ABSTRACT

Motivation – According to the UK’s National Skills for Life survey carried out in 2003, 16% or equivalent to 5.2 million of the UK population presented low levels of literacy (Williams, et al. 2003). In this study we investigate the differences in information seeking behaviours between low and high literacy users of an on-line social service system. Research approach – Ten volunteers participated in the study. Using the National Skills for Life Survey, five were classified as high literacy; five as low literacy. All participants were asked to think-aloud whilst carrying out the information search using the “Adviceguide” website. The four tasks were of varying difficulty; easy, medium and difficult. Observations, video recording, and a semi structured interview technique that uses cognitive probes were used. The qualitative data were transcribed and analysed using Strauss and Corbin’s (1998) Grounded Theory and Wong and Blandford (2002) Emergent Themes Analysis approach. Findings/Design – We identified eight themes or characteristics from this study; Verification, Reading, Recovery, Trajectories, Abandon, Focus, Satisfied, and Perception. Results showed that low and high literacy users demonstrated critically different characteristics. Take away message – To better support the low and high literacy users with information seeking, we plan to look at information seeking behaviour models as theoretical lenses to analyse their behaviour from the identified characteristics (Makri, Blandford & Cox, 2008). The behaviour models will better inform the development of interface design for low and high literacy users.

Keywords
Low & high literacy users, information search task, information seeking characteristics, information seeking behaviour models.

INTRODUCTION

The purpose of this study is to investigate information seeking behaviours of low and high literacy users of an online social service system called “Adviceguide”. This system is part of the services provided by the Barnet Citizen’s Advice Bureau which provide support to clients from socially disadvantaged backgrounds. According to the UK’s National Skills for Life survey carried out in 2003, 16% or equivalent to 5.2 million of the UK population presented low levels of literacy (Williams, et al. 2003).

A previous study by Kodagoda & Wong (2008) looked at the performance of low and high literacy users information search and retrieval. The results show that low literacy users perform significantly worse than high literacy users as they used a website to search for social service information intended for use by socially disadvantaged citizens. Key results indicate that (i) low literacy users take eight times more time than high literacy users to complete an information search task, and yet were significantly less accurate, (ii) low literacy users on average spent one-third more time on a web page than high literacy users, but did not seem to be informed by it, (iii) low literacy users employed a much less focused information search strategy than high literacy users visiting eight times more web pages in total, (iv) low literacy users back-tracked 13 times more frequently than high literacy users, and are four times more likely to re-visit web pages, and (v) low literacy users are 13 times more likely to be lost than high literacy users.

Another study carried out on low-literacy users reading strategies and navigational behaviours characteristics were; that they tend to read word by word, having a narrow field of focus, skip chunks of text if confronted by long and dense pages, quickly being satisfied with information found, minimise the amount of reading by skipping from one link to the another, and avoided searching as it requires spelling and typing (Summers & Summers, 2005).

© The Authors 2009. Published by the British Computer Society
Proceedings of NDM9, the 9th International Conference on Naturalistic Decision Making
London, UK, June 2009
347
As more government and social services information are placed online, the problems faced by the low literacy users should be considered. A comparative study of information seeking behaviour between five high literacy and five low literacy participants using the “Advice guide website is discussed in this paper.

**STUDY**

**Participants**

Ten clients of the Citizens Advice Bureau volunteered for the study. They comprised six females and four males with an average age of 45 years. Using UK’s National Skills for Life literacy assessment survey, five participants were classified as high literate, and the remaining five were classified as low literate. Low literacy participants showed a reading level below UK’s National Curriculum Level 4 while high literate users showed reading levels above 5. None of the participants had used the “Adviceguide” website previously, although all the participants had some experience in using the Internet.

**Design**

Four information search tasks were developed based on the highest type of advice provided to the clients of the Barnet Citizens’ Advice Bureau during April 2005 to May 2007. Information search task required the participant to find a specific piece of information, such as eligibility to receive benefits, money advice, assistance on giving up smoking, details of local child care availability, information on children hospital, and advice on council tax arrears. The information search tasks vary on task complexity: two of tasks were classified as easy (E1 and E2), one medium (M) and one difficult (D). For the easy Tasks (E1, E2), the user required a maximum of two navigational steps and minimum amount of reading to find a solution from the home page. For the medium task (M), the user required three to four navigation steps with multiple paths leading to the same information. For the difficult Task (D), the user required more than five navigation steps, same solution in multiple places.

