

Reflections on design workshops with older adults for touchscreen mobile text entry

A. Komninou, E. Nicol and M.D. Dunlop

Mobiquitous Lab, Computer and Information Sciences,
University of Strathclyde, Glasgow, Scotland, UK.

Abstract. *On touchscreen devices such as smartphones and tablets, text entry remains key to many tasks and is an important factor in the usability of such devices. The physical and cognitive issues associated with ageing can make the known problems of text entry particularly acute for older adults. Poor usability can present a significant problem for older adults where accessing services and social activities are concerned, both of which have implications for exclusion. In a study of mobile text entry where we aimed to develop novel keyboard layouts to address the particular requirements of this group of users we employed a variety of participatory design techniques. We report on our experiences from employing these methods and the methodological implications for further research in this area.*

Keywords: Older Adults, Participatory Design, Mobile Text Entry.

1 Introduction and Background

Text entry is still core to mobile interactions such as email and web searching. Most smartphones now do not have a physical keyboard, relying on on-screen touch keyboards. These are slower and more error prone than traditional mini-physical keyboards (e.g. [1, 2]) but are popular as they permit full screen services and larger reading area. While there have been many studies of text entry on touchscreens, there has been very little work studying the effects of ageing on text entry, and particularly on modern touchscreen phones where reduced visual acuity, motor control and working memory are likely to have more of an impact. Key recommendations by Fisk et al. [3] for designing for this age group are at a complete contrast with typical smartphone touchscreen interactions where, for example, soft buttons require almost zero force to activate and have zero travel, soft keyboards have no gap between active key regions and modern phones have surfaces that are high gloss, low friction, glass with interfaces that often have reduced contrast for styling – a problem that is exacerbated by sunlight and reflections. Taken together this means that natural ageing processes can make smartphone interaction particularly challenging for older adults.

While industry is focussed on targeting younger users, our project aims to investigate text entry with older users. Our initial studies showed that older users may actually be more willing to adopt new keyboard layouts than younger users, who appear more attached to soft-screen QWERTY. The results reported here were obtained during the early stages of a two-year investigation of text entry for older

adults. Using participatory design approaches we conducted research workshops with older adults to identify key design criteria and parameters for improved text entry. Our project aims to develop innovative keyboard designs that are more suited to older adults, e.g. by reducing input errors and disrupting flow less, and to quantitatively measure touchscreen target acquisition times for this age group, a key factor for text entry interaction modelling and interface design. This paper focusses on our workshops and lessons that can be learned both for the focussed area of text entry on mobiles and the more general domain of collaborative design with older adults.

2 Relevant Work

2.1 Designing with Older Adults

Older adults are traditionally stereotyped as technology-wary and generally disengaged from advancements in the integration of technology in everyday lives. Lindsay et al. [4] and Coleman et al. [5], found through experience in working with older adults, that they do not conform to the ageist stereotypes that designers often have in mind. They are keen to explore and use new technology, however, they differ from younger users in their decisions on accepting or rejecting new technology. For them, the dominant criterion is the technology's perceived usefulness to their daily life, rather than fashion or trends that strongly affect younger users. They also found that older users were particularly keen on the idea of the "disappearing computer" and specifically welcomed technology that looks more like artefacts they already play a role in their lives.

To design systems for use by older adults, Newell et al. [6] argued that the designers and developers of systems need to interact more closely with older users, as there is a distinct inability of the former to appreciate the issues that older users face without direct observation. Coleman et al. [5] also argue that low adoption of technology by older adults can be strongly linked to "*the fact that the push for new technology comes from developers, not from the end-users who may ultimately not be interested in the discussions or the outcomes*". They argue that involvement of older users in the conceptual and very early stages of an idea is highly desirable. Antona et al. [7] surveyed a variety of established design techniques and commented on their suitability. Of 12 techniques reviewed, only 5 were deemed appropriate without modification: direct observation, activity diaries & cultural probes, scenarios & personas, prototyping and art-based approaches. A further 6, including some of the most widespread methods, were deemed applicable only with modification: brainstorming, survey & questionnaires, interviews, group discussions, user trials and cooperative & participatory design. One technique, empathic modelling, was deemed unsuitable. Several empirically derived observations can be found to justify the need for these modifications to well known techniques. Here, we classify these factors into four categories: *physiological, psychological, cognitive* and *societal* (Table 1).

