A Trial Protocol for Evaluating Assistive Online Forms for Older Adults

Arthur Money  
*Brunel University*, arthur.money@brunel.ac.uk

Senaka Fernando  
*Brunel University*, senaka.fernando@brunel.ac.uk

Tony Elliman  
*Brunel University*, tony.elliman@brunel.ac.uk

Lorna Lines  
*Brunel University*, lorna.lines@brunel.ac.uk

Follow this and additional works at: [http://aisel.aisnet.org/ecis2010](http://aisel.aisnet.org/ecis2010)

Recommended Citation  
Money, Arthur; Fernando, Senaka; Elliman, Tony; and Lines, Lorna, "A Trial Protocol for Evaluating Assistive Online Forms for Older Adults" (2010). ECIS 2010 Proceedings. 90.  
[http://aisel.aisnet.org/ecis2010/90](http://aisel.aisnet.org/ecis2010/90)
A TRIAL PROTOCOL FOR EVALUATING ASSISTIVE ONLINE FORMS FOR OLDER ADULTS

<table>
<thead>
<tr>
<th>Journal:</th>
<th>18th European Conference on Information Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ECIS2010-0048.R1</td>
</tr>
<tr>
<td>Submission Type:</td>
<td>Research-in-Progress Paper</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Usability/usability evaluation, Human-computer interaction, Personalization, Web applications</td>
</tr>
</tbody>
</table>
A TRIAL PROTOCOL FOR EVALUATING ASSISTIVE ONLINE FORMS FOR OLDER ADULTS

Money, Arthur G., Brunel University, School of Information Systems, Computing and Mathematics, Uxbridge, UK, arthur.money@brunel.ac.uk
Fernando, Senaka, Brunel University, School of Information Systems, Computing and Mathematics, Uxbridge, UK, senaka.fernando@brunel.ac.uk
Elliman, Tony, Brunel University, School of Information Systems, Computing and Mathematics, Uxbridge, UK, tony.elliman@brunel.ac.uk
Lines, Lorna, Brunel University, School of Information Systems, Computing and Mathematics, Uxbridge, UK, lorna.lines@brunel.ac.uk

Abstract
The Delivering Inclusive Access to Disabled and Elderly Members of the community (DIADEM) project is funded through the Framework 6 European Union (EU) research programme. Its aim is to develop the DIADEM application which personalises the online form interface according to individual users’ needs, making the content more accessible for cognitively impaired older adults. In this paper, we present a trial protocol which has been specifically designed to identify cognitively impaired older adults, and to evaluate the usability of online form content from an older adult user’s perspective. To demonstrate the applicability of the trial protocol within the context of an ongoing research project, details of pan-European trials involving 77 eligible users, who evaluated DIADEM enabled online forms according to the trial protocol, are presented. The outcomes of the trials reveal a number of online form design recommendations, which will be incorporated into future versions of the DIADEM application. Although these have been primarily developed for informing the design of future versions of the DIADEM application, they also provide a valuable point of reference for online form developers more generally, which if followed, will ensure that content is more usable for the cognitively impaired older adult user group. To conclude, the lessons learned from implementing the trial protocol, and its appropriateness for this target user group is discussed, as well as how the findings may be incorporated into future versions of the DIADEM application.

Keywords: Assistive technology, Inclusive access, e-Government, Older adults, Online forms
1. INTRODUCTION
Declines in mobility, dexterity, and cognitive ability occur as a natural product of the ageing process (Craik and Salthouse 2007). Furthermore, the proportion of population that is made up of older adults is on the increase. Consequently, increasing numbers of older adults are living within their homes, but with restricted access to public services and in their ability to carry out independent living tasks such as banking, bill payment, accessing social services, and grocery shopping (Money, Lines et al. 2010). Government initiatives across Europe have identified web-based services as being a potential means of improving accessibility and quality of life for older adults (Money, Fernando et al. 2008). Until recently, accessing government services, such as welfare, housing, tax returns, and financial support has been achieved by completing and submitting paper-based forms. However, the wide spread availability of the Internet has meant that it is becoming increasingly common to expect users to access, complete and submit these forms on the web (Lines, Ikechi et al. 2006). The majority of online forms, however, are not designed for the specific needs of the older adult user, and thus commonly present significant usability challenges (Czaja and Lee 2008). Online forms differ from general web-content, due to their predominantly question and answer based structure. As a result, online forms pose usability challenges that differ from those posed when browsing more general web-content (Lines, Ikechi et al. 2006; Arch 2008). However, there seems to be very little research presenting online form specific design guidelines, and an even smaller amount specific to the older adult user (Sayago and Blat 2007).

