory information such as confirmation tones and alerts. | |
| Do not use pure white or rapidly changing contrast backgrounds. ‡ | |
| Make it easy for people to change the text size directly from the screen. | |
| Use high-contrast colour combinations of font and/or graphics and background to ensure readability and perceptibility; avoid using blue, green and yellow in close proximity. * | Two different points. One heuristic should be about not using high-colour contrast combinations to ensure readability and perceptibility. The other heuristic is confusing. The wordings are ambiguous. |
| Use colour conservatively, limiting the maximum number of colours in use to ~four. | |
| Make sure text uses types, styles and sizes appropriate to older adults, that is, for instance, but not exclusively: large-sized fonts, sans serif, non-condensed typefaces, non-italic, and left justified. * | |
| Make links and buttons clearly visible and distinguishable from other user interface elements. * | |
| Make information easy to read, skim (or) and scan. | |
| Group information visually (make good use of colour, text, topics, etc.). | |
| Allow sufficient white space to ensure a balanced user interface design. * | |
| Use user interface elements consistently and adhere to standards and conventions if those exist. * | Two different points in one heuristic. One heuristics should be on having consistent user interface elements; the other should be about adhering to standards and conventions (if those exists) |
| Use simple and meaningful icons. * | |

### *3.3. Results*

51 potential usability problems were identified in the iPhone prototype and 17 in the iPad prototype. Only one problem in each version was rated as 5 "very major", the remaining problems were divided between 4 (major, 52.9%) and 3 (minor, 42.6%) and only one problem was rated as 1 (very minor).

Only three potential usability problems identified were not addressed by the heuristics proposed by Silva et al. [22]. Of the 33 heuristics proposed, 20 were used in the evaluations. It is not surprising that only 20 were used in the evaluations because some heuristics were not relevant for the prototype. For example, the prototype does not contain any animations or moving objects. However, the experts had numerous difficulties in using the heuristics, which are summarized in Table 3. The experts also commented that the set of heuristics is very long and has no high level structure, which adds a further layer of difficulty in using them. Finally the experts noted that many of the heuristics feature in general usability heuristics and it would be helpful to have a set of heuristics which concentrate on the additional aspects important for older users, and not mix them up with heuristics for all users, which people are likely to already be familiar with.

## 4. Case Study 3: Think-aloud Protocol

The gold standard for the evaluation of new technologies is testing with potential users of systems, to see how they use a system and what problems they have with it. To understand users' experience, practitioners and researchers often use a technique called a "concurrent verbal protocol" or more informally "think aloud protocol" [7, 23]. In this technique, a user is given a number of typical tasks to do with a system, and as they do them they are asked to "think aloud", that is talk through what they are doing, what problems they are encountering. This of course adds to the mental effort of doing the task and some people find it quite difficult to do. A number of researchers [e.g. 23, 24] have investigated the mental effort, how this might change the task for the user and whether particular kinds of usability problems are detected with this technique. The effectiveness of the think aloud protocol has been investigated for blind users [25], but to the best of our knowledge, no research has investigated its use with older users and whether this raises particular issues.

Thus for Case Study 3, we investigated the use of think-aloud protocol with older adults. They evaluated a low-fidelity paper prototype to monitor their liquid intake, which had been developed based on information gathered in the focus groups reported in Case Study 1, the CHE evaluation in Case Study 2 and the heuristics developed by Silva et al [22]. Currently, many researchers include older adults early in the design process to try to create acceptable and usable systems for this user group [e.g. 26, 27]. However, a review Zapata et al. [28] found that only one out of 18 studies conducted a think-aloud protocol to evaluate the usability of mobile health-related "apps". Thus a very important stage of evaluation is typically being missed.

### *4.1. Method*

#### *4.1.1. Design*

A low-fidelity prototype of an app to support older adults to monitor their liquid intake was developed based on the suggestions given by the older adults in Case Study 1 and the CHE evaluation in Case Study 2. Twenty tasks were designed, for example registering a user account, updating liquid intake and viewing liquid intake history. Four of the tasks included different alternative design options (e.g. adding liquid intake via glasses or cups).

Each time a participant proposed a potential usability problem, they were asked to briefly explain the problem and rate it on the same five point scale used in the CHE. For each task with different design options, participants were asked to choose which option they prefer or they could suggest other possible designs.