**Methodology**

Multiple Cognitive Task Analysis (CTA) methods were used to extract and understand the human decision process during their cognitive work (Wong, 2006). CTA methods used in our study were process tracing, observation and interview methods which are broad families identified by Cook (1994). Process tracing captures participants thoughts during an information search task via the think-aloud protocol. Video captured the participants’ interactions with the “Adviceguide” website. Users were prevented from using the sites or other search facilities, users were asked to follow the web site menu links. User observations made during the information search tasks, assisted during interview sessions to prompt the user for further clarifications, reasoning for acting in a particular manner. All participants were asked to think-aloud whilst carrying out the information search task. The Critical Decision Method (CDM), a semi-structured interview technique that uses cognitive probes was used to extract information and experience of the interviewee. The interviews aimed to (i) find out an interesting or memorable (good or bad) experience they encountered during the information search task, (ii) explain an incident the facilitator found interesting which was identified during the observations. The facilitator played back the videos to the participants to make it easier for them to recall the incident and explain their decisions, actions or thoughts. The videos were played back for two reasons: to help the participant easily recall the incident and to avoid any confusion.

The qualitative data were transcribed using HyperRESEARCH, and analysed using Strauss and Corbin’s (1998) Grounded Theory (GT) and Wong and Blandford (2002) Emergent Themes Analysis (ETA) approach. Both process involved coding the data at different levels, where one used a bottom up approach while the other using a top down approach.

**Procedure**

The study was conducted at the Middlesex University, Interaction Design Centre Usability Lab. First, participants were given instructions and were informed that they could ask for assistance to read the information search tasks if required. The participants used the “Adviceguide” website home page as the starting point. Participants were asked to navigate through the website using the standard links and menus and were restrained from using any external or internal search facility. The participants had one practice trial to familiarise them with the thinking aloud protocol. All participants granted us permission for video recording. When the participants were ready to begin, one of the four information tasks were handed in a randomized order to minimize possible learning effects. The participants controlled the start and the end of the information search task. The participants were asked to notify the researcher when the information search task was completed or abandoned. The participants were requested to convey or write the specific target piece of information found. Each information search task commenced from the home page of the “Adviceguide” website. The information search task ended when the participant found the relevant information or abandoned the search. Participants were not given any clues or directions to solve the information search task.

The participants were also asked to think aloud during the information search task. At the end of each information search task, the participant handed in their answer. At the end of each session each participant was interviewed about
their experiences in carrying out the tasks. The interviews were based on a CDM like approach. Each participant was shown parts of the video to help them reflect and explain their behaviour at the time.

**THE USE OF COMBINED COGNITIVE TASK ANALYSIS METHODS**

In our study we used process tracing, observation and interview CTA methods. CTA is defined as “the extension of traditional task analysis techniques to yield information about the knowledge, thought processes, and goal structures that underline observable task performance” (Schraagen, Chipman & Shalin 2000).

The combination of the CTA methods enabled us to capture interesting insights of the two different user groups. Insights to some of the interesting findings were (i) low literacy users were unable to articulate themselves fully during the think-aloud protocol, (ii) while some high literacy users omitted some of the interesting reasoning. In both cases the observations made during the information search task, assisted us to probe the participant during the semi-structured interview sessions and were also able to show the clips of the recoded video to help them recall and explain their thinking during that time frame.

When observing the low literacy users during their information search task, we found that they were unable to articulate why they had abandoned a search. During the CDM interview sessions when the video clips were played and probed them to recall and explain what they were doing during that time, their explanations were they abandon the search due to insufficient information 30% of the time, or they believed the information they are looking for should be placed in a certain place on the website, and since they could not find any similarities of what they expected on the “Advice Guide” website they abandon the search 40% of the time. This group of low literacy users had a clear expectation of where the information should have been placed. But none of the low literacy users expectations were similar to another.

We have taken one low literacy participants quote to explain findings from the CDM interview. Participant - LL1 information search task - E1 "there should be one in here holiday rules and regulations, so you can click on holiday and find straight away how much you’re entitled for whatever".

When observing the high literacy users during their information search task, we observed that some of the high literacy users opted to make no comment during the think-aloud sessions, even when we saw they found information they required, but still continued to search further until they found another complementary set of information. When we replayed the video clips and asked them what they were doing during that time, their answer was they were trying to verify the information found previously. We have taken one high literacy participants quote to explain the findings from the CDM interview. Participant - HL1 information search task - E1 "then I went to paid holidays just to double check and confirm and I found the same thing".