A range of research has been carried out relating to online form design and usability. Wroblewoski (2008) presents a valuable contribution, suggesting that currently online forms are often developed from the ‘outside in’, reflecting the structure of an organisation’s database, expecting the user to provide information in a format and order that conforms directly to this. However, organisations should rather be designing online forms ‘inside out’, centred on the perspectives of the people outside the organisation and how they define being a member of the service offered to them. Das et al. (2008), explore the usability of different label positions within online forms, and recommend right-hand aligned labels. Interestingly, this challenges some of the findings presented by Wroblewoski (2008), highlighting the complex and context specific nature of online form design. Pauwels et al. (2009) explore whether coloured fields are more effective way of highlighting mandatory fields to the user, compared with the de-facto standard of using an asterisk. The results show that users more accurately identify and complete online forms that use colour to highlight mandatory fields. Bargas-Avila (2007) evaluate whether real-time error validation and feedback is preferable, as opposed to providing error feedback after the user has completed the whole form. Results revealed that the latter was more preferable and resulted in forms with lower numbers of errors being submitted.

Much research has been carried out to consolidate the knowledge contributions of studies in the area of online form design and usability by focusing on developing general web-content design guidelines. In some cases, this has considered the needs of the older adult to be fully incorporated by the needs of disabled users. For example, the World Wide Web Consortium Web Content Accessibility Guidelines (WCAG) (W3C 2008) provides 14 guidelines for the design of general web-content for disabled users. The United States general services administration section 508 (Administration 1998) also provides standards for supporting inclusive access for disabled users. The National Institute of Ageing (NIA), recognises older adults as a user group in their own right, and presents twenty-five guidelines for general web-content for users aged 60 (Morrell, Dailey et al. 2002). In terms of the older adult user group, only three small scale studies were found in research literature that present design guidelines specifically for online form content. Lines et al. (2006) and (2004), carried out two small scale studies, in the UK. The most recent of these (Lines, Ikechi et al. 2006), validated results from a previous study (Lines 2004) and as a result presented 13 online form design guidelines. Sayago and Blat (2007) also carried out a small scale study, involving one online form and seven Spanish older adult users. The findings revealed that users prefer checkboxes and radio-buttons as opposed to list-boxes. They also required optional fields to be grouped into separate sections.
These studies provide a valuable starting point for developing accessible online form content. However, the majority of e-Government web-content still fails to conform to the most basic design guidelines (Choudrie and Ghinea 2005). It seems that online form content will only be more accessible if service providers and designers choose to follow design guidelines that have been developed to focus specifically on the needs of the user. At present, this seems often not to be the case. It is therefore likely that older adults would benefit from a client-side application that adapts and personalises content so that they better conform to online form design guidelines. Such an application would provide older adult users with a standardised and familiar interface, which would enable them to access online form content more efficiently and effectively. The Delivering Inclusive Access to Disabled and Elderly Members of the community (DIADEM) project aims to develop a web-based application that adapts and personalises existing online forms based on the individual user’s preferences, and interactions with on-screen content. Section 2 provides a conceptual overview of the DIADEM application. A trial protocol designed to evaluate the application is presented in Section 3. In Section 4, the data analysis methods used, and the key findings of the DIADEM user trials are presented. Section 5 concludes the paper.

2. DIADEM APPLICATION OVERVIEW
From the user perspective, the DIADEM application is a plug-in to a web browser. It employs an Expert System (ES) element that monitors user interactions, and personalises the user interface so that usability challenges faced are alleviated. DIADEM operates alongside existing web-architectures. Thus, the DIADEM web-based services interface serves as an intermediary between existing application software, the DIADEM server, and ultimately the user or client site. The client component is data driven by XML control files from the server, so that it can interface to a range of services. From the client side, the DIADEM plug-in makes it possible to carry out the final enablement and transformation of existing online forms into the standardised DIADEM online form user interface.

The format and functionality of the initial DIADEM online form user interface is derived directly from the DIADEM online form requirements specification. Previous work related to developing the online form requirements specification is presented in (Money, Lines et al. 2010). The ES element of DIADEM monitors interactions whilst the user accesses online form content. Some Interaction Measures (IMs) may include: user inactivity; typing speed; key stroke accuracy; mouse click accuracy; frequency of backward navigation through the dialogue; frequency of submission errors; question completion rates. The DIADEM learning engine analyses these interactions and attempts to identify when the user appears to be in a problem state. Certain sequences of user interactions trigger ES rules, which in turn prompt the DIADEM application to provide the user with help (assistance). For example, a user may be considered to be in a problem state, if DIADEM detects that the user is typing in information via the keyboard, but does not have an input field selected. The current UK version of the DIADEM application has a number of ES rules, and which typically result in providing assistance to the user in the form of pro-active voice assistance (spoken instructions) that are aimed at helping the user overcome the detected problem. This functionality is still in early stages of development, and dependent on the outcome of this round of user trial results. Figure 1 provides an example of a DIADEM enabled Sheffield City Council online form.