#### *4.1.2. Participants*

The inclusion criteria were to be 65 years or over and living independently, either alone or with a partner. There were 20 participants (10 women, 10 men) with the mean age of 70 years. Six participants lived alone. Two participants had a highest education level of primary school, seven had secondary school, three had a bachelors degree, one had a post-graduate degree, and four had professional qualifications. Nineteen participants are retirees, one works part-time. In addition, 16 participants are Internet users with experiences of using the Internet from 2 years to more than 20 years. Twelve participants are computer users with experience of using the computers from 5 months to more than 30 years. Thirteen participants are tablet computer users with experience of using the device from 5 months to 5 years.

#### *4.1.3. Procedure*

The study took place in the Interaction Labs at the University or York, or at the participant's own home. Participants were first briefed about the study and completed an informed consent form. All participants viewed the same prototype. The first author guided the participants through the prototypes and gathered the usability problems, possible suggestions and comments made by the participants. Each session was audio-recorded for later detailed analysis. Each session lasted approximately 75 minutes. After completing session, participants completed a short demographic questionnaire. Participants were then debriefed and invited to ask any questions about the study. Participants were offered a gift voucher worth £25 (approximately USD 36) to thank them for their time and efforts.

### *4.2. Results*

With the use of think-aloud protocol, participants appeared to freely to critique, comment, and suggest the design of the app despite having different technology background, particularly in using touchscreen.

Apart from the usability problems that were found, the content analysis of the evaluations showed that this group of participants preferred an app that is simple and straightforward. These participants also prefer an app that uses lesser amount of steps, time and effort to complete a single task. For example, to register a user account, these

participants prefer to do the registration in one page rather than having to enter the required details, for example e-mail, password and username, one-by-one in different pages.

Participants also did not like long and repetitive instructions. For example, Figure 1 (left), shows how the user was instructed to enter their username and password and how to proceed in using the app. Participants preferred exploring the app via try-and-error to figure out what each feature does rather than reading instructions. This is contrary to the stereotype of older adults as requiring extensive instructions and being cautious about how to proceed. We noticed that participants would simply skim long instruction pages.

Most interestingly, we also found that seven participants did not like the use of a picker (or spinner) to enter numbers such as their weight and number of glasses of liquid consumed to the mobile app. Figure 1 (right) shows the use of picker in the app to update liquid intake. Participants suggested using keypad, buttons, or pull-down menu instead.

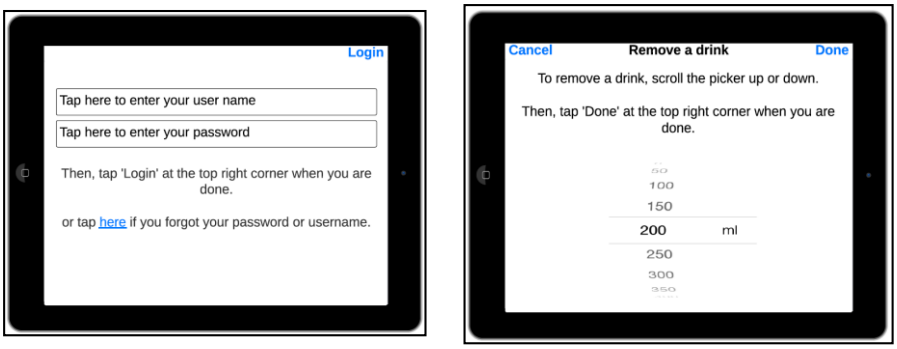


**Figure 1.** Usability problems identified on the iPad version by the participants in Case Study 3: long and repetitive instructions at login (left), and the picker as the interaction style to update liquid intake (right).

## 5. Discussions and Conclusion

This paper reports on three case studies with different methods for working with older adults to design and evaluate technologies to support their independent living.

