Paper:
http://dx.doi.org/10.3233/JAD-160545

This article is brought to you by Swansea University. Any person downloading material is agreeing to abide by the terms of the repository licence. Authors are personally responsible for adhering to publisher restrictions or conditions. When uploading content they are required to comply with their publisher agreement and the SHERPA RoMEO database to judge whether or not it is copyright safe to add this version of the paper to this repository.
http://www.swansea.ac.uk/iss/researchsupport/cronfa-support/
Administering Cognitive Tests Through Touch Screen Tablet Devices: Potential Issues

Amy Jenkins a, Stephen Lindsay b, Parisa Eslambolchilar b, Ian M. Thornton c

a Department of Psychology, Swansea University, Swansea, UK
b Department of Computer Science, Swansea University, Swansea, UK
c Department of Cognitive Science, University of Malta, Malta

Accepted 28 June 2016

Abstract. Mobile technologies, such as tablet devices, open up new possibilities for health-related diagnosis, monitoring, and intervention for older adults and healthcare practitioners. Current evaluations of cognitive integrity typically occur within clinical settings, such as memory clinics, using pen and paper or computer-based tests. In the present study, we investigate the challenges associated with transferring such tests to touch-based, mobile technology platforms from an older adult perspective. Problems may include individual variability in technical familiarity and acceptance; various factors influencing usability; acceptability; response characteristics and thus validity per se of a given test. For the results of mobile technology-based tests of reaction time to be valid and related to disease status rather than extraneous variables, it is imperative the whole test process is investigated in order to determine potential effects before the test is fully developed. Researchers have emphasized the importance of including the ‘user’ in the evaluation of such devices; thus we performed a focus group-based qualitative assessment of the processes involved in the administration and performance of a tablet-based version of a typical information processing speed (a multi-item localization task), to younger and older adults. We report that developers should consider factors surrounding user expectations, performance feedback, and physical response requirements in order to inform further research into such applications.

Keywords: Aging, attention, cognition, focus groups, qualitative research, tablet computers

INTRODUCTION

The past five years have seen a rapid growth in the number of people over the age of 65 using mobile devices. Almost one in five older adults in the United States possess a smart phone with increased usage driven by factors such as the advanced capabilities of smart devices, the value placed on the ability to communicate with relatives, and the perceived usability of touch screen technology [1, 2]. The trend opens new avenues for adjuncts to health-related diagnosis, monitoring, and intervention and thus the delivery of healthcare to a population that typically find it hard to access such services. This is of particular relevance for older adults who are increasingly at risk of developing dementia and associated disorders, and an often-corresponding reduction in both mobility and the ability to access healthcare services. As a result of increased engagement with digital technology devices such as tablets and smartphones, mass healthcare monitoring in older adulthood is a real possibility. Furthermore, healthcare solutions...
A. Jenkins et al. / Administering Cognitive Tests: Tablet Devices

scale up for a large number of users and demand.

Mobile technology (mHealth) has been applied to different healthcare challenges to address the growing availability of broadband, wireless internet, the idea of ‘information to support the extended beyond traditional medical care providers a platform for community-based, virtual patient/caregiver interaction [11, 12].

The test environment, and any administrator/patient interaction [11–13, 23–27]. Such factors are particularly pertinent to the test-taking experience of an individual's reaction speed and variability. RT speed and variability are measures regularly employed in behavioral indicators of the speed of information processing and the integrity of cognitive and motor skills [2, 18–22], all factors specific to the use of a touch screen tablet and thus this platform can introduce new biases may, for example, alter the accuracy, validity, and clinical relevance of the results, when used clinically. RT speed and variability appear to be behavioral indicators of the integrity (at least in part) of white and grey matter [17] in older adulthood and neurodegenerative dementia processes such as Alzheimer's disease, such measures may be of use clinically.