3. TRIAL PROTOCOL
To evaluate the recently developed first version of the DIADEM application, and to identify future design guidelines, a trial protocol was developed and user trials were carried out according to it. Figure 2 provides an overview of the trial protocol. It should be noted that some of the methods employed within this trial protocol are tried and tested usability evaluation methods, such as the think-aloud protocol, thematic analysis and the analysis of log-file information relating to user interactions with online form content. However, the value of this trial protocol lies within the overall protocol design, the combination of these complimentary methods, and the lessons learned as a result of implementing this protocol with the older adult user group.
Phase 1: Participant recruitment and briefing
- Local councils, voluntary organisations, local hospitals
- Recruit 30-40 participants per partner country

Phase 2: Data collection
- Inclusion criteria
  - Age 70+
  - Years in full-time education <=12
  - Decline in memory or cognitive function?

- Online-form interaction
  - Re-iterate user tasks
  - Begin tape recording session
  - Present online form to participant
  - If necessary prompt participant to comment on: efficiency, effectiveness, general comments.
  - Make notes time whilst observing

- Post online-form interaction
  - Complete SUMI questionnaire
  - Discuss the following: Strengths, weaknesses, potential solutions to weaknesses, any other issues
  - Make notes during discussion on the note template provided

Phase 3: Data analysis and key findings
- Analyse inclusion criteria
  - Score ACE-R exam
  - Identify participants that met one of three inclusion criteria

- Think aloud and semi-structured interview analysis
  - Carry out thematic analysis
  - Familiarise with data
  - Generate initial codes
  - Search for themes
  - Review themes
  - Define and name themes

- SUMI scores
  - Collate SUMI scores and send for analysis

- Log file data analysis
  - Number of ES rules triggered vs ACE-R scores

**Figure 1:** An example of a DIADEM enabled Sheffield City Council change of address form

**Figure 2:** DIADEM trial protocol
3.1 Participant recruitment and briefing

Phase 1 of the trial protocol involves the recruitment and briefing of users. In the case of the DIADEM user trials, 94 users were recruited to take part in the trials, mainly sourced from voluntary organisations in Italy, Norway, and the UK, providing access to older adult volunteers who offer their time to take part in research. All users had some experience of using computers. Typically, volunteers take part in two or three research projects per year via these voluntary agencies. Occasionally, these projects have required users to complete paper-based forms and questionnaires as part of the research. From a methodological perspective, there may be a perceived risk in that these users could potentially be more adept at filling out forms, compared with members of the average population. However, since form-filling has become a common task that is carried out as part of our day-to-day activities, it was felt that the increased form-filling activity incurred from taking part in voluntary research was relatively infrequent, and thus did not warrant excluding these users from taking part in the study.

Clear communication with candidate older adult users is important, to avoid potential confusion, and to ensure that users are fully informed and aware of the requirements of the task. Initially, trial information sheets outlining the goals and requirements of the study were sent to the voluntary agencies, who distributed this information on our behalf. Users then contacted the voluntary agency if they wished to take part in the trials, and gave consent for their contact details to be passed on so that further arrangements could be made by telephone.

3.2 Data collection

Phase 2 of the trial protocol involves collection of data from users. Data collected during DIADEM user trials consisted of, inclusion criteria data, online form interaction data, post online form interaction data, and log-file data relating to the online form interaction task was collected. Details of each of these are now provided.

3.2.1 Inclusion criteria

Previous trials carried out on the DIADEM project (Money, Lines et al. 2008) used The Mini Mental State Examination (MMSE) (Folstein, Folstein et al. 1975) as the primary inclusion criteria. However, the MMSE was found not to be sufficiently sensitive to identify mild levels of cognitive decline. Consequently, for these trials, multiple inclusion criteria were used, so that users presenting with mild cognitive declines were more likely to be identified. For data to be included at the analysis stage, users were required to present with at least one of the following three criteria:

1) User aged 70+. All users aged 70+ were automatically included at the analysis stage. The rate of cognitive decline is known to accelerate with age (Salthouse 1982). For example, a meta-analysis shows that by the age of 40, an individual’s cognitive speed is expected to drop by 20%, this is expected to drop further, to 60% by the age of 80 (Salthouse 1982). Therefore, it is considered likely that older adults aged 70+ would present with some level of cognitive decline.

2) The user has spent 12 years or less in full-time education, and has self-reported a noted decline in memory or cognitive functioning. A high level of education is believed to have a protective effect on the level of cognitive decline an individual may present (Albert, Savage et al. 1995). Therefore those who reported to have spent 12 years or less in full time education, were more likely to present with cognitive decline. Furthermore, users were asked whether they have noticed or have been aware of any decline in memory or cognitive functioning. Self-reporting of changes in cognitive functioning or memory is an accepted means of identifying individuals who are presenting with some level of cognitive decline (Reisberg, Ferris et al. 1982).

3) The user scores lower than average on at least one sub-scale of Addenbrooke’s cognitive examination (ACE-R). Although the ACE-R is not as widely tested as the MMSE, since it is a relatively new screening tool, it fully incorporates the MMSE and it is now widely accepted that the ACE-R is significantly more sensitive than the MMSE alone (Larner 2006). Like the MMSE,
the ACE-R is still considered to be a brief, sensitive, and inexpensive screening tool (Mioshi, Dawson et al. 2006) that can typically be administered in less than 10 minutes. It is also shown to be more effective in picking early cognitive dysfunction/mild cognitive declines (Larner 2006) compared with the MMSE.