By triangulating and orchestrating our methods, we were able to use more than one CTA method in order to overcome the limitations of the other methods. The information that was not found during the think-aloud, was identified during the observations and we were able to probe the participant during the CDM semi-structured interview session. The playback of the video clip further helped the participant recall what they were doing at that specific time.

**RESULTS**

Below we discuss low literacy users identified eight characteristics. The identified characteristics: Verification, Reading, Recovery, Trajectories, Abandon, Focus, Satisfied and Perception. They are described below.

**Verification:** Low literacy users do not verify the information found for correctness.

Verification takes place when users find information they need and examine other related links for correctness.

The verification made by the high literacy users on their information search tasks are shown in Table 1. In the medium and difficulty tasks, 100% of the high literacy users verified the information found for correctness. However, for the easy tasks, verification decreased to 80%.

None of the five low literacy users attempted to verify the information as shown in Table 2. As soon as the low literacy users found some relevant or interesting information they stopped their search and assumed the information they found was correct and abandon the search early, or abandon the search if they did not find any relevant or interesting information.
Identifying Information Seeking Behaviours of Low and High Literacy Users

We have extracted a few observations and quotes from our high literacy participants to justify our above conclusion.

Participant HL1, Information task E1. This high literacy user found the answer to the task, but still checked another link to verify the answer.

"then I went to paid holidays just to double check and confirm and I found the same thing"

Participant HL2, Information task E1. This high literacy user found the answer to the task, but still scrolls down to "how much paid holiday can you take" to verify the answer found.

"Hmm I am just trying to see if there is any other information regards to the paid holidays"

Participant HL5, Information task E1. The high literacy user found the answer but still wanted to verify the answer with other available links.

"I have already identified the information that is useful to me, which is holiday and holiday pay, there is a list now, with various information that I would like to find more on"

Reading: Low literacy users read word by word.
Reading takes place when users trying to read word by word, trying to make sense of the information they read. Scanning take place when a user takes a glance through headings and subheadings or start, middle of a paragraph until they find something relevant or interesting.

The scanning and reading made by the high literacy users during their information search tasks are shown in Table 3. In the easy and medium difficult task, 100% of the high literacy users scanned and read. However, for the difficult task, scanning decreased to 90% while the reading remained at 100%.

All low literacy users read word by word trying to make sense of what they were reading as shown in table 4. We observed that all 5 low literacy users pointed the mouse at the words being read or at the line they were reading. They showed a 100% reading on all task difficulties.

Recovery: Low literacy users were unable to recover from a mistake.
Recovery refers to recuperate from a wrong or irrelevant information search to a more focused or relevant one were the information search resulting in finding the information.
High literacy users were observed to be more likely to recognize wrong or irrelevant information and recover by finding the relevant information as shown in Table 5. We found 35% high literacy users identified wrong or irrelevant information at the early stages in their search and was able to recover. When high literacy users identified a wrong or irrelevant link they used appropriate keywords enabling them to back track or choose another link to make adjustments to their current search. However during the easy levels, recovery was not required.

The low literacy users were less likely to recognize wrong or irrelevant information and recover as shown in Table 6. We found 20% of low literacy users identified wrong or irrelevant information but were still unable to recover. After identifying irrelevant or wrong information 50% of the time they arrived at wrong result assuming the information they found was correct. The remaining 50% of the time they abandon their search assuming the information they have is insufficient to solve the problem.

Some remarks that highlight the effect of recovery of low and high literacy users are:

Participant HL4, Information task M. This high literacy user at the start followed links that were wrong and irrelevant but was aware, gained focus and found the necessary information. Appropriate keywords enabling to back track and choose correct links.

"I am going to go to “Frequently asked questions ....” A list have come up with questions and err so far this link has not been very useful to me. Now I am going to try the link on “Your Money” and “Employment” there is a list of the employment stuff there"

"I think this will give me the answer Does my employer need to give me a certain period of notice before he dismisses me, hmm here one week, if you’ve worked for your employer for one month but less than two years."

Participant HL5, Information task M. The participant identified irrelevant content but was able to recover.

"hmm I am just going to go back and see if there is something more relevant"

"I think this will give me the answer “Does my employer need to give me a certain period of notice before he dismisses me, hmm here one week, if you’ve worked for your employer for one month but less than two years."

Participant LL3, Information task E2. The participant identified irrelevant content, was unable to recover and gave up.

"does not have any information on the holiday pay, I went to the employment and from there to government employment schemes and there is no holiday pay or any information on pay"

Participant LL4, Information task M. The participant identified irrelevant content but was unable to recover assuming the information found was correct.