Arguably, RT and IIV\textsubscript{RT} testing appear particularly suited to delivery or presentation via a touch screen tablet as they tend to be cheaper and simpler to use than laptops or desktop computers and can have multiple advantages over computers for testing information processing in older adults [9, 18, 19]. However, it is also increasingly clear that factors unrelated to brain structure and function and a disease process can influence RT and IIV\textsubscript{RT} and that it is vital to determine, investigate, and ameliorate such effects with respect to the touch screen tablet platform, in order to ensure test validity.

Evidence already reveals that there are a number of challenges to be aware of when digital technologies are used by older adults including physical issues such as decline in manual dexterity and eyesight and decreasing cognitive capabilities, frustration, the need for specific training, age, gender, dry finger skin, and age-related cognitive motor skills [2, 18–22], all factors likely to affect the performance of RT and IIV\textsubscript{RT} tests using a touch screen platform and thus their clinical validity, usefulness, and robustness. Furthermore, RT research has revealed many participant and methodology-related factors capable of significantly affecting RT study outcome including: the test item, the environment, response requirements,
perform sequences of actions [29, 30] to other well-established paper-and-pencil (e.g., The Trail Making Task [31]) and cancellation tests [32] in requiring a sequence to be identified in a specific order.

A typical trial from the tablet implementation that was used in the current study would be to touch all in sequence, from one to eight. The general advantages of computer-based presentation as compared to paper-and-pencil tasks include the recording of RTs for each item, rather than simply overall completion time (e.g., [32]) and the ability to easily explore spatial patterns of search organization (e.g., [34]). In addition to these, the MILO task makes it possible to easily manipulate the sequence type (e.g., letters, digits, or both) and sequence behavior (e.g., items vanishing or remaining, sequence position fixed or shuffling between responses), to explore the temporal context of visual search [29]. Such a task therefore represents the type that might be considered for use in a clinical situation, providing information about RT speed and variability, and attention processing and other aspects of higher level, cognitive processing.

MATERIALS AND METHODS

For the purpose of the current study, we used a fixed sequence of the digits one to eight, and configured the display so that items vanished when touched. Although this MILO configuration was not initially designed specifically for use with older adults, we chose the task specifically because the display layout and physical response demands were appropriate for use with this population [35–37]. For example, there are a number of challenges to be aware of when digital technologies are used by older adults including physical issues such as decline in manual dexterity and eyesight and decreasing cognitive capabilities, both potentially hindering interaction with mobile platforms, which are not adapted to their needs [18, 19, 22]. In the MILO task, the target object size and spacing were well within these suggested limits and responses could be self-paced. More specifically, when the iPad was placed on a table 50 cm in front of participants, each 1.9 cm item subtended approximately 2° visual angle, with gaps between items varying between 0.8° and 8° visual angle. To successfully complete a trial, participants were required to touch each object following the numeric sequence one to eight as quickly as possible, but there were no specific time limits, so participants could calibrate their responses taking into account any motor limitations.

When an item was touched, it vanished from the screen, so that the set size, and search difficulty was reduced with each response. Touching an item out of sequence (i.e., a mistake) resulted in the termination of the trial and visual feedback in the form of a schematic sad face. There was a two second inter-trial interval and no feedback on speed or accuracy was provided for correct trials. Each participant completed 10 training and up to 10 experimental trials and at the start of each trial the position of all target items was randomized within the constraints of a virtual grid that was programmed to ensure items did not overlap. As our goal was to explore factors related to presenting a RT task using a touch screen tablet format per se, we did not record actual RT performance as participants were allowed to comment upon any aspects the task while they were doing it. Instead, as detailed below, we used a focus-group design to make a qualitative assessment of individuals' experiences and device usability.

