

# What's on Your Mind?: Investigating Recommendations for Inclusive Social Networking and Older Adults

Chris Norval<sup>1</sup>, John Arnott<sup>1</sup>

<sup>1</sup>School of Computing  
University of Dundee  
Dundee, Scotland, UK, DD1 4HN  
{chrisnorval, jarnott}@computing.dundee.ac.uk

Vicki L. Hanson<sup>2,1</sup>

<sup>2</sup>GCCIS  
Rochester Institute of Technology  
Rochester, New York, USA  
vlh@acm.org

## ABSTRACT

Social networking sites (SNSs) are becoming increasingly popular as a method for social interaction. While research has reported benefits associated with components of SNS usage, a digital divide has emerged between younger and older users. SNSs can be useful for communicating with family members and helping one feel digitally included; however, there are a wide range of reasons why many older adults choose not to use this kind of technology. We present a series of user studies investigating the barriers and challenges that SNSs can present to older users. These user studies led to the derivation of user recommendations to mitigate these barriers. The recommendations were then evaluated within a comparative evaluation which involved 25 older adults completing tasks on two interface versions of a simulation SNS. We present the recommendations and the methods of their creation and evaluation. Implications for developers of SNSs are discussed.

## Author Keywords

Social networking sites; older adults; recommendations; inclusive design; comparative evaluation

## ACM Classification Keywords

H.5.3. [User Interfaces]: User-centered design; Prototyping; Screen design (e.g., text, graphics, color)

## General Terms

Human Factors

## INTRODUCTION

Social Networking Sites (SNSs) have grown in popularity in recent years. Facebook, for example, currently has over 1.1 billion monthly active users<sup>1</sup>. Both in the UK and US it was reported that between 64% and 72% of online adults use SNSs [3, 18].

<sup>1</sup><https://newsroom.fb.com/Key-Facts>

Researchers from different fields have investigated reasons and implications for this rise in popularity. Such research has suggested that intensity of Facebook usage is positively associated with life satisfaction and social trust [21], and the use of some Facebook components are associated with increased social capital [5, 21] and reduced loneliness [5].

Additionally, researchers have found that SNSs can be a beneficial online platform for older adults as a communication platform with family members [1, 8, 9, 14, 17, 20]. Using social media, where, for example, children or grandchildren live far away, is suggested to help older adults who feel disassociated with family events [8].

Despite these positive associations and benefits of use, younger adults are considerably more likely than older adults to be using them. Ofcom found that 92% of 16-24 year old respondents have set up an SNS profile, compared with 25% of respondents over 65 [18]. Similarly, Pew Research Center's report observes that 89% of the respondents aged 18-29 use SNSs, compared with 43% of over 65s [3].

A lack of interest and purpose have been suggested as leading reasons for the low uptake amongst the older population [7, 9, 13, 16, 20]. This signifies a digital divide and a problem for older adults if SNSs are used more widely for communication and diffusion of information. We believe that by considering the needs and opinions of older adults, a more acceptable and useful platform for online communication can be provided.

This research aims to explore measures to make SNSs more inclusive by creating user-centred recommendations for developers of such sites. These recommendations attempt to avoid common barriers which can prevent older adults from choosing to use SNSs. These suggested improvements aim to support older adults who would potentially be interested in using a SNS, but don't for a variety of reasons, rather than those that find SNSs uninteresting or without purpose. While the recommendations do not contain a solution to every problem, we believe that they are a positive step in identifying potential solutions to key barriers.

Additionally, despite sharing similar concerns, such as privacy, many younger adults continue to use SNSs in order to enjoy social inclusion [10]. While many of the concerns may not be unique to older adults, we support that developing for 'extra ordinary' users helps improve the service for people of all ages [19].

This paper describes the process of creating and validating the recommendations, outlines an exploratory comparative evaluation to explore their impact and provides context into modifications that aided participants in completing tasks through observational notes and empirical results. We lastly open a discussion around design considerations, and how developers can improve SNSs for a wider range of people.

## RELATED WORK

### Exploring the Lack of Participation

From the literature, several barriers and reasons for not using SNSs are raised from participant discussions. For many of the studies, older adults could not see the purpose, or benefit, of using SNSs [7,9,13,16,20]. Participants were content with their existing methods of communication, such as telephone, SMS and email [13]. Lehtinen et al. argue that showing clear benefits and purposes of using a SNS is vital if we are to introduce older adults to this technology.

Lehtinen et al. observed that current SNSs “*do not seem to fit the everyday communication of older adults well*” [13]. This was supported by arguments that older adults find intentional self-presentation as not socially acceptable, and online communication as “cold”. SNSs were perceived as “places for dating or celebrities, not for themselves”. This was additionally observed by Gibson et al., with older adults commenting on negative feelings of being “on display” [9].

Preconceptions can be a barrier, such as if one was to encounter negative media stories involving SNSs [9, 24]. This can lead to older adults being put off from joining such a site. Mikkola and Halonen found that while Facebook was used very little by participants, they all knew what it was, viewing the site as “useless time consuming” [16]. At times, negative media coverage has led to misconceptions and concerns about the behaviour of SNSs [9,24].

Privacy has been shown to be a major barrier in the literature, being described as a “major concern” and “very important” [7, 9–11, 13, 14, 16, 17, 20, 23, 24]. A degree of distrust was observed by Mikkola and Halonen, reporting that some older adults were afraid of joining Facebook [16]. The group discussions held by Xie et al. had a strong emphasis on privacy, being described as the “primary concern” and “key perceptual barrier to adoption” [24].

While SNSs designed specifically for older adults do exist, they have generally not had a large uptake [6]. This may in part be due to a lack of family members on these sites, a lack of critical mass, people not knowing that they exist or little being done to address the actual barriers faced by older adults.

A number of other barriers exist, such as a fear of committing social blunders, a lack of confidence, a lack of skill, concerns over security and abusive other users [7, 13, 20, 24]. Much of the literature overlaps, with several emerging themes. These themes, such as a lack of purpose, incompatibilities, preconceptions of SNSs and privacy, are all suggested as important reasons as to why participation with older adults on SNSs is so low. We hold that exploring and understanding these themes is the first

step to providing an improved online environment and opportunity for older adults to explore SNSs.

