Available online at www.sciencedirect.com
SCIENCE DIRECT®
Interacting with Computers xx (2005) 1–22
Interacting with Computers
www.elsevier.com/locate/intcom

Introducing the Internet to the over-60s: Developing an email system for older novice computer users

Anna Dickinson[a,*], Alan F. Newell[a], Michael J. Smith[b], Robin L. Hill[c]

[a]*Applied Computing, University of Dundee, Dundee, UK*
[b]*Business Consulting, Fujitsu Services, Manchester, UK*
[c]*Human Communication Research Centre/ Institute for Communicating and Collaborative Systems, School of Informatics, University of Edinburgh, Edinburgh, UK*

Abstract

Contemporary technology offers many benefits to older people, but these are often rendered inaccessible through poor software design. As the Internet increasingly becomes a source of information and services it is vital to ensure that older people can access these resources. As part of project funded by the UK government, a multi-disciplinary team set out to develop usable software that would help to introduce older people to the Internet. The first step was to develop an email system for older people with no experience of Internet use. The project was intended to show that it is possible to design usable technology for this group and to explore some of the issues involved in doing so. Design and technical challenges necessitated various tradeoffs. The system produced demonstrated the success of the design decisions: it was significantly easier to use than, and preferred to, a commercial equivalent by a group of older people with no experience of Internet use.

Keywords: Design; Older people; Internet; Email; Digital inclusion; Usability

1. Introduction

Standard computer technology is inappropriate for a large, and growing, proportion of the population: people over 60, who are, as a group, unlikely to use computers, less likely

* Corresponding author. Tel: +44 1382 384787; fax: +44 1382 345509.
E-mail address: adickinson@computing.dundee.ac.uk (A. Dickinson).

0953-5438/$ - see front matter © 2005 Published by Elsevier B.V.
doi:10.1016/j.intcom.2005.09.003

to know about computers, more likely to be anxious about using them and will find them harder to use than younger people do (Czaja, 1997; Hawthorn, 1998; Marquie et al., 2002).

In this paper the development and evaluation of an exemplar email system for older people is described. The study was collaboration between researchers and industry. We present our conclusions on the design of a specific interface and the methodological issues encountered evaluating a computer system with older adults with little or no experience of computers. The specific issues of communicating and working with an industry team are reported in detail elsewhere (Newell et al., n.d.).

2. Background

Inappropriate interface designs are a fundamental barrier to digital inclusion: older adults find standard interfaces harder to use than younger adults do even when computer experience is controlled for (Chadwick-Dias et al., 2003; Worden et al., 1997). Older people are also significantly more likely to be inexperienced computer users (ONS, 2004; Fox, 2004). Morris describes this divide as the 'technological alienation' of older people (Morris, 1992).

Lack of experience and support make it relatively more likely that older adults will have negative computer experiences, a significant factor in computer anxiety (Todman and Drysdale, 2004). Conversely, more encouraging experiences engender a more positive attitude to computers (Danowski and Sacks, 1980; Morris, 1992).

As inappropriate technology not only reflects but also exacerbates social exclusion (Valentine et al., 2002) it is important to challenge technological exclusion by finding effective ways to include older adults in the development process and by providing examples of appropriate technologies. Such examples are important because legal pressure (Sloan, 2001) alone cannot solve the problems of inaccessible software: there are genuine difficulties inherent in designing well for excluded groups. It is, for example, important to break away from the 'accessibility options' paradigm, which presupposes a dichotomy between 'able' and 'disabled' (Newell and Gregor, 1997).

The project undertaken was to develop an example system which would attract older people to use the Internet, and encourage them to progress to more sophisticated Internet use. The system was to be:

> "attractive to older users (over 60 years of age) who were uninitiated and unconfident in the use of computers and for whom the internet was an alien territory"

3. Email: previous research

Email, the internet application most used by older computer users (Coyne and Nielsen, 2002; SeniorNet survey April 2002; Age Concern, 2002), was selected as the most likely internet application to attract non-users. There are clear social reasons for email being attractive to older adults: networks of families and friends are increasingly widely spread

(Grundy, 1996) and social isolation can be associated with ageing (Russell and Schofield, 1999; Andrews et al., 2003).

Given its popularity and clear social benefits there has been surprisingly little research in the human–computer interaction field on email systems for older people. In 1993 Czaja et al., showed that an appropriately designed and very basic text-based communication system, developed for use within a pre-established network of older women, could be used 'with a minimal amount of difficulty' (Czaja et al., 1993). The study demonstrated that a reduced-functionality system could be used relatively independently by a group of older women, but it did not address all of the issues of an email system. For example, by using a pre-established network the research team were able to avoid the problems surrounding email addresses by providing a printed list of names to which participants sent messages.

Hawthorn's SeniorMail is a redevelopment of the Microsoft Outlook Express interface to support use by older adults. SeniorMail alters the visual presentation of the system (bigger buttons, larger font size, etc.) and has a list of possible actions presented in a simplified menu system, and simplified navigation (Hawthorn, 2002, 2003). This work is an adaptation rather than a system developed from first principles, and is arguably therefore influenced by Outlook Express. In addition, the system is designed to offer the functionality that relatively experienced users require and, as such, is unlikely to be optimally designed for beginners.

Arnott and Khairulla developed a prototype email system with a group of older users of varying experience, basing their initial prototyping on existing email interfaces, and reached conclusions similar to those reached by Hawthorn (Arnott et al., 2004). Central to both systems was reduced functionality and an initial menu system.

In other studies computers have been introduced into homes for the elderly and these studies initially appear to suggest that the introduction of computers has a positive effect on elderly users' wellbeing (see for example, McConatha et al., 1994). In studies where interpersonal communication was offered, this was the most popular and widely used application among the elderly users (Danowski and Sacks, 1980; McConatha et al., 1994). Unfortunately these studies used standard commercial software, which necessitated a high level of support. The improvements in wellbeing could not be sustained because it was not possible to continue providing support once the research was concluded (Sherer, 1997).

Although the literature indicates a fundamental need for an appropriate email system that can be used autonomously by older users with little computer experience, there has been little recent work on how such a system might be developed. Vitally, research in this field lacks a comparative evaluation of a specially designed system with a commercially available one, even though the raison d'etre of specially designed systems is the assumption that they are more appropriate than standard systems.

4. Design

Most older users are inexperienced computer users. Inexperienced users are likely to encounter difficulty with terminology (Janicki, 2002; Crystal, 2001) and with a range of interface conventions, including: when to double-click with the mouse, how to move between text input fields and how to use the scrollbar (Ellis and Kurniawan, 2000; Coyne

and Nielsen, 2002). Specific difficulties with standard systems include a failure to notice small or peripheral changes (e.g. selection of an email) and consequent uncertainty. Multiple Windows also present a barrier: inexperienced users clicking on a background Window often fear that they have lost all their work in a now-hidden Window.