For Case Study 1, we aimed to concentrate on the issue of the size of focus group with older adults and how this affected the information elicited. Our findings showed that the smaller focus groups elicited more information per person but the larger groups elicited more information per unit of time of the focus groups. These results are logical in terms of people's opportunity to express their views. In terms of managing the focus groups with older adults, we found that the focus groups functioned in very similar ways to groups with younger participants. The moderator may need to work harder, to take into account the various difficulties that older participants might be having, for example hearing discussion or seeing visual materials. However, this is simply an extension of the work of the moderator. It is very helpful to know in advance what difficulties people might have, so appropriate arrangements can be made. But in several of our groups, we extemporized solutions.

For Case Study 2, a low-fidelity prototype was evaluated by a group of experts using the Collaborative Heuristic Evaluation (CHE) method [6] and particularly the heuristics developed by Silva et al [22]. However, the experts struggled to use these heuristics, although they did find that most of the problems identified could be categorised by them. In addition, the experts raised numerous concerns about the appropriateness, wording and clarity of the heuristics. Heuristics should be simple, clear and straightforward. Each heuristic should consist of just one point. However, the Silva et al. heuristics had numerous heuristics consisting of two or more points per heuristic. The experts also queried the evidence base for some of the heuristics, as Silva et al do not provide detailed information about how they were identified. For example, one heuristics is "Avoid the use of scrolling." This surprised the experts and Piper, Campbell, and Hollan [29] found that the older adults preferred scrolling using hands gestures as compared to using buttons whilst using a touch screen. However, Page [30] found that older adults struggle with the concept of scrolling on touchscreens. Thus the question of older adults use of scrolling seems unresolved.

For Case Study 3, we evaluated the usability of a low-fidelity liquid monitoring prototype with older adults using the think-aloud protocol [7]. We found that older adults did not appear to have any particular difficulties with the protocol, but a more detailed study is needed, also comparing concurrent think aloud with retrospective think aloud (in which participants do the tasks and then talk through them while watching a video of their experience). Our think aloud protocol study produced many usability problems, as well as new information about the older participants' attitudes to mobile technology. In particular, we found that numerous participants did not like interaction with the picker; this also needs further investigation.

Overall we have found that these three well known methods very typical of user-centred design work well with older adults, although care needs to be taken to adjust them appropriately for the particular needs of the participants involved.

## References


[1] United Nations. 2015. World Population Ageing 2015. Available at: http://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2015_Report.pdf

[2] Office National Statistics. 2014. Aging, June 26, 2014. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/ageing

[3] Lloyd-Williams, M., Kennedy, V., Sixsmith, A. and Sixsmith, J. The end of life: a qualitative study of the perceptions of people over the age of 80 on issues surrounding death and dying, *Journal of Pain and Symptom Management,* 34(1), 60-66, 2007.

[4] Fan, C., Forlizzi, J., and Dey, A. 2012. Considerations for technology that support physical activity by older adults. Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '12). New York: ACM Press.

[5] Lines, L. and Hone, K. S. 2004. Eliciting user requirements with older adults: lessons from the design of an interactive domestic alarm system, *Universal Access in the Information Society,* 3(2), 141-148.

[6] Petrie, H. and Buykx, L. 2010. Collaborative Heuristic Evaluation: improving the effectiveness of heuristic evaluation, In *Proceedings of UPA 2010 International Conference*. Omnipress. Available at: http://upa.omnibooksonline.com/index.htm.

[7] Lewis, C. 1982. *Using the "thinking-aloud" method in cognitive interface design*. IBM TJ Watson Research Centre.

[8] Elia, M., and Russell, C. 2009. Combating malnutrition: recommendations for action. In M. Elia, and C. Russell (Eds.) *Combating malnutrition: recommendations for action*. Nutrition Advisory Group on malnutrition.

[9] Preece, J., Sharp, H., and Rogers, Y. 2015. *Interaction Design: Beyond Human-Computer Interaction*. Wiley.