### Older Adults, SNSs, Requirements and Usability

A variety of evaluations and usability studies have been undertaken with a variety of different types of systems. Chou et al. conducted ‘think-aloud’ interviews with 5 adults over 55 years old in Taiwan while they used Facebook [7]. Inconveniences were categorised before being turned into user requirements, which were then combined with general web guidelines [12] generated from existing research. Chou et al. go on to argue that Facebook can be improved for older adults through the use of these requirements [7].

Karahasanović et al. report a list of needs and requirements at both an individual and group level for elderly people and sites with user generated content (such as SNSs) [11]. Requirements such as ease of use, a need for control and a concern about the skills required to stay up to date were all found to be important to participants.

Gibson et al. discuss how participants used Ning<sup>2</sup>, an open source social platform, to create their own closed-off social network [9]. This custom SNS had the underlying purpose of providing information on upcoming events and classes. Gibson et al. suggest that a clear purpose and a community focus, rather than on any individual, may be one of the key reasons for the site’s uptake. The SNS was additionally not publicly available to anyone outside the group, with users expressing greater feelings of security. A secure and focused SNS may be seen as more acceptable for this age group, conditional on careful behaviour and design considerations.

Many of the suggestions from the literature focus on technical requirements which do little to address issues relating to how the site behaves. Little attention has also been given to investigating the impact of these requirements in a controlled evaluation. There is a need for research which addresses both of these aspects. By evaluating our recommendations using a comparative evaluation, we can begin to understand how much of an impact these changes can make, and promote them to developers and researchers.

## CREATION OF RECOMMENDATIONS

### Focus Groups

To generate the recommendations, two focus groups were held with separate groups of Internet users over 60 years old. Participants were recruited from the SiDE user pool [22], a collection of over 800 volunteers, mostly over 65, who have been recruited through local organisations, groups or clubs for accessibility research. Both focus groups were audio recorded and transcribed so that qualitative analysis could take place. Participants were given a £10 gift voucher.

The first focus group involved 8 people (4 male, 4 female; aged 61-80,  $M = 67.9$ ,  $SD = 5.8$ ), two of whom were SNS users, two were previously SNS users and four were non-users. A mixture of backgrounds was sought to explore a range of opinions from different experiences. The

<sup>2</sup><http://www.ning.com/>

non-users had mostly high-level concerns (such as privacy, security) whereas the users had specific examples of barriers that they had experienced and overcome (e.g. difficulties accessing privacy settings, frequent unsolicited emails). These examples were generally more specific and informed. The experiences of those who had previously overcome barriers in their early stages of use were recognised as valuable for identifying potential problems for new users.

This led to SNS users being specifically recruited for the second focus group. 7 people were recruited (5 female, 2 male; aged 63-78,  $M = 68.7$ ,  $SD = 5.2$ ). Due to a misunderstanding during the selection process, one participant was not a SNS user. Many of the participants who used SNSs did so for communication with family and friends, describing it as beneficial for themselves.

Participants of both focus groups were asked to discuss positive and negative opinions of SNSs alongside improvements that could be made. Conversations were steered through a number of topics by the lead researcher, such as how the participants used SNSs, why they started using them, why they chose not to use them and how those sites could be made more suitable.

### Qualitative Analysis

To analyse the focus group transcripts, thematic analysis [2] was used to categorise and find themes relating to both positive and negative aspects of SNS usage from quotes in the transcripts. For example, one participant commented that “*You’ve got to keep the information really simple and with simple terminology*”. Quotes were then coded with one or more themes, such as “Terminology”.

A further stage of thematic analysis was conducted with two key qualitative papers in the literature [9, 13], chosen due to their relevance in the field. This aimed to investigate theoretical saturation and congruity. Text from the literature, e.g. ‘they were anxious to retain their privacy’ [9], was coded in the same way as text from the transcripts. Findings in the literature were coded as new themes, unless a new theme overlapped with one from the transcripts, whereupon the existing theme was used.

Including these two papers in the later stages of thematic analysis enabled us to verify the consistency of our themes with other research. A high amount of overlap appeared between our themes and those from the two papers. This suggests that many of our themes generalise to beyond the small sample sizes of the focus groups. Themes such as a lack of purpose, privacy concerns and negative media stories were present in both the literature and the transcripts.

Some themes were unique to the focus groups, such as frequent and unsolicited emails to the user, difficulties experienced when trying to deactivate an account, constantly changing interfaces, confusion over financial cost and feelings of invasion regarding friend suggestions. Other themes were unique to the literature, such as a fear of social blunders and a desire for partial anonymity, mass distribution of messages and off-the-record (temporary) messages.

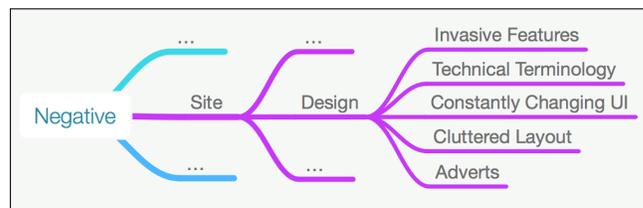


Figure 1. A subset of the themes. The sub-themes are all relating to people raising concerns over the design of specific sites.

While thematic analysis usually encourages refinement toward a small collection of abstract themes, maintaining the granularity and specificity of our themes was necessary to preserve the wide range of issues. Our solution was to categorise similar themes into sub-themes throughout the coding process, refining the analysis while maintaining granularity. This generated a large mind map structure, with abstract themes near the center and more specific themes near the outside. Figure 1 shows a subset of the themes.

### Identifying the Recommendations

Traversing from higher to the lower levels of the mind map provided context into each granular theme (e.g. ‘Negative’ > ‘Don’t like SNSs’ > ‘Worried’ > ‘Privacy’ > ‘Lack of Transparency’). Recommendations for developers were then derived from this process by a researcher. From the above example, a recommendation was identified to ‘Clarify and simplify privacy and account settings’, as this would likely help users who were worried about the lack of transparency regarding privacy settings. Additional context was gained by consulting the quotes which were coded under the sub-theme in question. This aimed to avoid misinterpretations during the process of deriving the recommendations.