Older adults can be affected by any combination of these factors, making the acquisition of a "representative sample" a difficult, if not impossible, exercise. Hence design should proceed with a sample that is representative of the characteristics that

Table 1. Factors that influence the use of design methods with older adults.

Factor Category	Issues
<i>Physiological</i>	Age factors that make self-reporting inaccurate Limited endurance Medical conditions that hinder motor skills, hearing or verbal expression
<i>Psychological</i>	Tendency of blaming themselves instead of designers for issues Fragility of confidence while using technology Anxiety towards computer use Perception that computers are not much use to them Difficulty in focusing on the design process if they feel that it is going towards a direction that is not valuable to them
<i>Cognitive</i>	Lack of understanding of technical language and metaphors Lack of underlying understanding of computer concepts Difficulty in envisaging new technology Disapproval of deep explorations in subjects that are forced on them by the designer
<i>Societal</i>	Tendency to diverge into unrelated subjects during discussions Participatory design meetings are seen as social events Positive predisposition towards prototypes and tendency to praise rather than offend researchers by offering objective views

researchers deem important. With these factors in mind, several guidelines on how to employ design tools with older adult participants can be retrieved from literature.

User Trials (lab): They should be conducted in a supportive environment where the designer is able to interact with participants [6]. Participants should not be left unattended. This requirement dictates that additional measures are used during observation, e.g. recording the level of assistance required by the participant [6].

Participatory Design, Interviews and Group Discussions: With such tools, hands-on approaches work best and are more valuable than relying on participant descriptions. To ensure that participants and designers establish a mutual understanding, the language and terminology used must be “compatible”, in other words, technical language must be avoided and questions should be short and simple, with wording that participants can understand [8]. One way to avoid such issues was recommended by Göllner et al [9], who suggested the use of metaphorical tools to alleviate issues of using terminology (e.g. a “carrier pigeon” that carries a message to another device, instead of “sending via Bluetooth”, or a “blackboard” to display a message instead of a “pervasive display”). Such sessions should start with a general introduction to the topic area that should be followed by general questions, before the researcher delves into more specific and focused questions [4]. Amongst other lessons they also highlighted that support must be given to participants during tasks that require creative thinking for envisioning new technology or intangible concepts [4].

Survey and Questionnaires: Newell et al. [7] report that older adult self-reporting on questionnaires is fundamentally different to younger adults: e.g. their responses include more “don’t know” options and less range extremes. Older adults also tend to need higher levels of certainty before responding to questions, hence it is

recommended that questionnaires are administered directly, or during an interview by the researchers, to ensure that participants are given the confidence to respond.

2.2 Use of Touch Screens by Older Adults

Touchscreens are a core feature of the majority of modern mobile devices, including basic mobile phones, smartphones and tablets. Hence their adoption by older adults is high for many of the same reasons as the general population: e.g. features, high affordability, and societal factors. Coleman et al. [5] present cases where technology can act as a mediator to connect older adults with their family or grandchildren and situations where uninterested older adults had been given a modern mobile phone as a gift by their children.

Murata et al. [10] investigated suitability of touch screens for older adults and showed that the time required to point to a target with a touch interface was similar for younger and older adults, in contrast to slower performance by older adults with other pointing device, such as a mouse. Similar findings were reported by Findlater et al. [11] investigating gestures performed by mouse or touchscreen, where it was found that older adults benefited greatly from interaction with a touchscreen. Murata et al. [10] also concluded that these benefits for older adults may only be applicable to targets larger than 50x50 pixels, as a smaller target size led to high error rates. Jin et al. [12] found that manual dexterity can affect speed and accuracy when a target is embedded within a row of adjacent buttons. They found that older adults performed optimally with a button size of 19 mm², although their expressed preference was for buttons that were large but not overly so (between 16.5 and 19 mm²). Interestingly they also found that larger spacing had a negative impact on reaction times and recommend adjacent buttons are spaced between 3.2 and 12.7 mm apart and optimally at 6.4 mm. Zero spacing did not affect response time, but participants showed the lowest accuracy and preference with this set-up. Chung et al. [13] found touch-based ATM keypads to be quicker and more comfortable for use by older adults than their physical counterparts. The relatively improved position for touchscreen use by older adults may be due to touch-screens requiring less physical effort and thus being more suitable for users with low muscular strength or motor impairments [14]. The researchers also argue that touch interaction errors can be mitigated partly by the provision of visual, auditory and tactile feedback, something that was also supported by Bonner et al. [15], but a combination of feedback modalities is desirable, as users may suffer from sensory impairment(s).