[10] Lazar, J., Feng, J.H. and Hochheiser, H. 2010. *Research Methods in Human-Computer Interaction*. Wiley.
[11] Adams, A. and Cox, A.L. 2008. Questionnaires, in-depth interviews and focus groups., In P. Cairns and A.L. Cox (Eds.), *Research Methods for Human-Computer Interaction*. Cambridge University Press.
[12] Brondani, M.A. MacEntee, M.I. Bryant, S. R. and O'Neill, B. 2008. Using written vignettes in focus groups among older adults to discuss oral health as a sensitive topic, *Qualitative Health Research,* 18(8), 1145-1153.
[13] Hawthorne, G., Davidson, N., Quinn, K., McCrate, G., Winkler, I. Lucas, R. Kilian, R. and Molzahn, A. 2006. Issues in conducting cross-cultural research: implementation of an agreed international procotol designed by the WHOQOL Group for the conduct of focus groups eliciting the quality of life of older adults, *Quality of Life Research,* 15(7), 1257-1270.
[14] Inglis, E.A., Szymkowiak, A., Gregor, P., Newell, A.F., Hine, N., Shah, P., Wilson, B.A. and Evans, J. 2003. Issues surrounding the user-centred development of a new interactive memory aid, *Universal Access in the Information Society,*2(3), 226-234, 2003.
[15] Lyons, K.J., Salsbury, S.A., Hondras, M. A., Jones, M. E. Andresen, A. A. and Goertz, C. M. 2013. Perspectives of older adults on co-management of low back pain by doctors of chiropractic and family medicine physicians: a focus group study. *BMC Complementary and Alternative Medicine,* 13(1), 225.
[16] Swallow, D., Petrie, H., Power, C., Lewis, A. and Edwards, A. Involving older adults in the technology design process: a case study on mobility and wellbeing in the built environment. (this volume).
[17] Bevan, M., Petrie, H., Cambridge, H., Cinderby, S., Croucher, K., Swallow, D., Gilroy, R. and Attuyer, K. (this volume). Co-Motion: Mobility and wellbeing in later life. (this volume).
[18] Nielsen, J. 1994. *Usability engineering*: Elsevier.
[19] Nielsen, J. How to conduct a heuristic evaluation. Available at: https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/
[20] Bertini, E., Gabrielli, S. and Kimani, S. 2006. Appropriating and assessing heuristics for mobile computing. *Proceedings of the Working Conference on Advanced Visual Interfaces (AVI '06*). New York: ACM Press.
[21] Petrie, H. and Power, C. 2012. What do users really care about?: a comparison of usability problems found by users and experts on highly interactive websites. *Proceedings of the 30th International Conference on Human Factors in Computing (CHI 2012)*. New York: ACM Press.
[22] Silva, P.A., Holden, K. and Jordan, P. 2015. 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, *Proceedings of the 48th Annual Hawaii International Conference on System Sciences*, 3237-3246.
[23] Krahmer, E. and Ummelen, N. 2004. Thinking about thinking aloud: A comparison of two verbal protocols for usability testing. IEEE Transactions of Professional Communications 47, 105-117.
[24] Olmsted-Hawala, E.L., Murphy, E.D., Hawala, S., and Ashenfelter, K.T. 2010. Think-aloud protocols: a comparison of three think-aloud protocols for use in testing data-dissemination web sites for usability. *Proceedings of SIGCHI Conference on Human Factors in Computing Systems*, New York: ACM Press. York.
[25] Savva, A., Petrie, H. and Power, C. 2015. Comparing concurrent and retrospective verbal protocols for blind and sighted users. *Proceedings of the 15th IFIP International Conference on Human-Computer Interaction (INTERACT 2015)*. Springer.
[26] Massimi, M., Baecker, R.M. and Wu, M. Using participatory activities with seniors to critique, build, and evaluate mobile phones, *Proceedings of the 9th International ACM SIGACCESS Conference on Computers and Accessibility* (ASSETS '07), New York: ACM Press.
[27] Davidson, J.L. and C. Jensen, C. 2013. Participatory design with older adults: an analysis of creativity in the design of mobile healthcare applications. Proceedings of the 9th ACM Conference on Creativity and Cognition. New York: ACM Press.
[28] Zapata, B.C., Fernández-Alemán, J. L., Idri, A. and Toval, A. 2015. Empirical studies on usability of mHealth apps: A systematic literature review, *Journal of Medical Systems,* 39(2), 1-19.
[29] Piper, A.M., Campbell, R. and Hollan, J.E. 2010. Exploring the accessibility and appeal of surface computing for older adult health care support. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. New York: ACM Press.
[30] Page, T. 2014. Touchscreen mobile devices and older adults: a usability study, *International Journal of Human Factors and Ergonomics,* 3(1), 65-85.