The sub-themes on the right of Figure 1 are all categorised as issues with the design of specific SNSs. Some recommendations that were derived from this subset include *using a simple layout* and *avoiding technical terminology*, *adverts* and *frequent changes to the interface*. ‘Invasive Features’ is an example of where quotes were utilised. Participants described friend suggestions and unsolicited emails as invasive. This led to a recommendation to avoid such invasive functionality.

Themes were treated equally when identifying the recommendations, regardless of originating from the transcripts or the literature. Each recommendation involved suggesting a way to either eliminate a concern or support a positive feature. In cases of themes relating to a positive feature, the recommendation suggested supporting that feature. In cases of themes relating to a concern, the recommendation attempted to identify the root cause of the problem and suggest a method to bypass this concern. If a possible solution could be identified by the researcher then it was included as a recommendation. Some themes, however, could not be addressed by a recommendation, e.g. ‘Negative’ > ‘Not Interested’ > ‘I am happy without it’. In this case, ‘I am happy without it’ is at the lowest level of granularity. Neither the parent themes or coded quotes identify a way to ‘solve’ this reason for not using a SNS.

| <b>Initial Recommendations</b> |   | <b>CUI</b> | <b>MUI</b> |
|--------------------------------|---|------------|------------|
| Design & Behaviour             | 1.1 Have increased privacy by default.  | X          | √          |
|                                | 1.2 Avoid invasive functionality and emails (e.g. friend suggestions).                | ⊗          | ⊗          |
|                                | 1.3 Clarify implications of functions to prevent social blunders.                     | †          | √          |
|                                | 1.4 Reduce and moderate fake accounts.  | ⊗          | ⊗          |
|                                | 1.5 Clearly label navigation to privacy, email and account settings.                  | X          | √          |
|                                | 1.6 Avoid targeting specific age groups (e.g. over 65).                               | √          | √          |
|                                | 1.7 Avoid pressurising techniques, reminders or features to encourage contributions.  | X          | √          |
|                                | 1.8 Avoid games, apps and irrelevant functionality.                                   | X          | √          |
|                                | 1.9 Avoid similar designs to sites where people hold negative assumptions. *          | X          | Sub        |
| User Control                   | 2.1 Provide control to delete content.  | †          | √          |
|                                | 2.2 Provide control to delete account if desired.                                     | †          | √          |
|                                | 2.3 Clarify and simplify privacy and settings.  | †          | √          |
|                                | 2.4 Allow control over what emails (if any) the user wishes to receive.               | †          | √          |
|                                | 2.5 Allow partial anonymity if desired.   | X          | X          |
|                                | 2.6 Allow simple removal of contacts.   | √          | √          |
|                                | 2.7 Allow categorization of contacts.   | X          | X          |
|                                | 2.8 Give users control of which contacts see what content.                            | X          | X          |
|                                | 2.9 Allow for mass distribution of messages. *  | ⊗          | ⊗          |
|                                | 2.10 Allow for session, or 'off-the-record' private messages. *                       | X          | X          |
|                                | 2.11 Provide moderation/report functionality. *                                       | †          | √          |
|                                | 2.12 Focus on a specific purpose (e.g. club or group). *                              | X          | X          |
| Generalisable Beyond SNSs      | 3.1 Use a simple, but bold, layout on the site.                                       | Sub        | Sub        |
|                                | 3.2 Reduce or avoid advertisements.   | X          | √          |
|                                | 3.3 Clarify if the site has a financial cost or is free.                              | √          | √          |
|                                | 3.4 Reassure users that suitable security measures are in place (e.g. SSL/https).     | ⊗          | ⊗          |
|                                | 3.5 Provide support to explain/help/guide users through tasks.                        | ⊗          | ⊗          |
|                                | 3.6 Avoid unnecessary complexity when changing settings or completing tasks.          | Sub        | Sub        |
|                                | 3.7 Avoid drastic and frequent changes to the site.                                   | ⊗          | ⊗          |
|                                | 3.8 Avoid technical terminology.  | √          | √          |
|                                | 3.9 Avoid using 'small print' checkboxes for additional services/information sharing. | √          | √          |
|                                | 3.10 Reduce or avoid emails to the user by default.                                   | X          | √          |
|                                | 3.11 Follow accessibility guidelines.   | ⊗          | ⊗          |

√ = Recommendation met    X = Recommendation not met    Sub = Subjective (debatable)

† = Recommendation somewhat met (e.g. difficult to find, unclear)    ⊗ = Difficult to implement/represent for study

**Table 1. The list of initial recommendations prior to the evaluation. Additionally shows which recommendations were met in each UI version. CUI = Control UI, MUI = Modified UI. Those marked with an \* are discussed further in the Results section, which should be consulted before use.**

The process resulted in the generation of 32 initial recommendations, shown in Table 1, to mitigate concerns or support positive features identified in the focus groups and literature. At this stage, the recommendations consist of potential solutions interpreted by the researcher from the analysis. The next step of the research is to evaluate the individual recommendations for participant agreement (validity), as well as the overall impact of the initial set of recommendations to investigate reliability.

## EVALUATION

To explore the impact of the recommendations, we developed a system and recruited non-SNS users to partake in a comparative evaluation. Non-SNS users were chosen to explore the impact of the recommendations for people who had never used a SNS before, simulating new users. Additionally, SNS users would likely skew results toward the control interface in an evaluation due to familiarity.

This approach led to a few challenges. It is difficult to develop a control site which does not follow any of the recommendations as this would require a conscious effort on the part of the developer to discard good design practice (SNSs commonly adhere to some, but not all, of the recommendations presented in this paper). Our solution to this was to investigate the impact that the recommendations may have on an existing SNS. This raises further challenges due to our requirement of non-SNS users as participants. Ethical considerations (the lack of consent from contacts or friends, requiring participants to sign up to a service) and the difficulty of conducting a controlled study on such a diverse site also surfaced in early stages of planning the study.