Nicolau and Jorge [16] examined the role of hand tremor in text input by older adults on smartphones and tablets. Amongst other findings, they showed that input speed and accuracy was better on tablets (attributed to the larger key size) and that speed was dependent on familiarity with QWERTY, but this did not affect accuracy. Hand tremor was also found to be linked to errors in mobile devices, but not so in tablets. In terms of the errors that older adults make, omission errors are the prevalent category and were attributed to cognition issues. Substitution and insertion errors were next to follow and were attributed mostly to poor aiming rather than finger slips for the former, and accidental touches for the latter. Most recently, Harada et al. [17] found that unexpected touch screen responses (either unregistered touches or

unintentional touches) are a major cause of frustration for older users and are not uncommon. Many of the unintentional touches occur due to the way the device is held and also because most modern screens support multi-touch. They also highlighted a problem with continuous focus on the keyboard instead of the text being input was a particular issue for older adults, who were observed to continuously type until they finished, before going back to check what has been entered for mistakes.

2.3 Conclusions from Literature

Our first finding from reviewing the relevant literature is that so far, a complete methodology for use with older adults has not been proposed or evaluated by researchers. Antona et al. [4] provided an interesting overview but did not proceed to piece together the necessary modifications of current design methods that they suggest. Our contribution is to bring together in one place justifications for modification to methods alongside a set of guidelines from the literature as to how to use the methods. Later, we will comment on their validity and add to these guidelines from our own experience, in the hope of formulating a more thorough methodology which can later be used by other researchers. In terms of our narrow field of study, i.e. text entry with touch-screens for older adults, we find that touchscreens are good for this age group because they require less manual dexterity and force than traditional keypads. However, target size and spacing is an issue particularly for text entry. Errors in text entry are unavoidable due to the particularly bad combination of multi-touch and physiological conditions. The work of Harada et al. [17] shows that perhaps what is most sorely needed, apart from algorithms for ironing out erroneous input, is support for older adults to be able to detect and manage mistakes during input review (i.e. after input has finished).

3 Studies

Following previous research with older adults who take classes at the university we established a good relationship with the Centre for Lifelong Learning (CLL) who organise learning activities on their Learning in Later Life (3Ls) programme. CLL staff allowed us to use their premises for studies, which we hoped would provide a familiar and supportive environment for participants [6]. They assisted with recruitment by allowing us to put up posters and by sending emails to their 2000 students. A majority of the people on the 3Ls programme are retired but there are also many who are in full or part-time employment, and many who do voluntary work. Many students attend the special interest clubs that have grown up around the classes and are enthusiastic participants in their educational and social activities. We acknowledge that there is an element of bias in our study that has arisen from the reliance on the CLL for recruitment given that the typical student who uses the CLL is an active, mobile older adult often from a skilled or professional background.

Table 2: Most popular locations for activities in mapping exercise

Location	Sticker Count
Train	23
Home	18
College	17
Café	14
Bus	11
Pub	11
Shop	9
Hospital	9

Activity	Sticker Count
Web searches	40
Sending a message	36
Making a call	31
Using social media	30
Creating a document	23
Other use	21

Cultural Probes: At the close of the first session participants were given cultural probe packs [18.] We allowed them time to explore the packs and discuss the contents with their groups before we explained how we wanted the materials to be used. Each pack contained five postcards. Four postcards had categories to be completed and were printed with the wording: *People contacted e.g. my brother or dentist; Sites/Applications used e.g. Google search or Facebook update or Text message; Problems experienced (Personally and that others might e.g. hard to read screen or buttons too wee).* The other postcard had the wording: *Desires for better mobile experience. What would make it easier? More fun? Quicker? Smoother?* In the packs were disposable cameras with instructions to take photos of: *Places where you entered text on your mobile device; The screen while entering text (remembering to be careful about privacy); Things going well; Things going poorly.* Participants were given a postage paid envelope and asked to post all materials to us a week after the session. With the exception of three participants who completed fewer postcards, all participants completed five postcards on time. Several mobile and text entry issues were uncovered. Firstly, we found frustration over the lack of support: in the absence of manuals users struggled to understand what their phones were capable of doing. They were also uncertain about using discussion forums etc. rather than the more traditional authoritative publications from manufacturers. Many commented that on-screen buttons, and in particular, keyboard buttons were too small to hit accurately. This was compounded by the extreme sensitivity of modern capacitive touchscreens and led to many problems with both text entry and accidentally tapping buttons (including soft home keys and premature sending of messages). All cameras were returned to us with between three and twenty photos taken. Some were of locations, of the participants or family members. Others showed aspects of the device screen they were using, illustrating the activity or problem occurring. Many photos were of poor quality, particularly those of devices - unfortunately, none of these provided a clear image of what was happening on the screen (often due to the effects of flash).