In an approach that is interdisciplinary and draws from Human Computer Interaction (HCI) and User Experience (UX) research traditions, a focus group approach was adopted in order to determine from the individuals themselves potential issues relating to the use of mobile technology for cognitive testing that may influence the RT results. To provide information of relevance to real life test scenarios, as it is common in MILO and similar computer-based tests of attention and cognition to provide on-screen feedback using a visual or auditory warning indicative of incorrect response, we also investigated the potential
The focus groups were audio-recorded, and a member of the research team took notes. A semi-structured predetermined framework of open-ended questions was used to ensure all aspects relating to the topic area were explored (Table 1). The focus group recordings were transcribed verbatim, and all identifiable information was either removed or consistently anonymized. Thematic analysis was employed on the interview data, which was realist driven, inductive, and bottom-up [38]. Two members of the research team read and re-read the transcripts making initial comments and codes. The process was repeated twice more until individual codes were identified. Subsequently these were grouped into three major themes that emerged across both younger and older participant groups, namely ‘views of test experience’, ‘testing situation and materials’, and ‘test performance’.

**RESULTS**

A number of themes and sub-themes have been identified highlighting categories rather than prevalence. Furthermore, the researcher administering the test typically sits close to the person taking the test; anecdotally this has been off-putting to the person taking the test. It may also be reassuring to know that feedback may be provided at any time during a test administration. This rationale may also be explained by the real-time feedback upon task acceptability and performance. Furthermore, the researcher administering the test typically sits close to the person taking the test, and it may also be reassuring to know that feedback may be provided at any time during a test administration. The three major themes that emerged across both younger and older participant groups, namely ‘views of test experience’, ‘testing situation and materials’, and ‘test performance’.

**Table 1**

<table>
<thead>
<tr>
<th>Focus group section</th>
<th>Questions and prompts</th>
</tr>
</thead>
<tbody>
<tr>
<td>iPad test feedback</td>
<td>-Has anyone used an iPad/similar device before?</td>
</tr>
<tr>
<td>questions</td>
<td>-How would you describe your experiences of using the test?</td>
</tr>
<tr>
<td></td>
<td>-Prompt – was it enjoyable or not?</td>
</tr>
<tr>
<td></td>
<td>-How well did you think you have done?</td>
</tr>
<tr>
<td></td>
<td>-Prompt - was it too fast? Hard to pay attention to, etc.?</td>
</tr>
<tr>
<td></td>
<td>-Was the iPad easy to use?</td>
</tr>
</tbody>
</table>

The focus groups were audio-recorded, and a member of the research team took notes. A semi-structured predetermined framework of open-ended questions was used to ensure all aspects relating to the topic area were explored (Table 1). The focus group recordings were transcribed verbatim, and all identifiable information was either removed or consistently anonymized. Thematic analysis was employed on the interview data, which was realist driven, inductive, and bottom-up [38]. Two members of the research team read and re-read the transcripts making initial comments and codes. The process was repeated twice more until individual codes were identified. Subsequently these were grouped into three major themes that emerged across both younger and older participant groups, namely ‘views of test experience’, ‘testing situation and materials’, and ‘test performance’.
Fig. 2. Views of test experience.

First is the sub-theme ‘absorbing’ which represents the view that some of the older participants said they were absorbed into the iPad test experience. For instance:

“[W]: I found it quite absorbing myself because you had to concentrate on what was in front of you to pinpoint what the next number was. I have to say it occupied all my thoughts I was just trying to do it as quickly as I could, and as accurately as I could. I was totally absorbed by those 1–8 numbers. Which is strange for me because my mind does tend to wander and it didn’t wander on that occasion”.

The second sub-theme reflects the older participants competing views that the test was a ‘challenge’, and the third sub-theme that it was ‘easy’. For instance:

Challenge: “[J]: I found it absolutely entertaining. I found it quite a challenge [mumbling]. I was sort of trying to do it quite quickly, I failed a couple of times but I think that was these [pointing out his fingers]”.

Easy: “[RA]: I thought it was easier than I thought it would be. I thought ‘I have never used an iPad before!’ And sometimes when I go onto the computer I press something and it goes off, I have done that a few times actually. The iPad I made a few mistakes”.