3.2.2 Online form interaction
The main trial activity took place in laboratory conditions, within designated office space provided by partner organisations involved with the DIADEM project. In the UK, trials took place within Sheffield City Council, In Italy, a usability laboratory was used, provided by CSI Piemonte, a technology service provider for the City of Turin. In Norway, trials took place within office space provided by a healthcare provider in Oslo. One researcher was present in the room at all times. Users were presented with a DIADEM enabled online form, which they were asked to complete whilst ‘thinking aloud’, and where possible provide comments on the effectiveness and efficiency of the form. Three online forms were used for these trials, one for each partner country. All forms may be considered to be e-Government related forms, i.e. forms used by public service organisations. In Italy, a form for Self-certification about the civil status of people was used. In Norway, a personal safety alarm application form was used. In the UK, a Sheffield City Council citizen’s change of address form was used. Any questions the user had at this stage were answered. Users were advised that minimal assistance from the researcher would be provided throughout the task, in order to achieve a realistic idea of typical challenges encountered when completing the online form. Each session was tape recorded, and notes were also taken, which would be used along with the tape recorded data at the analysis stage.

3.2.3 Post online form interaction
On completion of the DIADEM enabled online form, users completed a short satisfaction questionnaire. There are a number of tried and tested questionnaires, specifically designed to measure satisfaction of software usage. Some examples include, the Questionnaire for User Interaction Satisfaction (QUIS) (Schneiderman 1997), Website Analysis and MeasureMent Inventory (WAMMI) (Claridge and Kirakowski 2007), and Software Usability Measurement Inventory (SUMI) (Kirakowski and Corbett 1993). The WAMMI questionnaire was considered to be too simplistic, and is designed for evaluating the usability of web-content in general, as opposed to online form content. Although QUIS is a comprehensive questionnaire, it consists of 122 questions. After careful consideration, the QUIS was not considered as appropriate for the purposes of these trials and with this particular user group, due to the time and effort required to complete it, and the relatively complicated language used in the wording of the questions. It consists of 50 short and clearly worded statements, and can be administered relatively quickly (typically in less than 10 minutes). After reviewing a range of satisfaction questionnaires, the Software Usability Measurement Inventory (SUMI) questionnaire (Kirakowski and Corbett 1993) was chosen as most appropriate for these trials. One of the key benefits of SUMI, compared with other satisfaction questionnaires, is that it is made up of short statements that do not use irregular or low-frequency words, which as a result are likely to be easily comprehensible by the older adult target user group (Morrell, Dailey et al. 2002). Another benefit of SUMI is that the results can be compared with the benchmark SUMISCO database, which is a collection of over 3000 SUMI questionnaire results collected for a wide range of users and software applications. Finally, users were given the opportunity to elaborate on any of the think aloud comments made during the interaction task.

3.2.4 Log-file data collection
Log-files, recording user interaction behaviour and system responses to this behaviour, were created for each trial session carried out in the UK. The data recorded in the log-file served as a valuable source for gaining further insights into the older adult’s interaction experiences. One of the key features of the DIADEM application is to provide assistance to the user by monitoring user interaction measures, and triggering rules from the ES rule set, which result in triggering the system to provide the user with pro-active assistance. The log-file data contained a record of the number of times ES
rules were triggered throughout each session. Therefore, this made it possible to carry out an
evaluation of the frequency of ES rules triggered for each individual user, compared with their
respective levels of cognitive decline, as indicated by corresponding ACE-R scores. The outcome of
this evaluation would reveal whether the DIADEM application appeared to achieve its goal in
providing increased levels of assistance to users with comparatively high levels of cognitive decline.

4. ANALYSIS
After all trials had been completed, analysis and reporting activity was undertaken. This section details
this activity, and includes a brief summary of the key findings.

4.1 Inclusion criteria
The information collected for each user that took part in the trials was analysed during this phase.
Initially, based on the inclusion criteria, this involved identifying user data that was eligible to be
included at the analysis stage. Of the 94 users that took part in the trials, 77 were included at the
analysis stage of which 39 were female, 38 male, with an overall average age of 67.8. Overall, 31
users were aged 70+, 31 end-users self-reported to have experienced a decline in memory or cognitive
functioning and were in full-time education for 12 years or less. A total of 58 of the 77 end-users were
eligible for analysis by scoring below the ACE-R inclusion thresholds for one or more of the five
ACE-R subscales.

4.2 Think aloud and semi-structured interview
A thematic analysis was carried out on the ‘think aloud’, and semi-structured interview data collected
during the trials. Five key steps were taken to analyse the data. Initially all audio tapes were
transcribed into text format. The textual datasets (including observational notes taken by researchers
during trial sessions) were then initially perused to conceptualise the overarching themes that existed
within the textual datasets at a high-level, and noted in a coding template. The dataset was then
examined iteratively, enabling themes and sub-themes to be developed further. These were spliced and
linked together, and text relating to each category and sub-category were appropriately labelled. When
no further refinement of the categorisation could be derived,

4.2.1 Theme 1: On-screen information input
Theme description: Key issues that arose from users inputting information, and relating to the various
input mechanisms used in the online form (e.g., drop down boxes, radio buttons, free-text boxes).