“Ok that’s what I find it very it, dismissed, unfair dismissed or actual dismissed and again reads through the links, ha ha ok what is wrongful dismissal ok that’s the one I found it”

Trajectories: Low literacy user trajectories were dissimilar.

The trajectories are information search paths taken.

The High literacy users had similar clues which lead to very similar trajectories during their search paths as shown in Table 7. We found 75% of the high literacy users had similar trajectories. However in the easy levels the trajectories were 100% similar and as the tasks got difficult similarity trajectories decreased up to 40%. The low literacy users
showed dissimilar trajectories as shown in Table 8. We found, only 10% of the low literacy users had similar trajectories.

<table>
<thead>
<tr>
<th>High literacy users</th>
<th>Information search tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E1</td>
</tr>
<tr>
<td>HL1</td>
<td>T</td>
</tr>
<tr>
<td>HL2</td>
<td>T</td>
</tr>
<tr>
<td>HL3</td>
<td>T</td>
</tr>
<tr>
<td>HL4</td>
<td>T</td>
</tr>
<tr>
<td>HL5</td>
<td>T</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Low literacy users</th>
<th>Information search tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E1</td>
</tr>
<tr>
<td>LL1</td>
<td>T</td>
</tr>
<tr>
<td>LL2</td>
<td>D</td>
</tr>
<tr>
<td>LL3</td>
<td>T</td>
</tr>
<tr>
<td>LL4</td>
<td>D</td>
</tr>
<tr>
<td>LL5</td>
<td>D</td>
</tr>
</tbody>
</table>

Table 7. Comparison of similar and different trajectories carried out by the high literacy users. (T= similar trajectories, D=different trajectories)

Table 8. Comparison of similar and different trajectories carried out by the low literacy users. (T= similar trajectories, D=different trajectories)

One of the most interesting tasks that show the different behaviours between low and high literacy users is E2. While high literate participants presented the same trajectories in all the trials, low literate presented completely different trajectories.

The trajectory used in E2 for high literacy users was:

*Employment -> Basic rights at work -> holiday and holiday pay (users found required information)*

*More information link...* (Users went to this link to further verify the information found)

As shown below a) presents low literacy user LL1, b) presents low literacy user LL2 and c) presents low literacy user LL4 trajectories, respectively.

a) *Employment -> Basic rights at work -> holiday and holiday pay (users found required information)*

b) *Employment -> Government employment schemes -> Other help (user gave up the information search by this point)*

c) *Employment -> Dismissal -> Steps to work through to identify an unfair dismissal -> scrolls down Step two: have you actually been dismissed -> scroll up Dismissal -> scrolls down Step two: have you actually been dismissed -> scrolls down to What is wrongful dismissal -> scrolls Dismissed (user assumed the information extracted was correct)*

**Abandon, Focus, Satisfied, and Perception.**

*Abandon:* low literacy users showed a higher tendency to give up their search, unable to find information (Refer in Table 5 & 6. A stands for Abandon).

*Focus:* that they have a narrow focus of attention, and are not likely to notice content above, below or to the side of their focus.

*Satisfied:* they were likely to be satisfied and abandon the search as soon as they assumed the information found were relevant information.

*Perception:* low literacy users perception on where information should be available on the website was a mismatch to their perception, furthermore there were no similarities between them.

The behaviours described below lead us to believe the above described characteristics.

a) They assumed most of the time the information they are looking for is unavailable and abandon the searches at early stages.

b) They never verified the information they have found to confirm correctness and abandon the search early, this is due to quick satisfaction with information.

c) They got easily confused if the long dense pages which had anchor links.

d) The user had a perception where information should be on the website; if the information was not available in the places expected they immediately abandon the search assuming the information was not available.

Participant LL1, Information task E1. The user did not want to back track and find another solution expecting the answer to be found within the selected links.

"This is why I go for now for the employment, and it should be there whatever the government law, how long your entitle for holiday the information not coming up"
Participant LL2, Information task E2. The user was scrolling up and down in one page, clicking on anchor links and did not check other options available and abandon the search.

“hmmm same comes up ‘Help finding work’

scrolls up and down again and says

“same comes up again… there is nothing on holiday in the employment, I cannot find the information”

High and low literacy user Successful information searches and user assumed success.
a) Successful completion: High literacy users had a 100% higher successful completion for the easy level and the percentage decreased to 60% for difficult tasks. Low literacy users succeed only 40% of the time on the easy tasks and were not successful with the difficult ones. Please see Table 9.
b) User Assumed success: High literacy users did not assumed success for the easy level but the percentage increased to 40% for the medium and difficult levels. The low literacy users assumed success remained at 40% throughout all levels. Please see Table 9.