Based on these difficulties and on research which demonstrates that older adults are likely to encounter barriers to internet use in terms of vision (Carmichael, 1999; Hawthorn, 2000; Echt, 2002), cognition (Czaja, 1997; Czaja and Lee, 2003; Salthouse and Maurer, 1996; Salthouse and Babcock, 1991; May et al., 1999) and manual dexterity (Ranganathan et al., 2001; Chaparro et al., 1999), a number of guidelines were agreed at the beginning of the project.

Level of functionality:

- Only essential functionality for a working email system to be included
- Each screen to have a very clear primary function.
- The number of actions / buttons per screen kept to a minimum (fewer than 10) (Miller, 1956).

Accessibility:

- Larger than average clickable targets (32 and 26 pt size recommended)
- Larger than average fonts (14 point as a minimum).
- High contrast choice of colours for text and background.
- Accessibility features compatible with the W3C guidelines.

User interface paradigms:

- Simple and very consistent select and operate paradigms.
- Clear conventions for the positions of buttons and information.
- No new or poorly established interface paradigms which were unlikely to be understood by the user group.
- Avoid scroll bars if possible, and definitely do not use nested scroll bars.

Terminology:

- Terminology which was understandable by the user group.

Personalisation:

- Some personalization to allow for people with poor eye sight or dexterity, for example the ability to easily increase text size.

Central to the design was the need for simple screens. Research indicates that older adults are likely to find it harder to make sense of a display (Newell et al., 2003) and will spend longer searching through an array of information (Fisk et al., 2004). The more complicated and cluttered the display, the worse the performance of older adults compared to younger ones; very complicated displays may present an insurmountable barrier

(Chadwick-Dias et al., 2003). This central principle caused a difficulty, however: with ageing, spatial memory declines and deep interface structures create significant barriers to use for older adults who appear to lose their place and forget the purpose of the task (Zaphiris et al., 2003). Clearly for interface design this poses a challenge: minimal density of screen information dictates a deep structure, but such structures have been shown to degrade performance in older adults.

4.1. Development

The early development process depended heavily on expert feedback. Evaluating very early prototypes with this group of users may be counter-productive as the process implicitly assumes a great deal of technical knowledge, which this user group does not have (Zajicek, 2004; Eisma et al., 2004).

Fig. 1 shows the 'Message Tray' screen for the initial prototype. Separate sections at the top of the screen provided 'orientation' information to the user, telling them where they were and offering hints about what they could do there. A preferences button led to customization options and beneath it two buttons offered screen-specific actions: compose message and address book. In addition, the user could open messages by clicking on them. A small scrollbar allowed the user to scroll through the messages in the message tray.

Paper prototypes of the design illustrated by Fig. 1 were presented to a workshop of nine older people aged between 65 and 84 who had little or no experience of computer use. The evaluation sessions led to the following developments:

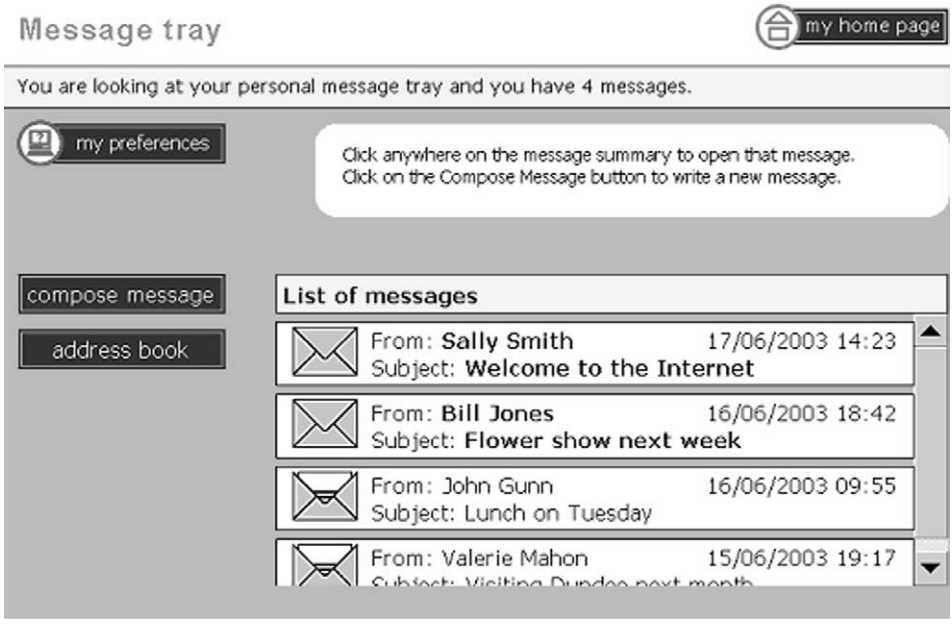

Fig. 1. Workshop design.

- Button text was made bolder and clearer, and changed to dark text on a pale background.
- The 'my preferences' button was moved from its intrusive position at the user's 'entry point' to the screen (Faraday, 2001).
- Language was both simplified and clarified. Although efforts had been made to avoid computer terminology, the language used was perceived as too formal by the workshop participants.
- The two 'orientation' panels were combined into a single instructions panel with an eye-catching icon.
- Terms like 'my home page' were removed because of uncertainty about who 'my' referred to.

4.2. Final Prototype

A final prototype email system was developed which conformed to as many of the initial design constraints as had been found to be technically feasible within the development period.

Fig. 2 above represents the email system. Functionality was conceived of as related to either reading or to sending a message.