The solution resulted in a combination of these options. An offline likeness to Facebook was developed, providing the main functionality (e.g. creating content, 'unfriending' contacts and changing settings) and controls of the system.

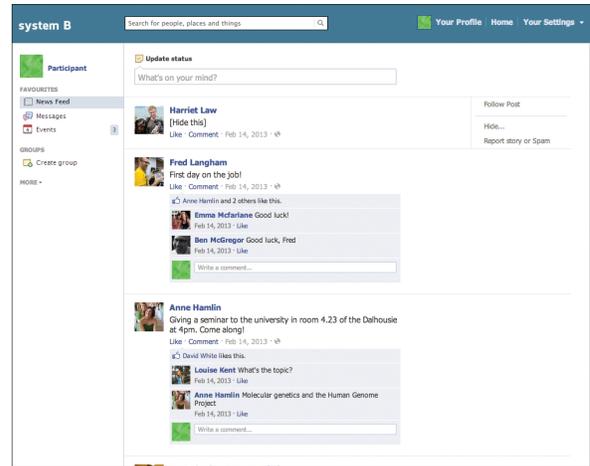


Figure 2. Homepages of the Control UI (left) and Modified UI (right).

A similar approach has previously been used to comparatively identify improvements to Facebook privacy controls [15]. Creating our own prototype allows alternative interfaces, and attempts to simulate the experiences faced by new users of the site while avoiding the above ethical issues. This method, however, comes with reduced real-world applicability, as participants communicate on the prototype with pre-generated users representing real-life contacts.

### Prototype System

The system used for the study has two user interface (UI) versions, the 'Control UI', which was strongly influenced by Facebook (as of December, 2012), and the 'Modified UI', which adapts the Control UI to follow recommendations as indicated in Table 2. These two interfaces are presented to the participants as 'System A' and 'System B' respectively.

The implementation of the Modified UI relied on identifying alterations to the UI that aligned with the recommendations. Since many of the recommendations are subjective in nature, subjectivity was needed in order to implement these into the prototype. While some followed a logical progression, e.g. recommendation 3.2 led to the removal of adverts, some were the interpretation of the researcher. An example may be how the navigation buttons to settings were labelled, or the simplification of the interface. The goal of this study was to explore how one might utilise the recommendations and the subsequent impact that these changes would have.

Not all recommendations are listed in Table 2. There are three reasons for this. Firstly, 5 recommendations (1.6, 2.6, 3.3, 3.8 and 3.9) were included in both UI versions (and therefore not listed as changes) due to being met by Facebook (and subsequently the Control UI). For example, non-technical terminology is used on both the Control UI and the Modified UI. Facebook also does not specifically target age groups in the same way that, for example, some SNSs do for adults over 65.

Secondly, 7 recommendations (1.2, 1.4, 2.9, 3.4, 3.5, 3.7 and 3.11) were outlined in Table 1 as difficult to represent in the study. As an example, the true implications of simulating

frequent changes to the UI were not practical in an hour-long session. Another recommendation suggested the use of a support section of the site. This was avoided to prevent tasks from being overly easy to complete, potentially detracting from other findings of the study.

Thirdly, since the system is a simplified representation, some of the more sophisticated features of Facebook were not

| Change   | Rec.(s)            |
|--|--------------------|
| Removal of advertisements and reminder bar   | 1.7, 3.2           |
| 'Pages' and 'Apps' moved to under 'More' on left navigation bar  | 1.8, 3.1           |
| Changed the colour of the header   | 1.9                |
| Header size increased  | 1.5, 3.1           |
| Font of content slightly increased in size   | 3.1                |
| Drop-down box for settings pages labelled with 'Your Settings'   | 1.5                |
| Add text to clarify who will be able to see a new status update  | 1.3                |
| Location and tagging removed when creating a status update   | 1.8                |
| Default frequency of email notifications set to least frequent option  | 3.10               |
| Default privacy setting set to 'Friends'   | 1.1                |
| Buttons to change privacy settings moved to the top of the privacy settings page                                 | 2.3, 3.6           |
| 'Quick Settings' tab added to settings page which contains suggested settings                                    | 2.2, 2.3, 2.4, 3.6 |
| Removal of 'Subscribers', 'Apps', 'Mobile', 'Payments', 'Advertisements' and 'Gifts' tabs from the settings page | 1.8, 2.3, 3.1      |
| Hovering the cursor over content makes the report and delete/hide box appear directly                            | 2.1, 3.6, 2.11     |

Table 2. Table showing the changes made to the Modified UI and the relevant recommendation(s). Codes can be looked up in Table 1.

implemented. Functionality such as the ‘Chat’ feature, uploading photographs and the ‘list’ controls were often simulated (i.e. an inactive link) or omitted. In such cases, neither UI version had the functionality to meet this criterion. Additionally, since some of the recommendations would drastically modify the focus or functionality of the prototype, such as by changing the purpose of the site and integrating anonymity and temporary private messages, the decision was taken to omit these from the prototype. This related to the aim of the study, an example of how developers can utilise the recommendations that are relevant to the design and purpose of their SNS. Hence, the prototype utilised the recommendations which maintained the main functionality and purpose of the Control UI.

Due to these reasons, the evaluation of the prototype is, in essence, a case study into the impact of the recommendations on a theoretical site. This site shares many of the same functionality and design aspects as Facebook, as is the case with many other SNSs. The study should therefore not be viewed as an evaluation of Facebook itself.

## METHOD

30 participants, who were all non-SNS users but Internet users with no severe cognitive impairments, were recruited from the SiDE user pool [22] to partake in the study. None were participants of either focus group. Of the 30 participants, one did not attend, three did not wish to complete the comparative evaluation part of the study and one wasn’t able to due to having a visual impairment. While having no visual impairment was a prerequisite for this study, this was not discovered for this participant during the recruitment process. Subsequently, they could not complete the comparative evaluation part of the study. The four participants who did not, or could not, complete the comparative evaluation wished to complete the questionnaire to evaluate the recommendations. This resulted in 25 (aged 64-86,  $M = 71.2$ ,  $SD = 5.07$ ) participants for the comparative evaluation and 29 (aged 64-89,  $M = 72.3$ ,  $SD = 6.30$ ) for the questionnaire. Participants were given a £10 gift voucher.