<table>
<thead>
<tr>
<th></th>
<th>High Literacy</th>
<th>Low Literacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Successful completion</td>
<td>100% M 80% D</td>
<td>40% E 20% M 0% D</td>
</tr>
<tr>
<td>b) User Assumed success</td>
<td>0% M 20% D</td>
<td>40% E 40% M 40% D</td>
</tr>
</tbody>
</table>

Table 9. High and low literacy user successful information search and user assumed success.

Some of the interesting findings summarised and discussed below. Please see Table 10 for some of the interesting summarised information.

a) Verification: High literacy users had 100% verification for medium and difficult levels. For the easy levels, verification was only 80%. The low literacy users did not verify their answers.
b) Reading: High and low literacy users both carried out reading at some point in their search. While scanning, High literacy users scanned for information 100% on the easy and medium difficult levels, however scanning decreased to 80% for the difficult level. The low literacy users did not scan information at all.
c) Recovery: High literacy users identified wrong or irrelevant information in medium and difficult levels. Their recovery decreased from 100% in the medium level to 67% in the difficult one. The low literacy users identified wrong or irrelevant information in the easy and medium levels. However, they were not able to recover.
d) Trajectories: High literacy users presented 100% similar trajectories at the easy level. However this decreased to 40% as the tasks got difficult. The low literacy users showed similar trajectories only 40% of the time if the task was easy.
e) Abandon: High literacy users did not abandon their search. Low literacy users abandonment rate increased from 20% for the easy level to 60% for the difficult level.

<table>
<thead>
<tr>
<th></th>
<th>High Literacy</th>
<th>Low Literacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Verification</td>
<td>80% E 100% M 100% D</td>
<td>0% E 0% M 0% D</td>
</tr>
<tr>
<td>b) Reading</td>
<td>100% E 100% M 100% D</td>
<td>100% E 100% M 100% D</td>
</tr>
<tr>
<td>Scanning</td>
<td>100% E 100% M 80% D</td>
<td>0% E 0% M 0% D</td>
</tr>
<tr>
<td>c) Recovery</td>
<td>- E 100% M 67% D</td>
<td>0% E 0% M -</td>
</tr>
<tr>
<td>d) Trajectories</td>
<td>100% E 60% M 40% D</td>
<td>40% E 0% M 0% D</td>
</tr>
<tr>
<td>e) Abandon</td>
<td>0% E 0% M 0% D</td>
<td>20% E 40% M 60% D</td>
</tr>
</tbody>
</table>

Table 10. Comparison of the summary results for high and low literacy users.

DISCUSSION
Our study was consistent with the findings of Summers and Summers (2005). Both studies showed that low literacy users read word by word trying to make sense of information and do not present the ability to scan (reading). They have a narrow field of view and are not likely to notice content above, below or to the side of their focus (focus). They were likely to be satisfied and abandon the search early assuming they found relevant information (satisfied). However,
we did not observed users tendency to skip chunks of text when faced with dense pages as described by Summers and Summers (2005).

Our study found when users were presented with dense pages with anchor links, they were very likely to get lost and disoriented, this resulted low literacy users to abandon the search. We also identified the following characteristics: Low literacy users do not verify the information found for correctness (verification). They were unable to recover from a mistake even if they did identify wrong or irrelevant content (recovery). They did not share similar clues that lead to very different (trajectories) during their search paths. Low literacy users mental representations of the categories were a mismatch to the system (Representation). Finally, low literacy users had several reasons to abandon an information search task, (a) unable to find the information, (b) unable to recover from a mistake, (c) mental representation of the categories being a mismatch to the systems representation, (d) being satisfied quickly.

CONCLUSION
In conclusion, low literacy users demonstrated a critically different strategy from high literacy participants when searching information using the “Adviceguide” website. They spend a lot of time reading instead of scanning, usually terminating the search before finding the right information. Verification was inexistent and a recurrent attitude to give up and terminate the search was presented. Their ability to recover from encountering wrong information was very low and they demonstrated a very narrow focus in all the cases. These behavioural patterns provoked low literacy users to use different search paths or trajectories. Future research aims to generalize these findings and provide a framework for designing interactive websites.

We suggest interfaces design should reduce low literacy users memory load, while maintain their attention by presenting less textual information; use high level linked clusters to afford rapid scanning so that they can see overall relationship structure; text which is simple to read; use of visual and audio where appropriate; assist users to recover from a search; try to match users mental perception with the interface design.

REFERENCES