Sub-theme 1.1: Drop-down boxes
Drop-down boxes provide a mechanism for inputting information without typing. These posed
frequent difficulties to users. In the majority of cases, users commented that they found it tricky to
select the drop-down arrow on the side of the menu and in many cases didn’t even realise the arrow
needed to be selected. The majority of which clearly stated that they found them awkward to use. For
example, some user comments included:

[Female, 72] ‘Why’s it do that? It’s not allowing me to type in the answer!’

Sub-theme title 1.2: Free text boxes
Multi-lined free-text boxes provide a mechanism for the user to provide information in free-
text format. In the change of address form, a multi-lined free-text box was used to allow the
user to provide their old address details. Users frequently encountered problems with multi-
lined free-text boxes, most commonly because they did not realise that the carriage return
button had to be pressed to get to the next line of the free-text box. One user summed up the difficulty as follows:

[Male, 66] 'Why can’t they just give you one long line to type in the answer?’

Sub-theme 1.3: Radio buttons
Users were positive about the radio-buttons within the forms. The standard radio-button sizes on the DIADEM enabled online forms are larger than those found in the original versions of the forms (50% larger). As a result, it seemed that users were more able to make accurate selections as a result of this. Occasionally, as in previous trials, users attempted to click on the area around the radio-button. Therefore, there may be a case for increasing the size of click sensitive area around the radio-button.

4.2.2 Theme 2: On-screen navigation
Theme description: Issues relating to how the user was expected to navigate their way around the form.

Sub-theme 2.1: Scrolling
Users experienced difficulties when required to scroll down the form. Users found it difficult to use the scroll bar function with the mouse, particularly to scroll to the segment of the form they needed. Consequently, they often attempted to use the arrow keys on the keyboard instead, which only allowed the user to scroll down the screen if they did not have a field selected. Whilst attempting to scroll, and often clicked on non-click sensitive parts of the form. Some users thought they had completed the form when they got to the last question displayed on the page, and then started to look for the ‘finish’ button. Users summed up their thoughts about the online form sections that required scrolling as follows:

[Male, 77] ‘It’s just not obvious when you have to scroll down’

Other issues with scrolling included occasionally, only part of a question was presented on-screen, whilst the other half of the question required scrolling down to become visible. Because some users were not familiar or able to use the scroll function effectively, they attempted to answer the question, without knowing what the full question was.

Sub-theme 2.2: Flashing cursor
The DIADEM application provides the user with an arrow indicator, either side of the selected input field, to help the user identify the location of the flashing cursor. Users found this to be a useful feature. However, a number of other users still managed to lose sense of the location of the flashing cursor. This may be, because users appeared to use different cues to locate the flashing cursor, some relying on the arrow indicators, and some ignoring the arrow, and looking specifically for the cursor itself. For the latter group of users, often the size of the flashing cursor was considered to be too small and caused problems.

Sub-theme 2.3: Mouse pointer size
Users also struggled to move between input fields when using the mouse. The key difficulty was that users often struggled to click on the exact on-screen location to enable them to enter into the next field. One user felt that it would help to increase the size of the mouse pointer, saying:

[Female, 70] ‘That (pointing at the mouse pointer icon) should be bigger and darker.’
Sub-theme 2.3: Moving between questions
The DIADEM application greys out questions once they have been correctly completed. Users found this feature extremely useful, and intuitive. A number of users made positive comments about this feature, for example:

[Female, 69] ‘The colour change helps you.’

Sub-theme 2.4: Moving between sections
Users often had difficulties in understanding how to move between sections of the online form. Although the DIADEM application standardised the ‘Next page’ and ‘Previous page’ buttons for the users to navigate the form, often the user was not aware of the location of these buttons within the form. The reason for this was that the buttons only appeared when the user scrolled right down to the end of the form, and thus taking into account the difficulties users had with scrolling, users were oblivious to the ‘Next page’ and ‘Previous page’ buttons. One other comment users made, was that the ‘Next page’ button should be located on the right hand side of the form, and only the ‘Previous page’ button should be located on the left hand side of the form.

The other mechanism provided for users to navigate between form sections was the tab system at the top of the form. Users were not familiar with tabs or how they worked. However, even those users that were familiar with tabs, very rarely used the tabs without eventually being prompted. The main cause for this seemed to be that when the user was ready to move to the next section of the form, they were near the bottom of the screen, and the tabs were no longer visible. Therefore the tabs did not always serve as a useful means of navigating to the next section of the form.

4.2.3 Theme 3: Assistance/help
Theme description: Comments regarding audio, textual, and visual prompts and assistance offered within the online form.