3.1.2 Session 2: Disambiguation of Text Entry Context

In preparation for our second session, we summarised the postcard data in tag-clouds (e.g. Fig 2) under 4 themes we identified as emergent: Desires, Problems, What (i.e. task for which the device was used), Who (i.e. who they were interacting with). Ahead of the second session, participants were assigned to groups that reflected the priorities expressed in their postcards. Participants were unaware of the reasons for their new groupings. We proceeded to perform the following activities:

Issue Identification: We began by thanking them for their submissions and showed them the tag-clouds for the 4 themes that emerged from their postcards. They were very interested in these. Presenting participants with the results of the research conducted with them has often been shown to be a fruitful part of the design process and so it was here. We made it clear that some issues they highlighted were beyond our control (e.g. battery life) and while acknowledging the importance and impact of these problems, we encouraged participants to focus on issues specifically concerned with text entry methods that they used, rather than issues beyond our control.

Photo organisation: We returned the photos to the participants and asked them to share and explain them to others in their groups. They were asked to organise the group's photos in themes of their own choosing and, using a post-it note to annotate photos where the activity required further explanation than was available from the image alone, which many of them did. Two groups did their photo organisation using the place or type of place where the interaction occurred as a category in which to group the photos. The third group based groupings on the type of task, e.g. text message, email. With the latter approach, there were fewer categories and a large pile of photos was created for each. We noticed that not all photos were included in groupings with participants retaining them in their own "private" bundle. The exercise was reasonably successful with participants mostly keen to share and discuss their photos. Coherent categories emerged consistent with those expressed in the postcards.

3.2 Observational studies and exploring alternatives

We explored some of the themes that had emerged from the initial workshops in an



Fig 2: Tag clouds of postcard responses

observational study. We were particularly interested in the issues of screen sensitivity that had been mentioned by 6 (i.e. nearly half) of the participants and were keen to understand better exactly what was happening when participants encountered this type of problem. To address these concerns, we organised a further session, inviting these 6 participants back to examine their text entry in more detail via observation and logging. Five (4M aged 61-70, 1F aged 68) agreed to participate. The following activities took place:

Observational studies: Participants did a series of text entry tasks using their own phones and we videoed their performance of these for later analysis. The tasks were a mixture of *copy tasks* where an onscreen phrase was to be copied, and *description tasks* where the participant was shown a set of images and asked to write a story about the people in the images. All participants used QWERTY for these tasks. Later in the session we explored alternative text entry methods. We followed up these focussed studies in subsequent weeks with individual studies to record in detail users typing behaviour using *copy tasks*, *description tasks* and *compose tasks* as per Vertanen et al. [19].

Structured interviews: We worked with pairs of participants asking them to try several input methods on an Android phone: an emulation of a physical 12 key phone, a 20 key keyboard and Swype. We conducted semi-structured interviews with each pair. None had seen any of the input methods before. Swype attracted the most favourable comments, the 12 key the least, being thought the most complicated to use. Participants thought it would take them over a week to learn the 12 key, up to a week for the 20 key and between a day and a week for Swype. None would be happy to give up their current keyboard for the 12 key, 2 would be happy to adopt the 20 key and 4 said they would give up their current input method for Swype. Participants were keen to try the new keyboards and were good at giving their opinions in interviews.

3.3 Exploring Alternatives and Paper Prototyping

Using Metaphor: In our next research sessions we investigated error correction and our participants' use of and attitudes to the associated issues. We met with participants (10F aged 53-73, 5M aged 55-70) in groups of 4-6 over 3 different sessions, one per group. We explained the idea of a "*blooper*": a term used in Hollywood for things that go wrong while filming, in a metaphorical approach similar to that described by Göllner [9]. We appropriated this term as a way of describing "things that go wrong" with text entry. We wanted to avoid the language of "mistakes" or "errors" to avoid implying blame on either participants or devices, and to address the twofold aspect of bloopers. Firstly, the input error by the user, caused by tapping the wrong key for example, and secondly the system's response to the error (and whether the response causes an unwanted word). We asked participants to write post-its describing typical things that go wrong for them when entering text. We then asked them to group the post-its into themes. In a similar exercise to that described in 3.1.1, participants were given 3 postcards to document bloopers for the week following the session. We moved on to talking about how to provide support to avoid and correct these problems, calling this "*blooper support*". We showed the groups 4 possible options: Autocorrect (on Apple devices), Ginger (an Android