The sub-theme ‘positive experience’ was a shared view of both the older and younger groups. For instance:

Positive experience (older): “[P]: it was quite enjoyable. [W]: and I think the more you did it the more you wanted to do it somehow”.

Positive experience (younger): “[R]: fab, thank you. Did you enjoy doing the test? [A]: it makes me want one [iPad]. [P]: it was interesting but I wouldn’t use the word ‘enjoy’ [laughter] I was just counting dots but it was a little more engaging that some can be. [S]: it made me wonder if they were dots or pool balls [laughter] I think it was nice that it changed on each trial. Like in a paper pencil version of a trail making there is only one set way of doing it and I like having the variation that it is new every time you do it, maybe it is more accurate that way”.

The sub-theme ‘boring and distractible’ is also a shared view in opposition to the test being a positive experience. For instance:

Boring and distractible (older): “[R]: so how did you find the test? [G]: a bit boring I found it, sorry. Repetitively boring there was obviously a sequence for that. I said that to [researcher] I said ‘is this um could you memorise these if you had a good memory and numerative memory?’ The problem is going too fast and then thinking something more interesting may come up next time. It was the same numbers just in a different location. Yeah I found it boring towards the end. [R]: yes and that is perfectly fine, I want you to be as honest as you can. Thank you [G]”.

Boring and distractible (younger): “[R]: ok, so would you say then something like that could be used on a regular basis or would you say no? [L]: I think it was boring”.

The sub-theme ‘game’ is unique to the younger group and represents the view that the test was like a ‘game’. For instance:

“[B]: it was like many games that you can get on the iPad already, like I have a few already that are similar. [R]: are there any that you think are similar to it? [S]: I wouldn’t know. [A]: not sure. [P]: when she was initially explaining it to me it did kind of remind me almost of like a word search type thing because you are obviously looking for like a 1 and then linking it. [B]: I have quite a few games where you have to link patterns between things and there is ummm well I have about 5 on here and there are millions available as well like [famous game]. [R]: yeah it is a similar thing”.
isn't it. [S]: see I was thinking well what the purpose of the game is, what it is going to be used as. For example, if it is something to do with cognitive training then I wondered what well if it would be of any use to have like a kind of positive feedback mechanism put in because I made a mistake and there was a little sad face and that was feedback too but you know to get people to play it maybe more regularly maybe it would have like increasing difficulty and a score. That would make them think don’t know if I would play it in the sake of doing it as it is now just like tapping the numbers and I want to know that I am doing good. [A]: yeah like in games you want to improve and beat your score. [S]: yeah like progression or how well I am doing. [B]: or different levels, like the next level could have like 10 numbers”.

Testing situation and materials

This second theme has three sub-themes developed from the findings of both the older and younger groups (Fig. 3). The first sub-theme reflects the views of both groups regarding the experience they had of using the iPad. For instance:

Device experience (older): “[R]: yes but she was doing scores, what’s more important is what I get from the feedback from the tests. Did you find it easy to use? [A]: yeah. [G]: well I did do very well but it was fine. [J]: I made two mistakes the same as you; as soon as I slowed down a bit I was more accurate. And these glasses [glasses], but it was difficult to read them back up. [R]: yeah ok so...

Device experience (younger): “[R]: ok, thank you. How about the positioning of the iPad? [L]: fine. [P]: I moved it. [R]: where did you move it to? [P]: I just moved it closer. The angle was a bit well I didn’t move the angle. For me it would have been better flat but maybe because it was quite far into the table. [RB]: it would have been helpful to have one of those holders, what are they called? [P]: like a copy holder? [RB]: yeah, just to have it in front of you. I wonder what that would be called...

Fig. 3. Testing situation and materials.
The first sub-theme ‘accuracy’ is based only on the older participants. For instance:

“[R]: so what did you think? Was it due to more accuracy or speed? [N]: a combination of both I think. [P]: yeah it is no good going fast if you’re going to get it all wrong is there. [J]: I was disappointed with the number of mistakes I did make, obviously trying to go too fast. [P]: I made one but I think it was because I didn’t press hard enough on the screen. The face came up [showing sad face]”.