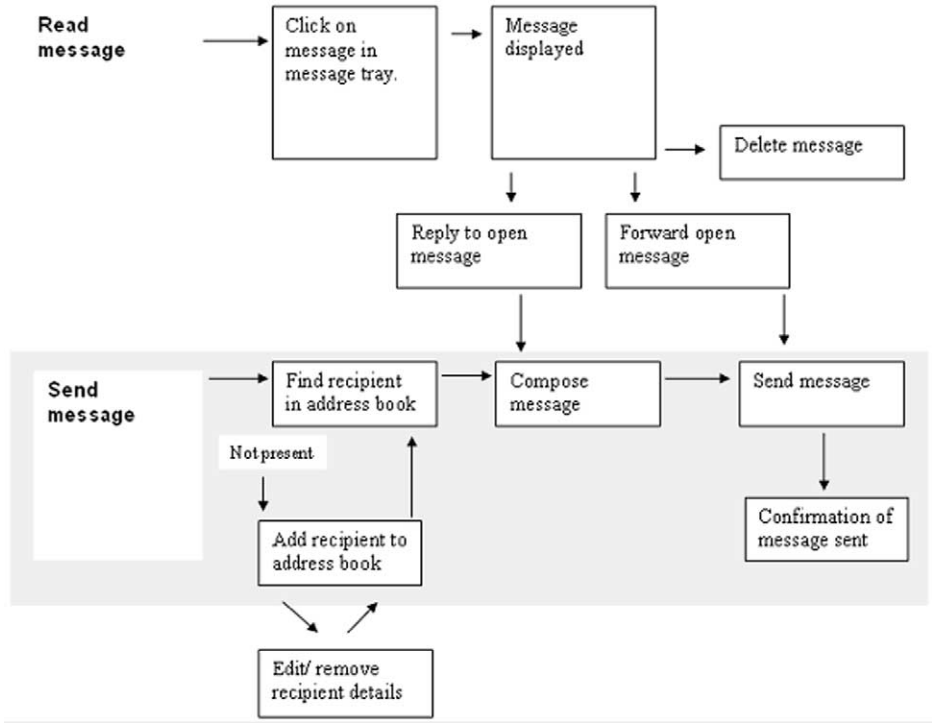

Fig. 2. Plan of email system.

4.2.1. Radically simple design

The design constraints had specified a simplification of the system, not only in terms of reduced functionality, but also in the presentation of information on each screen. The constraint that each screen should have a "clear primary function" defined what was possible.

Pages with a single primary function made it possible to reduce dependence on user knowledge, because the system could anticipate what the user would be likely to do on that page, e.g. enter the email address of the recipient. Single-use pages allowed a single text input field per page, removing the difficulties associated with uncertainty about where text would appear, and the problems of moving between multiple text-entry fields. In addition, mouse interaction was limited to a single click that always led to a new page. This provided clear feedback to the user that their action had had a result, and removed problems with noticing small changes, e.g. not realizing that text is selected.

Single-use pages, however, forced a deep interface structure, which raised the possibility of users becoming lost in the interface and forgetting their task. This was addressed by providing support for the user throughout the interface (see Fig. 3). The single-purpose screens made it possible to have screen-specific help which supported the user as they were guided along a linear path to completing the task.

Because the path was largely restricted to moving forwards (to proceed) or backwards (to correct), it was possible to present all screens in kiosk (full screen) style, avoiding the problems of multiple, overlapping Windows.

The ability to anticipate normal user movement through the page allowed the structuring of pages to support visual scan and a sensible progression. So, for example, the "send message" button was beneath the field for the message entry, making it the natural progression in terms of visual movement through the page from the top left (Fig. 4).

Knowledge about what the user would be likely to do on a single page also made it possible to anticipate likely points of weakness. One such point is the entry of email addresses, which are long, apparently nonsensical character strings, with sections separated by a small and unobtrusive piece of punctuation. On the email address entry page the system responded to common errors by providing instructions about what was wrong and how to correct it.

Simple, single-purpose pages made it possible to have larger default text (14 point) and buttons (32 point), and also to allow personalization options which would further increase them in size. In addition, it was possible to have slightly more text on the larger buttons so that button text normally made explicit the subject of the action; thus, for example "send message" rather than "send". Although allowing user personalization to take place gracefully, without impairing the coherence or aesthetics of a page, represented a challenge for the developers, it could only be done in this way because of the simplicity of the pages.

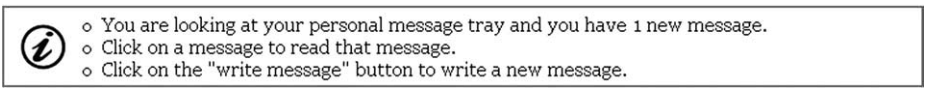

Fig. 3. Instructions to users on the Message Tray screen.

Fig. 4. Compose Message screen.

There were other aspects in the interface, which were not immediately connected to the single-use pages decision. These included relatively high contrast between text and background, capitalization of user names, and all emails and address book entries were represented as large buttons with a border to remove the ambiguity about what could be clicked on.

The system was a 'radically simple' interface, where single-purpose screens defined the user's interaction with the system by allowing the user to make a high level decision, for example 'write a message to Agnes', and supporting them throughout the process of carrying it out. Much of the responsibility for task success was given to the system.

The system had been designed specifically to be usable for older users, involving older users in its design. It was necessary to evaluate whether the design decisions made had been successful and would allow more autonomous and confident user interaction with the experimental system than with a commercial equivalent.

5. Evaluations

The Cybrarian system was evaluated in a controlled environment over a 2-week period with 15 older people. Each user attended two sessions of an hour and a half and on each occasion performed a series of email tasks with the experimental system and a comparative

Table 1
Participants by age and gender

	Female	Male	All
60–64	4	2	6
65–74	4	1	5
75–84	1	2	3
85+	–	1	1
All	9	6	15

system. The order in which people used the experimental and control system was controlled using a Latin square experimental design (ABBA) to avoid order effects.

Outlook Express, a system commonly provided by Internet Service Providers, was chosen as the control system.

5.1. Participants

Older people are an extremely diverse group; a User Sensitive Inclusive Design (Newell and Gregor, 2000; Gregor and Newell, 2001) methodology was employed, users were selected for study-relevant age-related characteristics.

Participants were recruited from a number of sources, including the UTOPIA project database. Fifteen participants were selected with minimal computer experience, lack of familiarity with the Internet and a combination of minor age-related impairments (Table 1).

All the participants had corrected vision, two had hearing aids, two had had strokes which left some minor motor control impairment and some concentration and memory difficulties, and another two participants had minor fine motor control impairment. Four people had other self-reported medical conditions.

Participants were paid £20 expenses per session.

5.2. System setup and apparatus

The apparatus used was a PC with 2.4 GHz processor, 512 Mb memory, 40 Gb disk, CDROM, 15″ LCD monitor, 10/100 Mbps UTP NIC running Windows XP Pro and the Internet Explorer 6 browser. Recording equipment consisted of a digital video camera and a dictaphone.

The system used was based on a POP3/SMTP server, which supported the email functionality. Emails were stored in an SQL database, which also contained information on the status of the emails (read/ unread).

Each user had an individual profile, which was available when they arrived for the first visit. Each email profile was set up identically with several messages already in the message tray, one of which was a ‘read’ message. Similarly, an identical list of ‘contacts’ was available in each of the address books. Among the contacts were fictional characters who featured in the task list, for example, participants were asked to send an email to an Agnes Kelly.

Outlook Express, the control system, was set up as it would be ‘out of the box’. It is unlikely that an inexperienced computer user will deliberately alter the set up of a program

so this was the most accurate and natural set up. As it comes 'out of the box' Outlook Express has five panes, including the preview pane and the contacts pane. Contacts and messages in the inbox were identical in both conditions.