keyboard that shows corrections for all incorrect words as the user types), the current Microsoft Word spellchecker and the modal spellchecker from an early 90s version of Word. After demos of each approach we asked 6 pairs and one group of 3 participants to evaluate each of the methods. Ginger was preferred by the biggest number of pairs (4) followed jointly by Autocorrect(2) and Word (1), with classic Word in last place (0).

Paper prototyping: We brought the three groups from the earlier blooper sessions together for a paper prototyping exercise on blooper support. As in previous sessions we made tag-clouds from the data and presented these at the start of the session. In groups of 4, pre-selected by us to reflect a diversity of phone ownership and experience we asked participants to take part in a *warm up activity* to familiarise them with the method and to get to know others in the group. We asked them to design an interface for a meal-sharing app then to present their ideas. Designs were to be drawn on paper facsimiles of mobile phone interfaces with blank spaces for participants to make drawings and annotations. They did this without the need for further instruction from us. We moved on to using the same method to design for ideal error correction. We reminded them of the concept of bloopers outlined in the earlier session and each group was allocated 2 tasks out of a possible 4. Task 1 is given as an example of the type here: *These suggestions are good, but imagine they were not. How could this be adapted to allow it to make alternative suggestions and allow you to pick them? Sketch your ideas.* Each group worked on their tasks then presented their ideas to the rest of those assembled. A design resulted for every task (e.g. Fig.3).

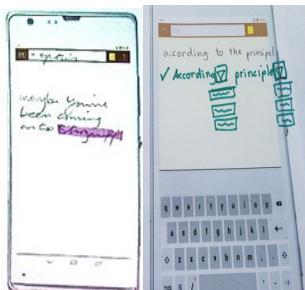


Fig 3: Designs for error correction produced in the prototyping exercise

In only one case the design produced was not drawn but was presented as a flow chart with written descriptions of how the interactions would take place. We think that the fact that 7 out of 8 tasks produced a design is a sign of the success of the method and that the warm up task succeeded in familiarising groups with the method.

4 Methodological Lessons

4.1 Recruiting and keeping participants

In a 2 year project requiring multiple user studies, recruitment and retention of participants was extremely important. We realise that without the assistance and

proximity of the CLL, recruitment would have been extremely difficult. We recognised the need to show our appreciation to participants and to make them feel like part of a valued ongoing team rather than just participants in a one-off operation. We offered remuneration to participants in the form of a shopping voucher, which we described as a thank you rather than a payment. We have since been told by several participants that they had not expected to be rewarded in this way and have received many messages of thanks. At all sessions we offered refreshments, which were enjoyed as a group activity in themselves. This experience confirms suggestions in [3] and [5] about session duration and the importance of breaks, as well as the perceived social aspects of participating in research [3]. We placed an emphasis on politeness and friendliness in our interactions with our participants, taking care to learn the names that people preferred to be known by, introducing them to people they had not previously met, making a point of greeting everyone who entered a session and remaining on hand to chat at the end of each session. We endeavoured to be quick to respond to correspondence received from existing and potential participants in order to gain and maintain respect and trust. We have taken some encouragement from the fact that the rate of return to our sessions has been almost 100%. Participants have shown a willingness to participate multiple times and have expressed disappointment when another engagement has prevented them from taking part.

4.2 Choosing tasks for studies

All studies reinforced the importance of providing clear and consistent instructions for tasks and the need to keep participants focussed, confirming guidelines in [1, 3, 5]. Participants, perhaps understandably, often became preoccupied with exploring and discussing the materials they had been given rather than paying attention to the instructions from the researchers. In the mapping exercise in particular, each group performed the task differently from the others, due perhaps to a lack of clarity from the researchers combined with a lack of attention to the instructions from the participants. So the results are not directly comparable but all are interesting. One method that we found not to be particularly useful was the photo exercise that participants did as part of their homework, negating the recommendation that cultural probes do not require modification to use [4]. The photos were not useful for learning about mobile use and activity. We thought the photos might have provided insights into participants' daily lives but none rose much above the banal and the photographs of screens were of poor quality, due to the effects of flash and that these cameras are unsuitable for close ups. Additionally, participants seemed much less keen to share their photos with other participants than they were to share their postcards.