The second sub-theme is the ‘use of hands’ whilst using the iPad. For instance:

Use of hands (older group): “[A]: the only problem I had with the touch screen is my nails. I have this problem at home, and that’s why I use a [brand name] pen because I find you have to develop a certain technique of touching. You can’t just go like that [action] because your nail would touch it and that doesn’t work so you have to slide off rather than...and I found that at home. But as I...”

The second sub-theme is the ‘use of hands’ whilst using the iPad. For instance:

Use of hands (older group): “[A]: the only problem I had with the touch screen is my nails. I have this problem at home, and that’s why I use a [brand name] pen because I find you have to develop a certain technique of touching. You can’t just go like that [action] because your nail would touch it and that doesn’t work so you have to slide off rather than...and I found that at home. But as I...”

The second sub-theme is the ‘use of hands’ whilst using the iPad. For instance:

Use of hands (older group): “[A]: the only problem I had with the touch screen is my nails. I have this problem at home, and that’s why I use a [brand name] pen because I find you have to develop a certain technique of touching. You can’t just go like that [action] because your nail would touch it and that doesn’t work so you have to slide off rather than...and I found that at home. But as I...”

The second sub-theme is the ‘use of hands’ whilst using the iPad. For instance:
The third sub-theme ‘speed’ is also shared by the older and younger groups. It reflects the speed participants thought they were supposed to go, or did go when using the iPad. For instance:

Speed (older): “[G]: we know ultimately what the tests are about and that’s cognitive impairment. [A]: or is it speed. [G]: I don’t think speed matters; it’s a balance between speed and accuracy. [M]: I think accuracy. [R]: there are lots of factors, there’s speed and accuracy. [R]: so how do you feel (J)? [J]: I would say about 85%, I think it was ok.”

Speed (younger): “[R]: so did you find the test enjoyable? [L]: in the beginning. [C]: yeah with my competitive edge to it. [L]: yeah I was a bit wish we was being timed and we knew we done. I got really competitive someone on to do better and faster but then other people will see that unhappy face and think ‘oh no’]. [P]: it put me off completely. [RB]: same [laughter]. I knew [researcher] was sat next to me and I didn’t want her to see the faces. [R]: do you think it would have made a difference if [researcher] was not in the room? [RB]: yeah, I didn’t want her to see it so I kept well at that angle she couldn’t have. [B]: it does show that the unhappy face does mean more”.

The sub-theme ‘tactic’ refers to the tactics both the older and younger groups had when completing the iPad test. For instance:

Tactic (older): “[JC]: I used the one finger all the time, I think I intuitively was picking out the first four numbers and then the other four. Also, I am very competitive, I was trying to go faster and faster so not much focus on being accurate so I had two errors.”

Tactic (younger): “[C]: yeah and also like how I went about it, like at the start I was just like looking 1, 2, 3, 4, as opposed to once I had an unhappy face it changed how I did it, like I was looking at groups so I would find 1, 2, then 3 and 4, then 5 and 6, and I found that I was quicker because it would take me an extra second to look but I tap quicker then because I already knew where the other one was. So I changed how I attended to it. [L]: changed your strategy. [C]: yeah”.

The final shared sub-theme is ‘performance feedback’ which relates to how much feedback they would ideally like to have had from performing the iPad test. For instance:

Performance feedback (older): “[N]: I have to say I would love to know how well I did. I would like to have some feedback on it. I think most of us who have done a test would like that. And what I assume is looking at how many mistakes someone makes is information I would like to have in feedback you know”.