Sub-theme 3.1: Pro-active voice assistance
Based on user interaction behaviour, the ES rules implemented for the pilot DIADEM application triggered a system responses in the form of pro-active voice assistance (spoken instructions). In principle users felt that the voice assistance was a valuable and positive addition to the form filling exercise. They felt that it was more useful to have pro-active voice assistance compared with just textual assistance, and gave a personalised feel to the interaction experience. They also liked the idea of having a dialogue with the system. One problem with the pro-active voice assistance was that it kept repeating instructions if the user did not perform the action the DIADEM application was expecting. This repeated instruction was seen to be invasive and distracting.

Sub-theme 3.2: Read out loud question text
DIADEM also offered users audio assistance by reading out load each of the on-screen questions, when the question text was clicked on. Users were very positive about this feature, and felt it helped to hear the audio version of the question as well as to read the textual version of the question on-screen. This function helped to add clarity in terms of what was required of them. This version of the DIADEM application, however, did not provide reading out loud of the text displayed in the help sections, which users seemed to be expecting, based on the fact that question text did offer this functionality.

4.2.4 Design recommendations
A number of online form design recommendations were developed as a result of the thematic analysis. These will be integrated into version two of the DIADEM application. A summary of these are presented in Table 1.

Although these design guidelines have been developed specifically for the DIADEM application, they also serve as a valuable point of reference for online form designers. Adhering to
these guidelines is likely result in the development of more usable online form content for the cognitively impaired older adult user group.

Table 1: Design recommendations

<table>
<thead>
<tr>
<th>Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Theme 1</strong></td>
</tr>
<tr>
<td>R1:</td>
</tr>
<tr>
<td>R2:</td>
</tr>
<tr>
<td>R3:</td>
</tr>
<tr>
<td><strong>Theme 2</strong></td>
</tr>
<tr>
<td>R4:</td>
</tr>
<tr>
<td>R5:</td>
</tr>
<tr>
<td>R6:</td>
</tr>
<tr>
<td>R7:</td>
</tr>
<tr>
<td>R8:</td>
</tr>
<tr>
<td>R9:</td>
</tr>
<tr>
<td>R10:</td>
</tr>
<tr>
<td><strong>Theme 3</strong></td>
</tr>
<tr>
<td>R11:</td>
</tr>
<tr>
<td>R12:</td>
</tr>
</tbody>
</table>

4.3 SUMI scores

An analysis of the SUMI questionnaire responses included a comparison of the DIADEM SUMI scores against the SUMISCO database. To follow is a description of the key findings.

4.3.1 SUMI findings

SUMI reports results on five sub-scales (Efficiency, Affect, Helpfulness, Control, Learnability) and one overall scale (Global), which is the sum of the five sub-scales. Efficiency represents the extent to which users feel that the software helps them with their work. Affect relates to the user's emotional reaction to the application. Helpfulness is the extent to which the application is self-explanatory. Control measures how in control of the application the user feels. Learnability is a measure of how quickly the user feels they were able to master the use of the application. Table 2 presents the overall SUMI scores achieved by the DIADEM application for each of the five sub-scales and the overall Global scale.

SUMI scores have been transformed so that they may be compared against the SUMISCO database average, which is fixed at 50 for each of the scales. The results may be interpreted as follows: If a scale score does not encompass 50 within its Ucl/Lcl range, it may be considered 95% certain that the results are higher (if the range is above 50) or lower (if the range is below 50) than SUMISCO database average. If the score incorporates 50, the results may be considered to be similar to the SUMISCO database.
As can be seen, Efficiency, Helpfulness, Control, Learnability and Global all incorporated 50 in the Ucl/Lcl ranges. Therefore it may be inferred that DIADEM enabled forms achieved levels of user satisfaction similar to the benchmark SUMISCO database. However, the Ucl value achieved for Efficiency was 50, which indicates that DIADEM only just achieved similar levels of efficiency compared with the SUMISCO database. Therefore, it is likely that future versions of the DIADEM application may benefit from focusing effort on improving the ways in which users are offered help by the application. The Ucl/Lcl range for Affect was 59/52, which was above the SUMISCO database average, indicating DIADEM enabled online forms on average were more likeable than the benchmark SUMISCO database scores.

4.4 ES rules analysis

Four ES rules were developed for the UK version of the DIADEM application. These allowed the DIADEM application to monitor user behaviour, and provide assistance, typically in the form of pro-active voice assistance (spoken instructions) where it was deemed necessary. The main aim of trialling such ES rules, was to gain insights into users’ reactions to pro-active voice assistance, and explore whether there appeared to be any correlations between the measures of user behaviour and each respective user’s cognitive abilities. The initial set of rules were derived by developing a set of use cases, representing the key challenges that users may present with when interacting with online forms. A total of 25 use cases were then discussed with the development team, in order to design rules that may be implemented in order to support the user, should these use cases arise within the trial setting. In total, four ES rules that were trialled.

**ES rule 1, User lost focus (keyboard based):** This was triggered if key-presses were detected on the keyboard, whilst the DIADEM monitoring logic believed that the user did not have a valid input field selected and the user was not viewing a summary of the information they had provided. Consequently, a pro-active voice assistance message was triggered to instruct the user that first an appropriate input field needed to be selected, before any information could be typed into the online form interface.