5.3. Visit structures

The session facilitator met the participant and guided them to the study room. The notetaker was introduced, an informed consent form was discussed and signed, and a short, informal discussion about computer use took place. In order to keep the environment relaxed only three people were in the room: the participant, the facilitator and a note taker. All sessions were videoed and audio-recorded with the participant's permission.

No training was given. Participants were told that they were being asked to try two different ways of sending messages using a computer. Throughout the sessions emphasis was placed on evaluating the two systems and it was made clear to the participant, repeatedly if necessary, that the *systems* were being evaluated. Neither system was identified as the experimental system and the facilitator denied any connection with either of the systems. This approach was intended to avoid, as far as possible, attempts to please the facilitator by praising 'their' system.

5.4. Tasks

Participants were asked to complete six tasks on each system in the first visit, and seven on each system in the second visit. Each task was read from an A4 card by the facilitator and the card was placed beside the monitor. Tasks explored aspects of the interaction on levels of increasing complexity.

	Visit 1	Visit 2
1	Please read the message from Sally Smith. Please reply briefly to the message from Sally Smith.	Open a message.
2	Invite John Gunn to come to your home for lunch.	Reply to a message.
3	Look at another new message and respond to it appropriately.	Send a message to Agnes Kelly.
4	Send an email to Agnes Kelly.[a]	Send a message to one of the people on this card.
5	Add the details of one of the people listed on this card to the address book.[b]	Forward a message.
6	Send a message to the person or friend you just added to the address book.	Delete a message.
7	–	Edit someone's details in your address book.

[a] Agnes Kelly was a contact in the address book from whom there was no message in the message tray.
[b] If the participant had brought the email address of someone else, this was used instead.

5.5. Protocol

Facilitators worked according to a script and a set of instructions. Notetakers also worked to instructions. A training period and 'dry run' took place. All notes were analysed by the study co-ordinator to check consistency. Where there was any uncertainty, videos of the sessions were referred to.

5.6. Measures

The central component tested was:

1. Was the user able to carry out tasks using the system?
   This was tested by recording:
   - ○ Tasks completed autonomously
   - ○ The number and frequency of errors, hesitations, requests for help and facilitator interventions.

Two secondary components were:

2. Which was the preferred system?
   This was evaluated with a 7-point Likert scale and a semi-structured interview.
3. Was there evidence that users would continue to use the system?

A semi-structured interview included questions on whether users would continue to use either system.

All sessions took place using a think aloud protocol and notetakers also recorded both verbal and non-verbal participant reactions to the systems, for example:

> Puts hand to mouth—he drops contact with mouse and keyboard.
>
> "Why am I not getting anything there? Am I not pressing it hard enough or something?"
>
> (user 14 notes. Visit 2, session 2)

6. Results and discussion

6.1. Autonomous usability

A critical quantitative measure of success in terms of the usability of the system was the number of tasks that a user managed to complete autonomously, without any intervention from the facilitator. As the tasks were designed to be progressive in both complexity and the level of system familiarity required, comparisons will focus on success rates across the different tasks.

Table 2 shows the number of tasks that participants completed without help from the facilitator during both visits. A 2 (System) $\times$ 2 (Visit) Analysis of Variance was conducted

Table 2
Tasks completed without intervention (max score in each cell = 15)

	Visit 1		Visit 2	
	Cybrarian	Control	Cybrarian	Control
Task 1	14	8	13	9
Task 2	13	7	14	9
Task 3	14	7	14	8
Task 4	10	7	11	8
Task 5	10	7	8	6
Task 6	7	6	12	8
Task 7	–	–	12	3
Mean	11.33	7.00	10.14	7.29

on the data to determine any difference in performance between using the Cybrarian system and using Outlook Express. The tasks set for the first visit, although comparable, did differ from those presented during the second visit, and hence Visit will be treated as a between-factor in the ANOVA.

A main effect of System [$F(1,11) = 9.366$, $p < 0.05$] indicates that more participants successfully completed the range of different tasks using the Cybrarian software (each task averaging a success rate of 10.69) rather than the Control (only 7.15; a little under half the participants on average). Performance did not alter significantly between visits [$F(1,11) = 0.162$, $p > 0.5$] and the lack of any interaction [$F(1,11) = 0.395$, $p > 0.5$] between the two variables confirms that the benefit of using the Cybrarian software was not restricted to either the first or second visit in particular.

A Spearman's correlation between task order and performance shows that increasing task complexity did reduce the number of participants successfully completing the tasks using Outlook Express on both visits [$r_s(6) = -0.845$, $p < 0.05$ and $r_s(7) = 0.879$, $p < 0.01$, respectively]. On the other hand, while there was a similar negative correlation for Cybrarian on the first visit [$r_s(6) = -0.883$, $p < 0.05$], there was no evidence of any correlation on the second visit [$r_s(7) = -0.090$, $p > 0.5$]. The Cybrarian system therefore appears to have minimised any negative effects associated with task complexity and required operator knowledge by the end of the experiment, indicative of faster learning and greater familiarity.

6.2. Error rates

The frequency of errors committed was another important component of participants' experience in using the email systems. For the purposes of recording error rates, errors were classified as either minor or significant. Minor errors, such as choosing an incorrect option while attempting to complete a task, are an indication of how well the participant understood the interface as well as how clearly they understood the purpose of the options available to them and the ways in which to carry out the tasks. Significant errors involve a more serious mistake, which would typically have caused the task to fail, implying a critical failure to interact correctly with the interface.

Table 3
Minor (ME) and significant (SE) errors

Task	Cybrarian					Control				
	N	ME		SE		N	ME		SE	
	Visit 1									
1	15	4	*0.27*	0	*0.00*	15	64	*4.27*	11	*0.73*
2	14	3	*0.21*	0	*0.00*	14	18	*1.29*	3	*0.21*
3	14	1	*0.07*	0	*0.00*	14	17	*1.21*	6	*0.43*
4	14	14	*1.00*	0	*0.00*	13	67	*5.15*	6	*0.46*
5	14	19	*1.36*	1	*0.07*	10	17	*1.70*	3	*0.30*
6	11	9	*0.82*	0	*0.00*	7	7	*1.00*	0	*0.00*
Mean	13.67	8.33	*0.62*	0.17	*0.01*	12.17	31.67	*2.44*	4.83	*0.36*
	Visit 2									
1	14	2	*0.14*	0	*0.00*	14	37	*2.64*	6	*0.43*
2	14	3	*0.21*	1	*0.07*	14	17	*1.21*	2	*0.14*
3	14	1	*0.07*	0	*0.00*	14	46	*3.29*	3	*0.21*
4	14	18	*1.29*	1	*0.07*	13	28	*2.15*	2	*0.15*
5	11	9	*0.82*	0	*0.00*	10	13	*1.30*	5	*0.50*
6	12	3	*0.25*	0	*0.00*	10	1	*0.10*	1	*0.10*
7	12	9	*0.75*	1	*0.08*	10	37	*3.70*	7	*0.70*
Mean	13.00	6.43	*0.50*	0.43	*0.03*	12.14	25.57	*2.06*	3.71	*0.32*

N denotes the sample size attempting the task.