4.3 Agreeing terminology

In any design exercise involving non-expert designers such as our participants, agreeing on common terminology to use when talking about technology or the tasks that one can perform using it can be key to the success of the exercise [5]. We found this was even the case when we reflected on our non-tech paper and pen exercises.

The mapping task led to arguments over what was a mobile, and a very in-depth discussion about what constituted a document. This shows that while approaches like [6] might alleviate some terminology issues, these issues can extend further than concepts related to technology but may be central to the subject under development. Such issues as did detract from performance of the activities but gave us insights into the motivations of the participants and their attitudes to technology and its place in their lives. When we designed the mapping exercise, we thought about places where our participants might consider certain mobile activities to be appropriate but we had not expected this to become the main topic of discussion in at least one of the groups and for it to influence so strongly their performance of the task. For this group the activity became about “*where would it be ok to do X*” rather than “*where do I do X*”, which is really what we’d had in mind.

4.4 Explaining where participants fit in the big picture

At the start of each session we showed a PowerPoint slide about where we had reached in the design process and re-stated the project’s aims and the design methods used. This was important because some participants had not been with us for every session, and was crucial to reminding them of the purpose of the research and their place within it. At every stage we explained what we would do next whether alone as researchers or together with them at the next session. We did this with a view to making them feel important, included and valued in the design process, as recommended in [4].

Some participants were openly sceptical of the methods we employed, particularly in the early sessions with activities such as mapping. This scepticism was much more pronounced than we have found with the participants with whom we typically conduct experiments, who are, for the most part, undergraduates. This scepticism may stem in part from surprise at being asked to carry out low-tech exercises when an exercise involving technology may have been expected. Perhaps older generations are less familiar with the sort of techniques that have become commonplace in schools, university and on workplace training courses which often involve such exercises. This has shown us the importance of managing participants’ expectations before and between sessions and of the need to provide feedback following each exercise, showing the participants what we learned from a task, that at the time may have seemed arbitrary or far removed from the overall design goal. Having a warm-up task such as that used in the paper prototyping exercise provided an effective of countering scepticism about the usefulness of the method, before the task proper was carried out. The warm-up task also addressed the fears of some participants about their ability to build prototypes, before the task proper was carried out.

4.5 Cross cutting lessons

On several occasions we gave participants homework to do between sessions. In all cases that homework was completed and in the majority of cases responses were pertinent to the questions asked and the topics that we asked them to think about. We

received a few emails asking us to clarify some aspects (e.g. wording on postcards), but the majority did tasks without requiring further instruction. Homework allowed us to uncover what was important to participants and to develop themes to focus on in future sessions. Responses received reinforced the need to develop a common language to describe concepts to participants. Arguably, doing homework meant that participants could keep the research in mind in the time between sessions, so that when they returned they still had a better chance of remembering the purpose of the exercise. The homework and artefacts produced were a good way to provide continuity between sessions. While some participants said they felt no need for improvement to their devices, we encountered enthusiasm from most participants at the prospect of new designs. The examples of novel input approaches that we introduced were met with interest. A majority of participants are open to trying new keyboard designs and, without being asked to, a few have begun using alternative methods such as Swype as their main input method since being involved in the project. The willingness to adopt something new bodes well for our plan to conduct longitudinal studies with prototype keyboards.

The importance of the social aspect has come to the fore. Participants told us that learning, while not quite secondary to the social activity associated with classes is certainly not the only reason for enrolling. Some of the same motivations explain the interest in our studies. While we were oversubscribed for group sessions, it was difficult to recruit for sessions where tasks were done individually. Those arriving for individual sessions were disappointed to find none of their peers there and not just because of the desire for informal socialising: participants preferred the more social group tasks. We observed peer tutoring such as participants giving tips, and training each other on a variety of mobile issues. Participants gave researchers tips about using mobile devices that were previously unknown to them. These activities did not however appear to distract the participants from the task we had set them. Allowing a degree of socialising in sessions is also helpful for retention, as participants begin to extract value from participation that greatly exceeds the remuneration.

5 Direct Lessons

The long-term goal of our project is to develop new approaches for text entry on smartphones that are suited to older adults. As we have developed the methodology, we learned some direct lessons about interface design for older adults.