The study was conducted in a private testing lab with two separate computers. Each computer had its own monitor, mouse and keyboard which were positioned perpendicular to each other. The participant could not see both screens at the same time, while the researcher could view both screens without moving and potentially distracting the participant.

For each participant, the computer running each version of the UI remained the same for the duration of the session. Participants could therefore associate each UI with a physical computer to make it easier to differentiate between ‘System A’ and ‘System B’ when later answering questions relating to that UI version. Each computer had an identical, but separate, prototype database which was reinitialised and repopulated using a ‘Seed’ file for every participant. This ensured that there were 16 (simulated) users and pre-generated content on each system, which remained consistent for every participant, regardless of deletions or modifications of content during prior sessions.

For the comparative evaluation, participants were asked to attempt 11 tasks on both UI versions, described in Table 4. These represented tasks that were identified in the earlier focus groups as important for the participants to be able to do. Participants attempted each task on one of the two UI versions before moving on to the second UI. The order of which UI version the task started with alternated for each participant and each task to counterbalance the study and mitigate practice effects. For the second part of the study, a questionnaire was completed measuring agreement with each individual recommendation.

## Data Collection

After each task was attempted on both systems, participants answered a Likert question on paper asking which system (UI version) they preferred completing the task on. This 5 point Likert question ranged from “System A greatly” (Control UI) to “System B greatly” (Modified UI) with a neutral point of “No Opinion”. This resulted in eleven of these answers for each participant (one for each task) for the study. After all tasks were attempted, participants completed two System Usability Scale (SUS) questionnaires [4], which each provided a 0-100 score relating to the usability of the corresponding UI version. In addition to this, we noted whether a participant successfully completed each task, allowing completion rates to be investigated. Observational notes were taken when problems were encountered.

The second part of the study consisted of a series of Likert questions aimed at measuring agreement (validity) with 31 of the 32 recommendations. One recommendation (3.11) which promotes the use of accessibility guidelines was not included in the questionnaire due to a strong existing body of support for accessible design, as seen in the WAI-AGE guidelines<sup>3</sup>. Recommendations with a median response that demonstrates agreement are considered verified and accepted. Those with a median response that demonstrates disagreement or no opinion with the recommendation are subsequently rejected.

To balance the questionnaire, an equal number of “Agree” and “Disagree” answers were expected to show full support for all recommendations by the participants. In order to achieve this balance, we used a method of polarising, where a positively worded question becomes a negatively worded question (e.g. should/shouldn’t), to ensure that an even mix of each is asked. Using this balancing method, however, means that for some questions it is not possible to assume that the recommendation has been supported, rather, it demonstrates that the polarised opposite of the recommendation is disagreed with. This is countered by creating two questionnaires. Each question is the polarised opposite (positive/negative) on the opposing questionnaire, resulting in the average of both standard and reversed questions being used to evaluate agreement for each recommendation. This allows us to avoid patterns which can be detected by participants and bias results.

<sup>3</sup><http://www.w3.org/WAI/older-users/>

| Question   | Control UI Median & (Q1 - Q3) | Modified UI Median & (Q1 - Q3) | Wilcoxon       |                   |             |
|--|-------------------------------|--------------------------------|----------------|-------------------|-------------|
| 1. I think that I would like to use this system frequently                                   | Disagree (1 - 2)              | Agree (1 - 3)                  | $Z = -2.592$ , | $p = .010^*$ ,    | $r = -.367$ |
| 2. I found the system unnecessarily complex  | No Opinion (1 - 3)            | Disagree (1 - 2.5)             | $Z = -2.496$ , | $p = .013^*$ ,    | $r = -.353$ |
| 3. I thought the system was easy to use  | Disagree (1 - 3)              | Agree (1 - 3)                  | $Z = -2.464$ , | $p = .014^*$ ,    | $r = -.348$ |
| 4. I think that I would need the support of a technical person to be able to use this system | No Opinion (1 - 3)            | No Opinion (1 - 3)             | $Z = -1.518$ , | $p = .129$ ,      | $r = -.215$ |
| 5. I found the various functions in this system were well integrated                         | No Opinion (1 - 2)            | No Opinion (2 - 3)             | $Z = -2.191$ , | $p = .028^*$ ,    | $r = -.310$ |
| 6. I thought there was too much inconsistency in this system                                 | No Opinion (1.5 - 2)          | No Opinion (1 - 2)             | $Z = -1.058$ , | $p = .290$ ,      | $r = -.150$ |
| 7. I would imagine that most people would learn to use this system very quickly              | Agree (1.5 - 3)               | Agree (3 - 3)                  | $Z = -2.359$ , | $p = .018^*$ ,    | $r = -.334$ |
| 8. I found the system very cumbersome to use   | No Opinion (1 - 3)            | Disagree (1 - 2)               | $Z = -2.251$ , | $p = .024^*$ ,    | $r = -.318$ |
| 9. I felt very confident using the system  | Disagree (1 - 2)              | No Opinion (1 - 3)             | $Z = -2.636$ , | $p = .008^{**}$ , | $r = -.373$ |
| 10. I needed to learn a lot of things before I could get going with this system              | Agree (1.5 - 3)               | Agree (1 - 3)                  | $Z = -1.155$ , | $p = .248$ ,      | $r = -.163$ |

\*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$

**Table 3.** Table showing the median responses, and first (Q1) and third (Q3) numerical quartiles, to the individual Likert questions in the System Usability Scale for each UI version. Results from a Wilcoxon signed-rank test between the two UI versions per question are included. For Q1 and Q3, 0 = ‘Strongly Disagree’, 1 = ‘Disagree’, 2 = ‘No Opinion’, 3 = ‘Agree’, 4 = ‘Strongly Agree’.