Performance feedback (younger): “[R]: fab ok, how did you find it? [B]: same here yeah and then I got an unhappy face then all of a sudden I was like “wow slow down” [RB]: I didn’t get an unhappy face. [B]: I got two. [L]: I got two. [C]: I got two. [L]: but I think my finger accidently went too far next to the other ball, basically I shouldn’t have had the second unhappy face. [P]: do you want to appeal the judgement? [Laugh]. [L]: 1176 A. Jenkins et al. / Administering Cognitive Tests: Tablet Devices

Fig. 4. Test performance.
The problem ‘search strategy’ is unique to older adults. It reflects the strategies employed by younger participants to perform the MILO task:

"it depends on how you attend to the positions. For instance, whether you’re a linear searcher and look at the holistic picture and then if they were split between left and right I found it easier to go from one side of the whole task whether you’re a linear searcher or whether you look at the holistic picture and group the numbers just in a different location. Yeah I found it boring towards the end. [R]: yes and that is perfectly fine, I want you to be as honest as you can. Thank you [G]." The younger participants also expressed the test experience as positive, for instance, “[R]: fab, thank you. Did you enjoy doing the test? [A]: it makes me want one [iPad]. [P]: it was interesting but I wouldn’t use the word ‘enjoy’ [laughter] I was just counting dots but it was a little more engaging that some can be. However, others also deemed it to be ‘boring and distractible’, thus “[R]: ok, so would you say then something like that could be used on a regular basis or would you say no? [L]: I think it was boring”.

Feedback

In the MILO test, performance feedback was given in the form of an unhappy face icon when a mistake was made. However, we can see from the comments made in this study that in real life, rather than providing a potential learning opportunity, via feedback, such an icon can have a demoralizing effect, with evidence that an individual experiences embarrassment if an observer can see the unhappy faces, i.e., their poor performance. These factors may detrimentally affect test results and render the individual less likely to want to do the task again. Related to this was the finding that people could feel very self-conscious when being watched; again the presence or not of an observer may affect an individual’s test performance. A number of participants were embarrassed at the thought that the researcher present could see if they had an unhappy face pop up. Although this might not be of importance if the tests are self-administered, it is a pertinent consideration when administered by another individual.
feedback per se and how it is perceived. It is certainly the case that one would best engage with the tablet in an individualistic manner. Some said they would prefer to be most alert and attentive early in the morning, others late at night. Using this test in a clinical setting may require one to take into account the test users’ preferred time of day and the actual time of day. Real-life exceptions may exist where the individual is exceptionally tired and allowances be made. For example, if one becomes too practiced, then test scores could be invalid. Conversely, if the participant thought was most important despite clear instructions given prior to the start of the test, then it should be made a priority that they fully engage with the instruction process prior to the start of the test. The inclusion of a practice trial could be implemented in the future.

These issues seem to suggest that participants might have treated the test more like it was a video game as opposed to a cognitive test with an approach that involves strategizing to maximize the score they receive and possibly an increased sense of motivation or competitiveness with other players to get a “high score”. Researchers have not examined the attitudes and motivations of people who engage with cognitive testing, however, the motivations for video game play are quite well understood. Engagement with video games can be intrinsically motivating with reward derived from simple actions and immersion in game [39] or motivation can be derived from a sense of challenge or competition in the game and the accomplishment that come with it [40]. In conventional video games, these motivators can drive people to practice/play more and become extremely skilled with the games, improving their scores and their visuospatial awareness [41].

Time of day

In this study have highlighted several factors pertinent to the development of tablet or digital tests of attention and reaction time tests used in the assessment of cognitive processing. These factors may introduce bias, variability and lack of clarity. The reported heterogeneity in the second strategy, e.g., the use of one or two hands. It is important therefore that highly specific instructions are given to the individual’s choice and execution of each strategy. This is also a factor to consider, i.e., does not make clear. For instance, the level of education about the systems purpose, i.e., is it the speed or the accuracy of their performance which is most important? There was much disparity regarding what the participants thought was most important despite clear instructions given prior to the start of the test. Their lack of clarity could have been due to their preoccupation with the testing situation. If so, then it should be made a priority that they fully engage with the instruction process prior to the start of the test. Another factor which may detrimentally affect task validity is the individual adopt the same search and response strategy each time?, a factor which may detrimentally affect performance.