Heterogeneity of devices, approaches and uses: A strong outcome is the diversity of text entry activities conducted by older adults. Many participants have been clear about adopting different approaches to text entry, error correction and grammatical correctness depending both on the recipient and mode of communication. One participant stated *“On my twitter posts I am very careful. Because of the character limit, I work to fine tune my posts. On Facebook, I have also learned to be careful after some posts were posted and taken out of context. But, in SMS and personal email I am much more casual about errors as I am having a conversation with people I know”*. Others agreed and added that some friends/relatives use abbreviations and “textspeak” while others use more complete

English. Our participants tend to adapt their style to match their conversation partner. Some participants were notably adept at swapping between devices depending on circumstance: using tablets, smartphones, or laptops for different situations. In the mapping exercise, we observed a high level of social awareness of the impact of devices on others: places such as cinemas were considered inappropriate for device use. Other locations were unexpected hotspots, e.g. the hospital: a participant said she wrote notes and questions for her appointment in advance, checked these during the appointment and added notes during and after the appointment. As with many smartphone users, our group were split on app use with many restricting themselves to apps that came with their phones rather than exploring app stores.

Responsibility for errors: Our users were sceptical of the need for error correction. As is commonly seen with users of poorer interfaces, users tended to blame themselves, as observed in [6]: as one user stated, “*it’s just fixing my mistakes*”. Despite an overall acceptance that buttons were too small and too sensitive earlier in the project, users in our error correction workshops still tended to feel that managing spelling and grammar was their own responsibility and that they should learn to type better rather than relying on spell correction. As such, we met with surprising scepticism that error correction should be a focus of our work. One promising avenue, however, particularly when considering that older users tend to focus on typing before going into a “review” mode of what they typed [17], is replacing or augmenting some forms of automatic error correction with better support for reviewing. This appears to fit better with user desire for less *predictive* entry but would still support sending high quality error-free messages as required. In the workshop where we showed users different approaches to error correction there was a strong preference for in-line style review mode similar to red-underlining in word processors (e.g. using different colour highlighting for spelling mistakes, autocorrected and autocompleted words).

Dislike of prediction: Pre-project workshops identified predictive text as an issue with our older adult participants. This was confirmed in both the postcard exercises and when showing alternative keyboards to users. The reasons are not clear and require further investigation: the visual noise of rapidly changing suggestions, low quality “random” suggestions and a dislike of early predictive text (e.g. the original 12-key T9) are all possible explanations. However, the groups tend to agree on disliking forms of predictive texting.

Support: User support and documentation have also emerged as important with users left unaware of the full capabilities of their devices and who to call when things inevitably go wrong.

6 Future work and conclusions

In doing this research we became more aware than ever that the older adult population is a diverse group, who engage in a wide range of highly interactive mobile activities and who, while keen to question methodologies and technologies, are open to new approaches. Future work will involve several longitudinal studies complemented by short one-off lab studies of text entry with this user group and a comparative study with younger users. One study will investigate typical daily input errors using a

standard QWERTY Android keyboard with logging capability and later, we plan to investigate long-term use of our review mode-supporting prototype keyboards. In the studies to date we have found it difficult to recruit for those sessions where tasks were done on an individual basis and where the group social aspect was missing. This is a challenge for us to address. In literature, the tendency of older users to see participation as a social event is treated as a pitfall, however, we believe that there are ways that this characteristic can be used to the researcher's advantage. We are considering a modified approach where we carry out text entry studies in group sessions, allowing participants to mix individual and group activities, or to share results of individual activities carried out with sessions with other peers, to keep them engaged by exploiting the social aspect of participation.

This paper has discussed the outcomes from a series of workshops with older adults. In these works we brought together a range of techniques from the literature to conduct an in-depth study of mobile text entry for older adults. We set out to learn about how older adults use text entry on their mobiles and what problems they face. All the methods, with the possible exception of the photographic element of cultural probes, were successful in giving us insights into usage patterns and problems. Methodologically we learned, in line with literature, not to underestimate the social aspect of group-based activities, to take care in wording of activities, to set expectations appropriately and to take great care in treating our participants politely and keeping them informed so that they felt part of the on-going research project.