## RESULTS

### System Usability Scale Scores

This study generates two SUS scores [4] (one for each UI version) per participant. Using a Wilcoxon signed-rank test, we can show that there was a significant difference in the SUS scores for the Control UI ( $Mdn = 40$ ) and the Modified UI ( $Mdn = 55$ );  $Z = -3.056$ ,  $p = .002$ ,  $r = -.432$ .

We can additionally look at the individual answers to gain insights into the different components of the questionnaire. Table 3 shows the median response to each question per UI version, as well as a Wilcoxon signed-rank test comparing the two UI versions for each question. Seven of the ten tests confirm a significant difference between the responses. Of those seven questions, five have a different median response for each system (questions 1, 2, 3, 8 and 9). The remaining two questions are significantly different on each UI version while having the same median response option (questions 5 and 7).

The individual Likert questions of the two SUS scales also allow for more specific comparisons into the different components of the SUS’s usability model over just the score alone. For example, participants mostly agreed that they would like to use the Modified UI regularly, but disagreed with the same question regarding the Control UI. Similarly, participants generally agreed that the Modified UI was easy to use, but did not agree that the Control UI was.

### UI Preference Per Task

While examining the participant responses to the Likert-based preference question for each task, we categorize answers into 3 categories: (1) Control UI is preferred (“greatly” or “slightly”), (2) Modified UI is preferred (“greatly” or “slightly”) or (3) “No Opinion” is selected. We group the Likert points together to compare the three overall perspectives, while not limiting choice.

A Friedman test was used after the assumptions of a one-way repeated ANOVA were not met. The results of the Friedman test show that there was a statistically significant difference between the number of instances where participants preferred either the Control UI ( $Mdn = 1$ ), the Modified UI ( $Mdn = 5$ ) or “No Opinion” ( $Mdn = 4$ );  $\chi^2(2, 25) = 24.344$ ,  $p < .001$ . Post-hoc analysis, using three Wilcoxon paired rank tests, was conducted with a Holm-Bonferroni correction applied. There were statistically significant differences between the Control UI and Modified UI categories,  $Z = -4.053$ ,  $p < .001$ ,  $r = -.573$ ), between the Modified UI and “No Opinion” categories,  $Z = -2.043$ ,  $p = 0.041$ ,  $r = -.289$ , and between the Control UI and “No Opinion”,  $Z = -2.989$ ,  $p = .003$ ,  $r = -.423$ ).

The high frequency of “No Opinion” being selected may be partially explained by considering that not all methods of completing the tasks were significantly modified by the recommendations. This resulted in some tasks having the

| Task   | CUI | MUI | Wilcoxon       |                    |             |
|--|-----|-----|----------------|--------------------|-------------|
| 1 Leave a 'Post' on a contact's profile page.                  | 22  | 20  | $Z = -1.414$ , | $p = .157$ ,       | $r = -.200$ |
| 2 Change privacy settings.                                     | 16  | 24  | $Z = -2.828$ , | $p = .005^{**}$ ,  | $r = -.400$ |
| 3 Report a 'Status Update' from a contact.                     | 5   | 11  | $Z = -2.449$ , | $p = .014^{*}$ ,   | $r = -.346$ |
| 4 Create a Status containing text.                             | 24  | 23  | $Z = -1.000$ , | $p = .317$ ,       | $r = -.141$ |
| 5 'Like' and comment on a contact's Status.                    | 6   | 7   | $Z = -1.000$ , | $p = .317$ ,       | $r = -.141$ |
| 6 Change how often the site emails the user.                   | 8   | 22  | $Z = -3.742$ , | $p < .001^{***}$ , | $r = -.529$ |
| 7 Hide a Status from the 'News Feed'.                          | 14  | 19  | $Z = -1.890$ , | $p = .059$ ,       | $r = -.267$ |
| 8 Delete a Status created by the user.                         | 22  | 24  | $Z = -1.414$ , | $p = .157$ ,       | $r = -.200$ |
| 9 Delete a Post created by someone else on the user's profile. | 20  | 22  | $Z = -1.414$ , | $p = .157$ ,       | $r = -.200$ |
| 10 'Unfriend' a contact.                                       | 22  | 22  | $Z = 0.0$ ,    | $p > .999$ ,       | $r = 0.0$   |
| 11 Deactivate the user's account.                              | 7   | 8   | $Z = -1.000$ , | $p = .317$ ,       | $r = -.141$ |

\*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$

**Table 4.** A brief description of each task each participant attempted along with the number of participants who completed each task on the Control UI (CUI) and the Modified UI (MUI). Results from a Wilcoxon signed-rank test between the two UI versions are also included (N = 25).

same steps to completion on both versions of the interface. This does, however, raise interesting implications for the modifications which were only aesthetic (e.g. a reduction in advertisements). Such results suggest that these aesthetic alterations did not lead participants to prefer the Modified UI over the Control UI. Figure 3 shows the number of each Likert response selected on each task and which tasks had a high number of "No Opinion" responses.

### Task Completion Rates

It is worth considering that neither interface is meant as a solution, only as a means for examining the guidelines. As a result, increasing task completion rates remain secondary to the main goals of the study. However, they are included for completeness and for further analysis of the findings.

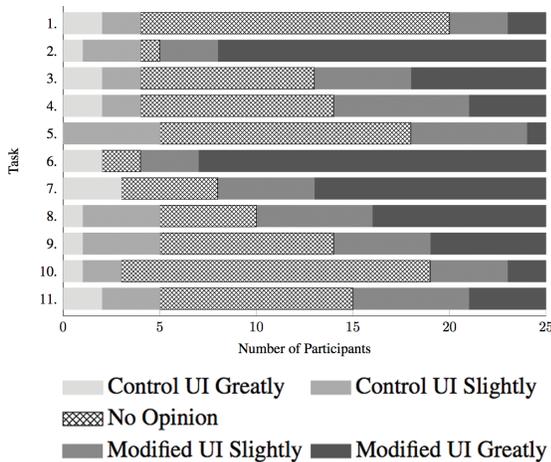
Initially, we look at the number of tasks that each participant completes on each UI version. A significant Shapiro-Wilk test of normality confirmed that the assumptions of a paired t-test were not met. Thus, the Wilcoxon signed-rank test was used, which shows that a statistically significant difference

exists between the median number of tasks completed on the Control UI ( $Mdn = 6$ ) and the Modified UI ( $Mdn = 8$ ) across participants;  $Z = -4.025$ ,  $p < .001$ ,  $r = -.569$ .