Factors such as arthritis or long fingernails may also affect task validity. It was also apparent that individuals in our focus groups definitely presented may influence performance. It is certainly the case that one would best engage with the tablet in an individualistic manner. Some said they would prefer to be most alert and attentive early in the morning, others late at night. Using this test in a clinical setting may require one to take into account the test users’ preferred time of day and the actual time of day. Real-life exceptions may exist where the individual is exceptionally tired and allowances be made. For example, if one becomes too practiced, then test scores could be invalid. Conversely, if the participant thought was most important despite clear instructions given prior to the start of the test, then it should be made a priority that they fully engage with the instruction process prior to the start of the test. The inclusion of a practice trial could be implemented in the future.

These issues seem to suggest that participants might have treated the test more like it was a video game as opposed to a cognitive test with an approach that involves strategizing to maximize the score they receive and possibly an increased sense of motivation or competitiveness with other players to get a “high score”. Researchers have not examined the attitudes and motivations of people who engage with cognitive testing, however, the motivations for video game play are quite well understood. Engagement with video games can be intrinsically motivating with reward derived from simple actions and immersion in game [39] or motivation can be derived from a sense of challenge or competition in the game and the accomplishment that come with it [40]. In conventional video games, these motivators can drive people to practice/play more and become extremely skilled with the games, improving their scores and their visuospatial awareness [41]. The questions this raises for the digital tests are first, whether the test motivates practice in the same way a game does, and second, whether this practice invalidates the test. For example, if one becomes too practiced, then test-performance ceiling effects can be induced.
that elderly people with dry or wrinkled fingertips had a significantly higher touch recognition error rate on some tablets. This could also be related with the layer types of the resistive touch-screen technology. Harada et al.’s [46] study also support dry-finger and users’ frustrations with unresponsive taps.

CONCLUSION

Arguably iPad-based tests may be an ideal base for home testing, with subsequent increased compliance in clinical trials, longitudinal clinical and research follow up, and the ability to signal deterioration and thus to facilitate intervention, but many factors need to be considered in their development if such tests are to reliable, valid, and objective. The participants in this study highlighted several issues pertinent to the development of tablet or mobile-based tests typical of those used in the assessment of cognitive function in older adults, which can then be used to inform more specific development for testing in individuals with cognitive impairment and dementia. In order to inform those considering developing tasks of RT and other aspects of cognitive function on touch screen based tablets, we summarize the information gained from our focus groups in the following section in a series of bullet points. It is clear from this information that many factors, which may not be currently taken into account when designing such tasks for use on touch screen tablets, but which, without being addressed could significantly influence task performance and thus adversely affect the clinical validity of such a test.

- Without highly specific instructions, response strategy to test components and stimuli can vary between individuals, despite clear instructions given.
- Users may adapt strategies, they may become a better match.
- As such, allowances need to be put in place for reliable, valid, and objective results.
- Arguably iPad-based tests may be an ideal base for home testing, with subsequent increased compliance in clinical trials, longitudinal clinical and research follow up, and the ability to signal deterioration and thus to facilitate intervention, but many factors need to be considered in their development if such tests are to reliable, valid, and objective.
- The participants in this study highlighted several issues pertinent to the development of tablet or mobile-based tests typical of those used in the assessment of cognitive function in older adults, which can then be used to inform more specific development for testing in individuals with cognitive impairment and dementia.

In a multi-directional tapping task on an Android tablet, Burkhard and Koch [45] asked 30 older adults to perform eleven single taps (eleven targets) in a specific order around the table. The authors used Fitts' Law in the assessment of cognitive function in older adults, which can then be used to inform more specific development for testing in individuals with cognitive impairment and dementia.

In this paper entered text on the mobile phone, which may also affect the physical ability to respond appropriately As such, allowances need to be put in place for reliable, valid, and objective results.