It is worth emphasising that one of the reasons for the reduction in errors in later tasks, particularly with the Control email system, is that fewer people attempted the later tasks. As Table 3 shows for example, during Visit 1, 11 participants attempted Task 6 using the Cybrarian system, whereas only 7 out of the 15 participants were able to begin the identical task using the Control system. Therefore, in absolute terms, more errors were recorded on Cybrarian's later tasks, partly because more people tried them. In order to compensate for this and enable more meaningful comparisons, a standardised measure of *errors per user* is given in italics in Table 3. Errors per user were significantly lower using the Cybrarian system [$F(1,11)=17.423$, $p<0.01$]; indeed, the rate was often negligible under Cybrarian.

6.3. User hesitations

The number of times people hesitated when using a system is another indication of user competence and confidence. A hesitation was recorded if a participant paused whilst performing their prescribed task. These could either be due to a delay in decision-making or the need to scan the interface in greater depth for information that was not immediately apparent. Significant hesitations were recorded when the participant paused for an extended period of time and there was greater user-awareness of a problem; quite often these hesitations were accompanied by the participant sitting back from the monitor and indicating that they were confused or lost. The recording procedure was fine-tuned to take the interaction style of individual participants into account. For example, some participants would pause while they explained an action they had just taken as they 'thought aloud' and these were not recorded as

Table 4
Minor and significant hesitations

Task	Cybrarian					Control				
	N	MH		SH		N	MH		SH	
	Visit 1									
1	15	7	*0.47*	0	*0.00*	15	17	*1.13*	64	*4.27*
2	14	4	*0.29*	0	*0.00*	14	15	*1.07*	3	*0.21*
3	14	2	*0.14*	0	*0.00*	14	18	*1.29*	3	*0.21*
4	14	24	*1.71*	3	*0.21*	13	41	*3.15*	9	*0.69*
5	14	17	*1.21*	1	*0.07*	10	8	*0.80*	5	*0.50*
6	11	5	*0.45*	0	*0.00*	7	3	*0.43*	0	*0.00*
Mean	13.67	9.83	*0.71*	0.67	*0.05*	12.17	17.00	*1.31*	14.00	*0.98*
	Visit 2									
1	14	1	*0.07*	2	*0.14*	14	14	*1.00*	17	*1.21*
2	14	5	*0.36*	4	*0.29*	14	13	*0.93*	1	*0.07*
3	14	12	*0.86*	2	*0.14*	14	11	*0.79*	22	*1.57*
4	14	17	*1.21*	6	*0.43*	13	15	*1.15*	8	*0.62*
5	11	13	*1.18*	5	*0.45*	10	6	*0.60*	4	*0.40*
6	12	1	*0.08*	0	*0.00*	10	1	*0.10*	0	*0.00*
7	12	4	*0.33*	2	*0.17*	10	10	*1.00*	19	*1.90*
Mean	13.00	7.57	*0.59*	3.00	*0.23*	12.14	10.00	*0.80*	10.14	*0.82*

N denotes the sample size attempting the task.

hesitations. Recorded hesitations are therefore restricted to pausing directly related to a user's *general interaction with the interface*.

Again, to enable a more meaningful comparison, a standardised measure of *hesitations per user* is included in italics in Table 4. Given that the Cybrarian system was a simpler design and there was therefore less to look at on the screen, the significantly smaller number of hesitations per user that it produced [$F(1,11)=7.981$, $p<0.05$] is arguably unsurprising. Nonetheless, the number of hesitations must be regarded as a significant component of a user's interaction with a system and there is a stark contrast in the results between the two systems: using the control system led to higher levels of uncertainty, a greater lack of confidence and greater confusion.

Measuring hesitations while using the think aloud technique can potentially cause confusion between hesitations due to confusion with the interface, and those due to consideration of how to *articulate* responses to the interface. While, in practice, distinguishing these is rarely problematic, the within subjects balanced design used in this study highlighted differences *between* conditions, and a significantly greater number of hesitations while using Outlook Express which cannot easily be attributed to the think aloud protocol in isolation.

6.4. Requests for help and facilitator interventions

Participants were instructed not to request help unless they were genuinely stuck and unable to proceed, and would therefore have to abandon the task without help from the facilitator. Requests for help are consequently interpretable as a more extreme form of hesitation, indicating a serious difficulty in understanding how to operate the system

Table 5
Requests for help (RH) and facilitator interventions (FI)

Task	Cybrarian					Control				
	N	RH		FI		N	RH		FI	
	Visit 1									
1	15	5	*0.33*	6	*0.40*	15	30	*2.00*	40	*2.67*
2	14	0	*0.00*	1	*0.07*	14	10	*0.71*	14	*1.00*
3	14	2	*0.14*	2	*0.14*	14	4	*0.29*	10	*0.71*
4	14	3	*0.21*	9	*0.64*	14	22	*1.57*	24	*1.71*
5	14	3	*0.21*	12	*0.86*	14	13	*0.93*	9	*0.64*
6	11	1	*0.09*	7	*0.64*	11	8	*0.73*	9	*0.82*
Mean	13.67	2.33	*0.17*	6.17	*0.46*	13.67	14.50	*1.04*	17.67	*1.26*
	Visit 2									
1	14	2	*0.14*	1	*0.07*	14	6	*0.43*	12	*0.86*
2	14	1	*0.07*	4	*0.29*	14	9	*0.64*	11	*0.79*
3	14	1	*0.07*	1	*0.07*	14	14	*1.00*	19	*1.36*
4	14	4	*0.29*	6	*0.43*	13	8	*0.62*	14	*1.08*
5	11	0	*0.00*	3	*0.27*	10	5	*0.50*	14	*1.40*
6	12	1	*0.08*	1	*0.08*	10	2	*0.20*	3	*0.30*
7	12	1	*0.08*	1	*0.08*	10	3	*0.30*	9	*0.90*
Mean	13.00	1.43	*0.11*	2.43	*0.19*	12.14	6.71	*0.53*	11.71	*0.95*

N denotes the sample size attempting the task.

and a diminished level of confidence in using the interface. Table 5 lists the Request for Help data, including the *requests per user* in italics. Requests relating to technical issues not specific to the software being tested, such as those concerned with the position of keys on the keyboard (for example, whether to press the *enter* key to get a new line), were made using both interfaces but were not specifically recorded in this experiment. At their own discretion, a facilitator could spontaneously offer guidance in order to minimise participant frustration and distress, or if the participant had become obviously lost or sidetracked; these are recorded as Facilitator Interventions.