This result gives us an idea of the number of successfully completed tasks on each UI version by participants, but we can also compare the number of successful task completions for the Control UI and the Modified UI on participants' first attempts at each task only, reducing the practice effect. Since half of the participants started each task on each version of the UI, the variables of the test cease to be paired.

A significant Shapiro-Wilk test of normality indicated that the data were not normally distributed, therefore a Mann-Whitney U test is used to show that there was a significant difference in the median number of tasks completed on the Control UI ( $Mdn = 3$ ) and the Modified UI ( $Mdn = 5$ ) on participants' first attempts at tasks;  $U = 137.0$ ,  $Z = -3.556$ ,  $p < .001$ ,  $r = -.503$ .

Table 4 additionally shows how many participants completed each task on each UI version. Wilcoxon signed-rank tests are used to identify tasks where a significant difference occurred in the number of participants who were able to complete it. Three such tasks are statistically significant.



**Figure 3.** Bar chart showing UI preference responses for each task.

### Accepted Recommendations

The data have been processed to take into account the two versions of the questionnaire and the reversed questions, providing a 0-4 value representing the degree of agreement participants had with, and subsequent validity of, each recommendation. Recommendations are classed as supported when the median response is 3 or above, which represents a median response of "Agree" or "Strongly Agree" on the Likert item. 27 recommendations are found to be supported, while 5 have been listed as "Rejected" due to a median response of "Strongly Disagree", "Disagree" or "No Opinion". The codes for these recommendations are listed on Table 5 and can be looked up on Table 1, where rejected recommendations are marked with an asterisk.

| Accepted (SA) | Accepted (A) | Rejected        |
|---------------|--------------|-----------------|
| 1.1 - 1.5     | 1.6 - 1.8    | 1.9 (NO)        |
| 2.1 - 2.5     | 2.6 - 2.8    | 2.9 - 2.11 (NO) |
| 3.1 - 3.6     | 3.7 - 3.11   | 2.12 (D)        |

**Table 5. Recommendations accepted and rejected by the evaluation. Codes can be looked up on Table 1. SA = ‘Strongly Agree’, A = ‘Agree’, NO = ‘No Opinion’, D = ‘Disagree’. Those rejected should not be used.**

## DISCUSSION

This study investigated potential solutions to barriers faced by older adults when using SNSs and explored how these recommendations could be utilised by developers. This study should be viewed as a case study into the impact of the recommendations on a theoretical but representative SNS site. The study suggests that improvements can be made when taking the recommendations into consideration, at least in terms of perceived usability, task completion and user preference (for completing certain tasks).

The third subset of the recommendations (3.1 - 3.11) generalise beyond SNSs. A number have been identified as good design principles for older adults, such as in the WAI-AGE guidelines<sup>4</sup> or those by Kurniawan and Zaphiris [12]. While these recommendations may be generally applicable, they were identified as barriers to SNS usage, and are therefore included in Table 1.

Our contribution is firstly to identify these potential solutions as a collection and then to investigate the impact that they may have on a SNS. Other research in the literature corroborates some of our findings.

### Considerations for Development

While the results suggest that the recommendations led to improvements in usability and task completion rates, such improvements depend on how these recommendations are applied to SNSs. Thus, an important aspect of what can be gained from this research is likely in the observations of how these improvements are made and the implications for the design of inclusive SNSs.

There were significant differences in the number of participants who completed a task for each UI version in 3 of the 11 tasks, as seen on Table 4. Two of these are related to changing settings. Observational notes from these tasks highlight that many participants who did not complete the task could not find out how to navigate to either the settings or the privacy page. Modifying the navigation panel in the top right to include text in addition to the white arrow of the drop down box led to participants discovering the privacy and settings pages. Using unambiguous text for buttons, rather than abstract icons or symbols (e.g. ‘Your Settings’ rather than a gear or arrow icon) can avoid this issue.

For task 2, participants who couldn’t find the ‘Settings’ drop down box on the Control UI often then found a ‘Privacy’ link at the bottom of the page, which isn’t always possible when sites utilise AJAX to generate new content when the

bottom of the page is reached. While AJAX allows for a fluid experience, developers should be conscious of any text, link or information at the bottom of the page which will be difficult or unreachable because of this feature, and consider moving them to a static location on the page (e.g. a sidebar).

While task 3 (reporting content to an administrator) was statistically higher on the Modified UI, the completion rates for either UI were not encouraging. Our method of having the collapsed set of options appear when hovering over content did not go far enough to allow a majority of participants to complete the task. Observing the participants revealed that they simply did not know what to do, and it was not apparent that hovering the mouse over that content would uncover additional features, as was the case for both UI versions. Developers should therefore aim to reduce mouse-hover functionality for older adults, or leave clear indicators that functionality is there. An example may be an always-visible ‘Delete’ link within content, which expands into different options when the user hovers over the area.

Some tasks did not have a high completion rate on either system, such as reporting a status, deactivating the account and ‘liking’ and commenting on a contact’s status update. For the latter, many participants had difficulty differentiating their profile from that of other people. Many created their own status as a reply, which wasn’t linked to that of the original. Many participants also struggled due to the lack of a visible ‘Submit’ button when writing a comment and did not know that pressing the ‘Enter’ key on the keyboard was needed. Again, this assumption of prior knowledge made the task difficult for participants to complete.

The findings of this study suggest that careful design considerations can increase the usability and inclusivity of a SNS. Assumptions of prior knowledge, such as the meaning of icons, features which appear on mouse hover and enter-to-submit text boxes, should be removed or reconsidered. These assumptions are increasingly common on the Internet, and thus, these recommendations may generalise beyond SNSs. The Modified UI version was preferred more frequently than the Control UI. We believe that this demonstrates that common practices used on SNSs can be improved upon, leading to a more inclusive online environment. By investigating how we can improve these sites for older adults, we can provide an online environment where a wider range of people can feel comfortable communicating and sharing content with family and friends, with implications for social capital, loneliness and life satisfaction. Developers of such sites can consider and implement these recommendations as they feel appropriate.