**ES rule 2, User lost focus (mouse based):** This focused on user interactions made via the mouse. The goal was to identify when the user was experiencing difficulties with the form filling task, and was triggered if the user persistently clicked or double clicked the mouse button, but did not appear to be successfully select an input element on the form interface. Pro-active voice assistance was triggered to assist the user in the event of this case arising.

**ES rule 3, Form complete - user idle:** Identified when the online form appeared to be completed, however, the user appeared to be idle or inactive. As a result, a pro-active voice message was triggered to inform the user that the form appeared to be complete, but had not yet been submitted.

**ES rule 4, Form incomplete - user idle:** Identified when the form appeared to be complete (all input fields contained values) but some input fields appeared to contain errors and the user appeared to be idle or inactive. As a result, voice assistance was triggered to inform the user that the form appeared to be complete, but contained errors which required amendment prior to submitting the form.

---

### Table 2: SUMI scores compared with SUMISCO database

<table>
<thead>
<tr>
<th>Scale</th>
<th>UF</th>
<th>Ucl</th>
<th>Median</th>
<th>Lcl</th>
<th>LF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficiency</td>
<td>80</td>
<td>50</td>
<td>47</td>
<td>44</td>
<td>14</td>
</tr>
<tr>
<td>Affect</td>
<td>84</td>
<td>59</td>
<td>56</td>
<td>52</td>
<td>21</td>
</tr>
<tr>
<td>Helpfulness</td>
<td>82</td>
<td>55</td>
<td>52</td>
<td>49</td>
<td>15</td>
</tr>
<tr>
<td>Control</td>
<td>77</td>
<td>51</td>
<td>49</td>
<td>46</td>
<td>20</td>
</tr>
<tr>
<td>Learnability</td>
<td>91</td>
<td>55</td>
<td>51</td>
<td>47</td>
<td>4</td>
</tr>
<tr>
<td>Global</td>
<td>84</td>
<td>55</td>
<td>52</td>
<td>49</td>
<td>17</td>
</tr>
</tbody>
</table>
4.4.1 Log-file findings

An important outcome of these trials was to evaluate the ES rule triggering mechanisms employed in the DIADEM application. In particular, it was important to establish whether the DIADEM ES rules appeared to be sensitive to users who presented with comparatively high levels of cognitive decline. In order to explore this, the frequency of ES rules triggered in each session was compared with the ACE-R scores achieved by each respective user (which was used to represent level of cognitive decline). It was hypothesised, that if the ES rules were being triggered effectively, then the number of triggers within a session would increase as the users ACE-R score decreased (the lower the ACE-R score, the higher the level of cognitive decline).

In order to test this hypothesis, the log-files recorded for the UK participants were analysed and the number of ES rules triggered were counted for each user. A Pearson's r correlation test was carried out, comparing the number of ES rules triggered with ACE-R scores achieved by each of the 33 UK users. Since we hypothesised that frequency of ES rules would increase as the ACE-R scores decreased, a one-tailed test was used. There was a significant negative correlation between ACE-R scores and the number of ES rule triggers during a session (r = -0.298, n = 33, P < 0.05). Therefore the DIADEM application was successful in triggering significantly more ES rules for users that achieved comparatively low ACE-R scores (presenting with higher levels of cognitive decline). These results are promising, as they indicate that the DIADEM application appears to be sensitive to the user level of cognitive decline, and provides increased levels of assistance accordingly.

5. CONCLUDING DISCUSSION

This paper has presented a trial protocol and demonstrated how it has been used to identify cognitively impaired older adults, and evaluate the online forms produced by the DIADEM application. It should be noted that some of the methods employed within this trial protocol are tried and tested usability evaluation methods, such as the think-aloud protocol, thematic analysis and the analysis of log-file information relating to user interactions with online form content. However, the value of this trial protocol lies within the overall protocol design, and the combination of these complimentary methods, in order to provide a comprehensive protocol for the evaluation of the DIADEM application. The trial protocol supports the collection of a range of subjective and objective data, including ACE-R scores, think aloud data, semi-structured interview data, SUMI scores, and log-file data. Moreover, the lessons learned from implementing the various methods employed by this trial protocol, particularly with an older adult user group, has provided valuable methodological insights which provide a valuable point of reference for future research studies involving this target user group. Methodological lessons learned as a result of implementing the trial protocol include the following:

- To avoid confusion, and maximise participation rates, it is important to provide clear verbal and written instructions to older adult users, including written confirmation of agreed meeting dates, times, and locations.
- The three-fold inclusion criteria used in the trial protocol, provided a significantly higher ratio of user data being included in the analysis phase, compared with previous trials that used the MMSE as the primary means of including or excluding user data. Given that the ACE-R cognitive examination is recognised as being more sensitive than the MMSE, this allowed user data to be included at the analysis stage, with a greater degree of confidence.
- The majority of older adult users appeared to grasp the concept of the talk-aloud protocol and seemed to provide ‘think aloud’ data very readily.
- Carrying out a thematic analysis on the ‘think aloud’ data, and semi-structured interview data, proved to be an extremely useful approach to analysing such qualitative data.
- The SUMI questionnaire provided a useful measure of user satisfaction, and seemed to be well received by the older adult users. Older adult users seemed to understand the questions, and generally completed the questionnaire in less than 10 minutes.
• Collecting post-task data helped to triangulate and confirm the themes identified within the think aloud data and observational notes taken by the researcher during the trial, and provided a useful additional point of reference whilst carrying out the thematic analysis.