The physical challenges reported above are consistent with Weilenmann [42] in the context of mobile phones. The senior informants read text on the mobile phone, which relied on sequential pressing of keys within certain time-frames. Participants reported issues regarding rhythm of key-pressing: (1) Doing pressing was not a straightforward task, they tended to press too slowly or pressing one finger, or several fingers on the same or different hands, was common when participants were relying on the skin conductance of their fingers. This indicates the importance of considering when developing such a test.

Without highly specific instructions, response strategy to test components and stimuli can vary between individuals, despite clear instructions given.

- Without highly specific instructions, response strategy to test components and stimuli can vary between individuals, despite clear instructions given.
- The participants in this study highlighted several issues pertinent to the development of tablet or mobile-based tests typical of those used in the assessment of cognitive function in older adults, which can then be used to inform more specific development for testing in individuals with cognitive impairment and dementia.
likely that developers will need to take into account that feedback is not given in the form of simple actions and immersion is not derived from performing well or competition in the game and that comes with it [40]. In conventional tests, these motivators can drive people to practice more and become extremely skilled, improving their scores and their visual and dexterity. The questions this raises for self-administered cognitive tests are whether the practice in the same way a game does with similarity of a given game or the use of visual aids is of great importance when developing such tests, see also [42]. A suggestion from some of the participants was that the tablet should be placed in a tilted stand, and indeed spontaneous tried to hold it in this position so they could see the stimuli. However, although this position may ameliorate some physical difficulties, it is possible that it may affect performance in other ways as yet investigated and thus once again consistency of positioning would be highly important. The positioning of the tablet in relation to lighting in the room can also interfere with the ability to see the stimuli, thus lighting becomes an important consideration when selecting the testing environment.

There are of course limitations with our focus group study. For example, individuals living with dementia or cognitive impairment were not included, and it is possible that test administration, reaction to it, and performance varies with the integrity of cognitive function. Future studies should include a wider range of tests and their validation with other forms of computerized testing, groups representative of a wider range of age-related changes such as those found in relation to vision (such as cataracts, wearing glasses, color blindness), hearing, mobility and dexterity, memory function (what happens if individuals forget the instructions?), and levels of motivation and response confidence (e.g., examining the potential for guessing the response). Other pertinent factors for developers to consider in the future.
tion, whether individuals always use the same response strategy throughout the test, or further away to compensate for changes in their viewing distance and lighting, technical aspects such as the display and operating systems [11], the feasibility of using the internet to access the test or to view it [9], how used to using the internet a person is [9], how to ensure the person taking the test can understand the intrinsic design of the iPad can be varied [11]. Finally, it is important to consider whether a test to be included in routine use in research practice, the needs of the test (e.g., patient, clinicians, scientists, programmers/developers) need to be investigated and considered in the development stage of such tests with the development of quality criteria for the usability of such tests.

The results of this small study lead to the suggestion of such factors relevant to tablet-based tests of cognitive function. Future work will need to focus on better understanding the impact of physical challenges to technical familiarity as the number of older adults who regularly engage with such technology rises.


terms of verbalization, whether individuals always use the same response strategy throughout the test, or whether people use different strategies. Methodological considerations regarding the optimal performance such as fixed viewing distance (at an angle or flat on a table), vibrations and lighting, technical aspects such as screen resolution and operating systems [11], the feasibility of using the internet to access the test or view it [9], how used to using the internet a person is [9], how to ensure the person taking the test can understand the intrinsic design of the iPad can be varied [11]. Finally, it is important to consider whether a test to be included in routine use in research practice, the needs of the test (e.g., patient, clinicians, scientists, programmers/developers) need to be investigated and considered in the development stage of such tests with the development of quality criteria for the usability of such tests.

The results of this small study lead to the suggestion of such factors relevant to tablet-based tests of cognitive function. Future work will need to focus on better understanding the impact of physical challenges to technical familiarity as the number of older adults who regularly engage with such technology rises.

ACKNOWLEDGEMENTS