The need for any form of help from the facilitator was significantly lower for the Cybrarian system [$F(1,11)=27.103$, $p<0.001$]. Additionally, while there were some participants who were more inclined to ask for help than others, all participants requested more help when they were using the control system rather than Cybrarian. The level of support did not differ significantly between visits [$F(1,11)=3.184$, $p>0.05$] and the difference between the two systems remained constant for both visits [$F(1,11)=0.761$, $p>0.05$]. Thus, overt and conscious indicators of difficulty also reveal that fewer problems are encountered with Cybrarian.

6.5. Ratings

After each series of tasks was carried out with a system, participants were asked to rate the system they had just used for Ease of Use, Pleasantness and Ease of Remembering on a 7-point scale where 1 was a strongly positive response and 7 a strongly negative response. The ratings contrast between the two systems was analysed using Wilcoxon Signed Rank

Table 6
Mean ratings for Cybrarian and control system and statistics for Wilcoxon Signed Rank Test

	Ease of use		Pleasantness		Ease of remembering	
	Visit 1	Visit 2	Visit 1	Visit 2	Visit 1	Visit 2
Cybrarian	2.07	1.86	1.93	1.5	2.27	2.21
Control	4.33	4.71	3.33	3.79	3.93	4.71
Difference	−2.26	−2.85	−1.4	−2.29	−1.66	−2.50
Wilcoxon *Z*-Score	−3.205	−3.192	−2.992	−3.211	−2.989	−2.992
Probability	0.001	0.001	0.003	0.001	0.003	0.003

Tests. The results indicated a statistically significant preference for Cybrarian on all three measures across both visits. The means and mean differences are displayed in Table 6, along with the standardised *Z*-scores and probabilities derived from the Wilcoxon tests.

The quantitative measures consistently indicated that Cybrarian was used more autonomously and with less evidence of uncertainty or confusion as measured by observed behaviour (errors and hesitations) and by the user's own judgement (requests for help and preferences ratings). The difference between the Cybrarian system, specifically designed for this group, and the control system was thus both reliable and pronounced.

6.6. Qualitative and subjective responses

At the end of each visit a short semi-structured interview took place to explore user reaction to the systems. Responses again indicated a strong preference for Cybrarian. In addition, respondents clearly would have used the system again: all of the participants who answered the question "Would you want to use Cybrarian again?" ($n=12$) said that they would, often very decisively. Participants also reported that they felt more inclined to explore other things that computers could do after they had used Cybrarian: " I would feel confident if it [a different computer application] was like that e-mail system [Cybrarian]" and "Once I became confident with the email I could move on to other uses of the computer—to take the next step…if it was like the mail of Cybrarian I would certainly use it!" This response confirms Morris' prediction that a positive initial experience with software encourages users to continue using computers (Morris, 1992 p.75).

The subsequent phase of the Cybrarian POC involved the development and evaluation of a web search and navigation system with a similar interface to the email system; this system is currently the basis for the UK government's MyGuide system. We hope to report upon the development and evaluation of the system in a later paper.

7. Methodology notes

7.1. Thinking aloud

Thinking aloud was a methodology that had been successfully used by the industry usability engineers on the project; in general its use with the older participants was successful

but task difficulty was negatively correlated with the ability to think aloud coherently. One example of this is Participant 2 struggling to reply to a message using Outlook Express:

Task1: Visit 1 (Session2)
FAC: I was wondering why you were clicking on that [inbox]… if you just tell me what you're thinking…
P clicked that…and I…went here…
FAC …yes…
P: …to uh… because I thought… I was trying to… this would bring up the message
FAC: right…
long pause
FAC: can I ask what you're looking for?
P: [sighs irritably] next step after… having done…
FAC: ok
P: now there's obviously a step that I've got to take that…
long pause
P: no… [sits back from the computer and drops the mouse]

The use of the think aloud protocol with older users can generate rich and interesting information, but will increase the cognitive complexity of the task. For participants with cognitive impairments, for example, those who have suffered a stroke, it may be preferable to separate, as far as possible, the process of "thinking aloud" from the first attempt at a task (see Fisk et al., 2004). We have recently used such a methodology in an eye-movement study (see Hill et al., n.d.) with considerable success.

7.2. Divided attention

Interface complexity also appeared to have an impact on some users' ability to use the keyboard and interface elements. Participant 8 repeatedly forgot small details about using the computer and the keyboard when she was using the control system:

"I can't remember…" [about subject lines]
Significant hesitation— P is confused
Request for help— she stops being able to use the scroll bar
FAC intervenes
Writes message in caps (has forgotten to remove them)
Significant hesitation to send— can't find send button
…
Forgets RTRN
Minor hesitation "now I've lost my…"
FAC prompts
"I've lost how to send it!"

This effect of task complexity interfering with other activities has been documented (Hawthorn, 2000), but identifying it can be difficult without a comparison system.

Such small errors, apparently unrelated to the system comparison, were not consistently recorded, but observations like these suggest that a more holistic approach to the user experience might yield interesting supporting information.

7.3. Comparison system

Working with this user group has well-documented difficulties (Eisma et al., 2004). One fundamental barrier to successful communication is the participants' lack of knowledge about the area, which makes it very difficult to carry out traditional user-centred design activities or to get useful subjective feedback from users about a system under evaluation. Using a comparison system provides one way to overcome these barriers, showing users that one task (for example, sending an email) can be achieved in more than one way. Using two different systems allowed users to compare and, vitally, using Cybrarian successfully gave them the confidence to criticise Outlook Express.

One participant commented: "using Cybrarian would encourage me to use a computer far more. Using Outlook Express only may well have put me off - although not having seen Cybrarian I would have had nothing to compare it to. I would have thought 'this is something I need to get used to…'."

8. Interface tensions

The evaluation demonstrated formally that, for older novice users, a specially designed system is more usable and attractive than a commercial system. The comparative evaluation illustrates, and confirms, the value of such research in confirming the assumption that standard systems do present considerable barriers for this user group. In addition, the fact that the system was developed in an industrial setting with insufficient time for high levels of user involvement and evaluation suggests that, while it is certainly non-optimal, it is possible to successfully transfer knowledge and aspects of methodology to an industrial setting.

The development process crystallised a number of tensions that exist in developing appropriate software for older novice users. These tensions constitute an important aspect of the process and are discussed briefly here.

The first tension encountered was the desirability of simple, one-purpose screens. The use of such screens dictates a deep structure, which has been shown to be difficult to navigate. The Cybrarian development team addressed this problem by providing onscreen user support throughout. The efficacy of this approach needs to be further evaluated in more controlled circumstances but initially appears to offer a solution.