• Although from a technical perspective, there was some overhead in configuring the DIADEM application so that appropriate log-file data could be collected, this data proved to be valuable in validating the trigger mechanisms used for existing ES rule set.

• The thematic analysis carried out on user think aloud data and semi-structured interview data revealed a number of positive system design features for the overall design of the DIADEM interface, such as the pro-active voice prompts, enlarged radio-buttons, and read out loud questions feature.

As a result of carrying out the trial protocol, a number of design guidelines that will be applied to the future versions of the DIADEM application were also identified, some of these include: extending the read out loud questions feature to include help text, and reducing the repetition of pro-active voice prompts. Indeed these guidelines provide valuable guidance not only for the DIADEM application, but for online form content more generally. Online form content designed according to these guidelines, are likely to result in more usable online form content for the cognitively impaired older adult user group. Although a detailed comparison of DIADEM guidelines against a range of other web accessibility guidelines is beyond the scope of this paper, future DIADEM project activity will focus on further developing an expanded set of design guidelines. This will be achieved by developing and trialling updated versions of the DIADEM application. Valuable contributions relating to this activity, may be to develop a comprehensive set of design guidelines derived from the DIADEM user trials activity, which may then be compared to more general and well established web accessibility guidelines for older adults, such as those proposed by the National Institute on Ageing (Morrell, Dailey et al. 2002), Nielsen and Norman (2002), and Kurniawan and Zaphiris (2005). This will provide valuable insights into the extent to which DIADEM guidelines overlap and potentially add to current knowledge within this research domain.

The results of the SUMI satisfaction questionnaire revealed that four out of the five satisfaction sub-scales (Efficiency, Helpfulness, Control, Learnability) and the overall satisfaction scale (Global) achieved levels of user satisfaction similar to the benchmark SUMISCO database. However, the score achieved for Efficiency was the lowest of these four, which indicates that there more work could be done to improve the applications performance relating to this sub-scale. Since the Efficiency sub-scale relates to the extent to which users feel the application assists them in their work, it is likely that future DIADEM development effort could be best spent on improving the ways in which the user is assisted in completing online forms, and the ways in which this assistance is provided to the user. Conversely, the Affect sub-scale achieved above average the SUMISCO benchmark. These results are extremely encouraging, when considering that these scores were achieved from an older adult user group, that were likely to present with some level of cognitive decline. Indeed, a study by Nielsen (2002), found that web-content is twice as difficult to use for older adults aged 65 and over, compared with younger users. Hence, if DIADEM was not effectively catering for older adult’s needs, SUMI satisfaction scores may well have been expected to be considerably lower than the benchmark SUMISCO database averages.

An analysis of the UK log-file data revealed that the DIADEM application provides more frequent assistance to users that present with higher levels of cognitive decline. These are promising results, which provide valuable support to the overall direction of the project. In particular, the findings indicate that the measures used to monitor user interactions, and the mechanisms used to trigger ES rules appear to be particularly sensitive to users that present with high levels of cognitive decline.

Future research, in the short term, will involve making preparations for user trials that will be carried out to evaluate DIADEM version two against legacy online forms. This will allow us to verify the value of the DIADEM application compared with the original versions of the online forms adapted by DIADEM. In the medium term, one of our key research activities will be to evaluate the usability
of DIADEM enabled forms within a variety of real-world settings. This will allow us to develop guidelines specific to the environment in which the DIADEM application is used. For example, it may be the case that audio assistance is only desirable, when completing online forms in private or within a setting that does not have other audible influences such as radio, television, and crowd noise. As a step towards this, we have already commenced with the development of DIADEM enabled online form browsing kiosks, which will be positioned within foyer areas of County Council buildings in the UK. Indeed, although the trials presented in this study were carried out in laboratory conditions, a valuable future research activity will be to trial the DIADEM application within a variety of real-world settings. This may well help to trial the appropriateness of some of the problem state definitions that are employed by DIADEM currently, and evaluate the extent to which DIADEM continues to provide support to users that may be subjected to a number of real life interruptions whilst completing a DIADEM enabled form. Other valuable research directions may include exploring the potential of exploiting developments in the modal/continuation-based web-server research domain. This may make it possible to continually monitor the state of the user’s progress within the context of the web-based application functionality as a whole and open up new opportunities to expand on the existing DIADEM application functionality.

References