### Future Work

Future research which evaluates the recommendations on a wider range of sites, to varying degrees of application and with different age ranges of older adults will improve our understanding of the impact such measures can have. Due to the fast changing nature of SNSs, it is also likely that more barriers to SNS usage may present themselves in the future. Future iterations can expand these recommendations as SNSs evolve by using the methods outlined in this paper.

<sup>4</sup><http://www.w3.org/WAI/older-users/>

## CONCLUSION

While social networking sites (SNSs) have become increasingly popular, a digital divide has emerged between younger and older people regarding SNS use. For this study we identified barriers faced by older adults and defined novel solutions to these by exploring how recommendations could be developed for use in SNS design. We verified and evaluated these recommendations using a comparative evaluation with older participants, finding that an SNS version which utilised these recommendations received a higher usability score, had a higher task completion rating and was preferred more frequently than one that did not. Observational notes were used to outline key considerations for developers, and we have argued that utilising these recommendations can improve SNSs for a greater number of people. We have presented these findings and discussed the implications for developers of social networking sites.

## ACKNOWLEDGMENTS

The authors wish to thank colleagues and reviewers for their feedback, and all participants of the study. They gratefully acknowledge support from RCUK Digital Economy Research Hub EP/G066019/1 "SiDE", a UK EPSRC doctoral training grant, and a Royal Society Wolfson Award to the third author.

## REFERENCES

1. Blit-Cohen, E., and Litwin, H. Elder participation in cyberspace: A qualitative analysis of Israeli retirees. *J. of Aging Studies* 18, 4 (2004), 385 – 398.
2. Braun, V., and Clarke, V. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101.
3. Brenner, J., and Smith, A. 72% of online adults are social networking site users. Tech. rep., Pew Internet & American Life Project, 2013.
4. Brooke, J. SUS: A quick and dirty usability scale. In *Usability evaluation in industry*. Taylor and Francis, London, 1996.
5. Burke, M., Marlow, C., and Lento, T. Social network activity and social well-being. In *Proc. CHI 2010*, ACM Press (2010), 1909–1912.
6. Chen, Y. Usability analysis on online social networks for the elderly. Tech. rep., Helsinki University of Technology, 2009.
7. Chou, W. H., Lai, Y.-T., and Liu, K.-H. User requirements of social media for the elderly: A case study in Taiwan. *Beh. & Inf. Tech.* 32, 9 (Sept. 2013), 920–937.
8. Cornejo, R., Favela, J., and Tentori, M. Ambient displays for integrating older adults into social networking sites. In *Collaboration and Technology*, vol. 6257 of LNCS. Springer, 2010, 321–336.
9. Gibson, L., Moncur, W., Forbes, P., Arnott, J., Martin, C., and Bhachu, A. S. Designing social networking sites for older adults. In *Proc. HCI 2010*, BCS (2010), 186–194.
10. Hoofnagle, C., King, J., Li, S., and Turow, J. How different are young adults from older adults when it comes to information privacy attitudes and policies? *UC Berkeley Center for Law and Technology* (2010).
11. Karahasanović, A., Brandtzæg, P. B., Heim, J., Lüders, M., Vermeir, L., Pierson, J., Lievens, B., Vanattenhoven, J., and Jans, G. Co-creation and user-generated content-elderly peoples user requirements. *Computers in Human Behavior* 25, 3 (2009), 655 – 678.
12. Kurniawan, S., and Zaphiris, P. Research-derived web design guidelines for older people. In *Proc. ASSETS 2005*, ACM Press (2005), 129–135.
13. Lehtinen, V., Näsänen, J., and Sarvas, R. A little silly and empty-headed: older adults' understandings of social networking sites. In *Proc. HCI 2009*, BCS (2009), 45–54.
14. Lewis, S., and Ariyachandra, T. Seniors and online social network use. In *Proc. CONISAR 2010* (2010).
15. Lipford, H. R., Besmer, A., and Watson, J. Understanding privacy settings in Facebook with an audience view. In *Proc. UPSEC 2008* (2008), 2:1–2:8.
16. Mikkola, K., and Halonen, R. "Nonsense?" - ICT perceived by the elderly. In *the European, Mediterranean & Middle Eastern Conference on Information Systems* (2011), 306–317.
17. Nyemba, E. S. F., Mukwasi, C. M., Mhakure, S., Mosiane, S., and Chigona, W. Golden baby boomers' perceptions of online social networking sites. *Int. J. Comp. Tech. App* 2, 3 (2011), 695–703.
18. Ofcom. Adults' media use and attitudes report 2013. Tech. rep., Ofcom, 2013.
19. Pullin, G., and Newell, A. Focussing on extra-ordinary users. In *Universal Access in HCI. Coping with Diversity*. Springer, 2007, 253–262.
20. Sundar, S. S., Oeldorf-Hirsch, A., Nussbaum, J., and Behr, R. Retirees on Facebook: can online social networking enhance their health and wellness? In *Proc. CHI EA 2011*, ACM Press (2011), 2287–2292.
21. Valenzuela, S., Park, N., and Kee, K. F. Is there social capital in a social network site?: Facebook use and college students' life satisfaction, trust, and participation. *JCMC* 14, 4 (2009), 875–901.
22. Watson, P., Hanson, V., Olivier, P., Tanaka, A., Li, F., Arnott, J., Gregor, P., and van Moorsel, A. Social inclusion through the digital economy. In *Digital Futures 2010* (2010).
23. Wherton, J., and Prendergast, D. The building bridges project: Involving older adults in the design of a communication technology to support peer-to-peer social engagement. In *HCI and Usability for e-Inclusion*, vol. 5889 of LNCS. Springer, 2009, 111–134.
24. Xie, B., Watkins, I., Golbeck, J., and Huang, M. Understanding and changing older adults' perceptions and learning of social media. *Educational Gerontology* 38, 4 (2012), 282–296.