Second, there is a tension between the reduction of functionality and support of the user. Functionality-laden systems reduce usability dramatically, especially for older users. However, the logical opposite is a system that only allows the sending and receiving of messages. Such a system also places excessive pressure on the user, in this case by making it necessary for them, for example, to remember email addresses each time they want to send an email. Finding the balance between sufficient and excessive functionality is not a simple process.

Providing a process structured as a guided path works well for beginners but the trade off is that the processes are cumbersome, offering security and support rather than immediacy. Within the one and a half hours that Cybrarian participants spent with the system, some began to find aspects of it intrusive rather than helpful. Although these users represented a minority of the evaluators this suggests that such 'training' pathways may reach obsolescence fairly rapidly for some users, but may retain their usefulness for others, especially for those with mild cognitive impairments.

The Cybrarian system was designed to support users in progressing to 'mainstream' applications. Although it is not inevitable that all users will want to do so, some will. One difficulty encountered was providing appropriate terminology to allow a 'bridge' between the learning system and mainstream applications. Later stages of the Cybrarian evaluations asked participants to 'forward' a message and to 'edit' details in the address book. User responses to these terms were strongly negative, criticizing them as 'computer jargon'. There appears to be, therefore, a tension in introducing mainstream language within a training environment. It seems likely that this can be overcome with appropriate support, but should be approached with caution.

9. Conclusion

These evaluations demonstrated that a specially designed system, Cybrarian, was significantly more usable than an industry-standard equivalent. The positive experience of the Cybrarian system led participants to be more confident about computer use generally, and about Internet use. To have taken a group of older people with little knowledge of computers and a perception of the Internet as 'alien' and received such positive responses to the experimental system, and demonstrations of successful use, suggests that it is possible to develop appropriate systems for many of the people who are currently unlikely to use the Internet.

This is only the first step in investigating this area: the Cybrarian system differed from the commercial system in many ways, including a simplified interface, reduced clutter on the screen, reduction of terminology, clear and simple navigation paths, and a particular type of 'help'. More work is needed to identify what the most important parameters were.

Many of the characteristics of the current Cybrarian system are also appropriate for other non-users. A system that demands no prior knowledge and provides instructions and supportive feedback is useful for a wide range of the people who do not currently use computers: non-users of all ages are likely to be helped by a system that makes no assumptions about their knowledge. As well as being helpful for novice users, the Cybrarian system is also likely to be useful for those with mild cognitive or motor control impairments, while the existing interface and the easy availability of options to change the visual parameters make it useful also for those with visual impairments, including, in some cases, people with dyslexia (Wilkins, 1995; Keates, 2000).

While expert users are likely to demand greater functionality than that currently provided by the Cybrarian system this demand could be addressed with a layered interface where experts were responsible for adding what they wanted but novices were not faced with

the full complexity of the system in the way that they currently are. Systems that are rarely used might benefit from the sort of onscreen support offered by the Cybrarian email system.

If older people and others who are non-users of the Internet are to be encouraged to explore its benefits, applications must be available which offer clear value. While attractor applications are necessary, they are not sufficient. Such applications must also provide an encouraging, rather than intimidating, introduction to computer, or Internet, use. Both research and experience indicate that initial impressions are very important in deciding whether or not people will persevere with computers. In our experience, once learners have developed confidence, they are capable of further exploration: it is in the initial stages that people can be discouraged for good.

Systems like Cybrarian represent first steps in attracting people and encouraging them to explore further. The real challenge of producing the Cybrarian email system was to hide the complexity of the functionality. Companies selling commercial software normally emphasise the new functionality that it offers, and measure success in terms of the new options being offered. One of the most difficult processes during the development of the system was recognising when our preconceptions about what was 'necessary' for an email system allowed complexity to creep into the interface designs. There is no doubt that contemporary computers are powerful enough to hide most of their excessive complexity from the user, but it is equally true that this involves a paradigm shift in managers', designers' and developers' understanding of what a good computer system is, and a recognition that a market exists for appropriately design, simple interfaces.

Acknowledgements

We would like to thank SHEFC (the Scottish Higher Education Funding Council) for funding the UTOPIA project and the DFES for Funding Cybrarian.

References

Age Concern, 2002. IT, Internet and older people: management survey, Age Concern 2002.

Andrews, G.J., Gavin, N., Begley, S., Brodie, D., 2003. Assisting friendships, combating loneliness: users' views on a 'befriending' scheme. Ageing and Society 23, 349–362.

Arnott, J.L., Khairulla, Z., Dickinson, A., Syme, A, Alm, N., Eisma, R., Gregor, P., 2004. Email Interfaces for Older People. Proceedings of the IEEE SMC International Conference on Systems, Man and Cybernetics, The Hague, The Netherlands, October 10–13.

Carmichael, A., 1999. Style Guide for the Design of Interactive Television Services for Elderly Viewers. Independent Television Commission, London.

Chadwick-Dias, A., McNulty M., Tullis T., 2003. Web usability and age: how design changes can improve performance. Proceedings of the ACM Conference on Universal Usability, Vancouver, Canada, November 16–17, 2003, pp. 30–37.

Chaparro, A., Bohan, M., Fernandez, J., Choi, S.D., Kattel, B., 1999. The impact of age on computer input device use: psychophysical and physiological measures. International Journal of Industrial Ergonomics 24, 503–513.

Coyne, K.P., Nielsen, J., 2002. Web Usability for Senior Citizens: Design Guidelines based on Usability Studies with People Age 65 and Older. Nielsen Norman Group.

Crystal, D., 2001. Language and the Internet. Cambridge University Press, Cambridge.

Czaja, S.J., 1997. Computer technology and the older adult. In: Helander, M.G., Landauer, T.K., Prabhu, P.V. (Eds.), Handbook of Human–Computer Interaction. Elsevier, Amsterdam, pp. 797–812.

Czaja, S.J., Lee, C.C., 2003. Designing computer systems for older adults. In: Jacko, J.A., Sears, A. (Eds.), The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications. Lawrence Erlbaum Associates, pp. 413–427.

Czaja, S.J., Guerrier, J.H., Nair, S.N., Landauer, T.K., 1993. Computer communication as an aid to independence for older adults. Behaviour and Information Technology 12 (4), 197–207.

Danowski, J.A., Sacks, W., 1980. Computer Communication and the Elderly. Experimental Aging Research 6 (2), 125–135.

Echt, K.V., 2002. Designing web-based health information for older adults: visual considerations and design directives. In: Morrell, R.W. (Ed.), Older Adults, Health Information and the World Wide Web. Lawrence Erlbaum Associates, pp. 61–87.