7 References

1. Hoggan, E., Brewster, S. A., Johnston, J.: Investigating the effectiveness of tactile feedback for mobile touchscreens. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 1573--1582. ACM (2008)
2. Lee, S., Zhai, S.: The performance of touch screen soft buttons. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 309--318. ACM (2009)
3. Fisk, A. D., Rogers, W. A., Charness, N., Czaja, S. J., Sharit, J. Designing for Older Adults: Principles and Creative Human Factors Approaches. CRC Press (2009).
4. Lindsay, S., Jackson, D., Schofield, G., Olivier, P.: Engaging older people using participatory design. In: Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems, pp. 1199--1208. ACM (2012)
5. Coleman, G. W., Gibson, L., Hanson, V. L., Bobrowicz, A., McKay, A.: Engaging the disengaged: how do we design technology for digitally excluded older adults?. In: Proceedings of the 8th ACM Conference on Designing Interactive Systems, pp. 175--178. ACM (2010)
6. Newell, A., Arnott, J., Carmichael, A., Morgan, M.: Methodologies for involving older adults in the design process. In: Universal Access in Human Computer Interaction. Coping with Diversity, pp. 982--989. Springer Berlin Heidelberg (2007).
7. Antona, M., Ntoa, S., Adami, I., Stephanidis, C.: User Requirements Elicitation for Universal Access. In: The universal access handbook, Stephanidis, C. (Ed.), pp. 15-1--15-14, CRC Press (2009)
8. Barrett, J., Kirk, S.: Running focus groups with elderly and disabled elderly participants. Applied ergonomics, 31(6), pp.621--629 (2000).

9. Göllner, S., Lindenberg, J., Conradie, P., Le, J., Sametingner, F.: The Enchanted Neighborhood: Using Metaphorical Devices for the Inclusion of Seniors in the Co-designing Process. In: Proceedings of the 3rd International Conference for Universal Design, pp. 1—10. IAUD (2010)
 10. Murata, A., Iwase, H.: Usability of touch-panel interfaces for older adults. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 47(4), pp. 767—776 (2005)
 11. Findlater, L., Froehlich, J. E., Fattal, K., Wobbrock, J. O., Dastyar, T.: Age-related differences in performance with touchscreens compared to traditional mouse input. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 343--346. ACM (2013)
 12. Jin, Z. X., Plocher, T., Kiff, L.: Touch screen user interfaces for older adults: button size and spacing. In: *Universal Access in Human Computer Interaction: Coping with Diversity*, pp. 933--941. Springer Berlin Heidelberg (2007)
 13. Chung, M. K., Kim, D., Na, S., Lee, D.: Usability evaluation of numeric entry tasks on keypad type and age. *International Journal of Industrial Ergonomics*, 40(1), pp. 97--105 (2010)
 14. Caprani, N. O'Connor, N. E., Gurrin, C.: Touch Screens for the Older User, In: *Assistive Technologies*, Dr. Fernando Auat Cheein (Ed.), pp.95-118, InTech (2012)
 15. Bonner, M.N., Brudvik, J.T., Abowd, G.D., Edwards, W.: No-Look Notes: Accessible Eyes-Free Multi-Touch Text Entry. In: *Pervasive Computing 2010, LNCS*, vol. 6030, pp. 409--426, Springer Berlin Heidelberg (2010)
 16. Nicolau, H., Jorge, J.: Elderly text-entry performance on touchscreens. In: Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility, pp. 127--134. ACM (2012)
 17. Harada, S., Sato, D., Takagi, H., Asakawa, C.: Characteristics of Elderly User Behavior on Mobile Multi-touch Devices. In: *Human-Computer Interaction—INTERACT 2013, LNCS*, vol. 8120, pp. 323--341. Springer Berlin Heidelberg (2013)
 18. Gaver, B., Dunne, T., Pacenti, E.: Design: cultural probes. *Interactions*,6(1), pp.21—29. ACM (1999)
 19. Vertanen, K., Kristensson, P: Complementing text entry evaluations with a composition task. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 21(2), pp. 8
-

[KA1]Lesson: Nobody has yet suggested a complete methodology for use with older adults. Antona et al have made comments but did not proceed to piece together the necessary modifications that they suggest. Our contribution here is to bring together in one place the justifications for modifications to the methods alongside a set of guidelines from the literature as to how to use these methods. Later, we will comment on their validity and add to these guidelines from our own experience.

[KA2]Lesson: Touchscreens are good because they require less manual dexterity and force than traditional keypads. However, target size and spacing is an issue particularly for text entry. Errors in text entry are unavoidable due to the particularly bad combination of multitouch and physiological conditions. Harada's work shows that perhaps what is most sorely needed, apart from algorithms for ironing out erroneous input, is support for detecting and managing mistakes during review (after input has finished).