Eisma, R., Dickinson, A., Goodman, J., Syme, A., Tiwari, L., Newell, A.F., 2004. Early user involvement in the development of information technology related products for older people. Universal Access in the Information Society 3 (2), 131–140.

Ellis, R.D., Kurniawan, S.H., 2000. Increasing the usability of online information for older users: a case study in participatory design. International Journal of Human-Computer Interaction 2 (12), 263–276.

Faraday, P., 2001. Attending to web pages. Proceedings Computer-Human Interaction conference, Seattle, Washington, USA, 31 March–5 April, pp. 159–160.

Fisk, A.D., Rogers, W.A., Czaja, S.J., Charness, N., Sharit, J., 2004. Designing for Older Adults: Principles and Creative Human Factors. CRC Press, Boca Raton.

Fox, S., 2004. Older Americans and the Internet. Report for Pew Internet and American Life Project, available online at http://www.pewinternet.org/.

Gregor, P., Newell, A.F., 2001. Designing for dynamic diversity—making accessible interfaces for older people. Proceedings of the EC/NSF Workshop on Universal Accessibility of Ubiquitous Computing: Providing for the Elderly, 22–25 May, Portugal, 2001, pp. 90–92.

Grundy, E.M.D., 1996. Population Review (5): The population aged 60 and over. Population Trends 84, 14–20.

Hawthorn, D., 1998. Cognitive ageing and human computer interface design. Proceedings of the Australasian Conference on Computer Human Interaction, Adelaide, Australia, 29 November 4, 1998, p. 270.

Hawthorn, D., 2000. Possible implications of aging for interface designers. Interacting with Computers 12 (5), 507–528.

Hawthorn, D., 2002. Designing usable applications for older users—an example. Proceedings of Human Factors, Melbourne, Australia, 25–27 November.

Hawthorn, D., 2003. How universal is good design for older people? Proceedings of the ACM Conference on Universal Usability, Vancouver, Canada, November 16–17, 2003, pp. 38–45.

Hill, R.L., Dickinson, A., Gregor, P, McIver, L., Milne, S., Nind, T., (in preparation). Making Sense of Content: using eye-tracking with older, novice computer users to evaluate an alternative web browser design.

Janicki, K., 2002. A hindrance to communication: the use of difficult and incomprehensible language. International Journal of Applied Linguistics 12 (2), 194–217.

Keates, A., 2000. Dyslexia and Information and Communications Technology: A Guide for Teachers and Parents. David Fulton Publishers, London.

Marquie, J.C., Jourdan-Boddaert, L., Huet, N., 2002. Do older adults underestimate their actual computer knowledge? Behaviour and Information Technology 21 (4), 273–280.

May, C.P., Hasher, L., Kane, M., 1999. The role of interference in memory span. Memory and Cognition 27, 759–767.

McConatha, D., McConatha, J.T., Dermigny, R., 1994. The use of interactive computer services to enhance the quality of life for long-term care residents. The Gerontologist 34 (4), 553–556.

Miller, G.A., 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. The Psychological Review 63, 81–97.

Morris, J.M., 1992. The effects of an introductory computer course on the attitudes of older adults towards computers. Proceedings of the 23rd SIGCSE technical symposium on Computer science education, Kansas City, Missouri, USA, 5–6 March, 1992, pp. 72–75.

Newell, A.F., Gregor, P., 1997. Human computer interfaces for people with disabilities. In: Helander, M.G., Landauer, T.K., Prabhu, P.V. (Eds.), Handbook of Human Computer Interaction. Elsevier, North Holland, p. 813.

Newell, A.F., Gregor, P., 2000. User sensitive inclusive design in search of a new paradigm. (Proceedings of the ACM Conference on Universal Usability, Arlington, VA, USA, 16–17 March, 2000) pp. 39–44.

Newell, A., Carmichael, A., Gregor, P., Alm, N., 2003. Information technology for cognitive support. In: Jacko, J., Sears, A. (Eds.), Handbook of Human Computer Interaction. Erlbaum, Mahwah, NJ, pp. 464–481.

Newell, A.F., Dickinson, A., Smith, M and Gregor, P. (under revision). Designing a Portal for Older Users: A Case Study of an Industrial/Academic Collaboration, under revision for ACM Transactions on Computer–Human Interaction.

Office of National Statistics (UK), 2004. Adults who have ever used the Internet by sex/age (Great Britain) figures for April 2004. Downloaded from: http://www.statistics.gov.uk/statbase/ssdataset.asp?vlnk=6928&More=Y on December 18 2004.

Ranganathan, V.K., Siemionow, V., Sahgal, V., Yue, G.H., 2001. Effects of aging on hand function. Journal of the American Geriatrics Society 49, 1478–1484.

Russell, C., Schofield, T., 1999. Social isolation in old age: a qualitative exploration of service providers' perceptions. Ageing and Society 19, 69–91.

Salthouse, T.A., Babcock, R.L., 1991. Decomposing adult age differences in working memory. Developmental Psychology 27, 763–776.

Salthouse, T.A., Maurer, J.J., 1996. Aging, job performance, and career development. In: Birren, J.E., Schaie, K.W. (Eds.), Handbook of the psychology of aging, Fourth ed. Academic Press, New York, pp. 353–364.

SeniorNet, 2002. Japan–US Comparison Survey. SeniorNet.

Sherer, M., 1997. Introducing computers to frail residents of homes for the aged. Educational Gerontology 23, 345–358.

Sloan, M., 2001. Web accessibility and the DDA. The Journal of Information, Law and Technology (JILT) p. 2. Available at: http://www2.warwick.ac.uk/fac/soc/law/elj/jilt/2001_2/ last accessed August 8, 2005.

Todman, J., Drysdale, E., 2004. Effects of qualitative differences in initial and subsequent computer experience on computer anxiety. Computers in Human Behavior 20, 581–590.

Valentine, G., Holloway, S., Bingham, N., 2002. The digital generation? children, ICT and the everyday nature of social exclusion. Antipode 34, 2.

Wilkins, A.J., 1995. Visual Stress. Oxford University Press, Oxford.

Worden, A., Walker, N., Bharat, K., Hudson, S., 1997. Making computers easier for older adults to use: area cursors and sticky icons. Proceedings of ACM Conference on Computer–Human Interaction, Atlanta, GA, USA, 18–23 April, 1997, pp. 266–271.

Zajicek, M., 2004. Successful and available: interface design exemplars for older users. Interacting with Computers 16, 411–430.

Zaphiris, P., Kurniawan, S.H., Ellis, R.D., 2003. Age related differences and the Depth vs. Breadth trade-off in hierarchical online information systems. Lecture Notes in Computer Science. (Proceedings of seventh ERCIM Workshop 'User Interfaces for All', 2003. Heidelberg, Germany) pp. 303–314.