

Uncertainty-tolerant design: Evaluating task performance and drag-and-link information gathering for a news writing task

Simon Attfield (s.attfield@cs.ucl.ac.uk*), Ann Blandford (a.blandford@cs.ucl.ac.uk), John Dowell (j.dowell@cs.ucl.ac.uk), Paul Cairns (p.cairns@ucl.ac.uk)

UCL Interaction Centre, University College London, Remax House 31/32, Alfred Place, London, WC1E 7DP, UK.

*Corresponding author (Telephone: +44 (0) 207 679 5242, Fax: +44 (0) 207 679 5295)

Keywords: Information Behaviour, Writing, PIM, Journalism, System Evaluation.

Abstract

Part of the challenge of designing systems to support knowledge work is to do so in a way which is sympathetic to users' uncertainty. NewsHarvester is a test-bed system designed to support news research and writing in a way that accommodates uncertainty in relation to information gathering. It does this using 'drag-and-link'; a simple feature by which text extracts copied from source locations are appended with hyperlinks to force the re-display of the source. We describe the rationale for using drag-and-link within NewsHarvester based on a previous ethnographic study of journalists, describe its implementation within NewsHarvester, and report a user-evaluation which compared drag-and-link with printing and standard drag-and-drop as information gathering mechanisms.

We found that users wanted to relocate information they had not previously identified as useful in order to include it in their report, to better understand the context of information already extracted, and as part of a more serendipitous search for information to add to a near-complete report. Users also considered drag-and-link an easier method for gathering information than printing, and considered that drag-and-link made it easier to relocate information. They also considered that drag-and-link promoted more flexible and dynamic working and increasing user enjoyment. An assessment of the quality of their work showed a trend that favoured drag-and-link over the other two methods, although this was not statistically significant. We conclude that drag-and-link improves user-experience during research and writing tasks in the face of information gathering uncertainty.

1 Introduction

The goal of knowledge work is to acquire and assimilate information about some aspect of the world and to respond by initiating change back into the world. The response might be to recommend or implement a course of action, such as legal strategy or a healthcare treatment, or it may be to generate a new information artefact, such as a marketing report, or a student essay. The information that is acquired and assimilated may include relevant law, the facts of a case, presenting symptoms, medical knowledge, or academic literature as it relates to a topic or question.

Within this framework of assimilation and response, though, knowledge work is beset with uncertainty. Under the compass of 'response' lies an indefinite range of possibilities which lack clear definition. The knowledge worker must shape and judge their response according to emerging constraints and their developing understanding of the presenting situation constructed from information. This understanding can be fractured, partial and transient as new information is encountered, ideas for a solution are formulated and revised, and new information becomes pertinent for discovery. Knowledge tasks typically feature indeterminability concerning task outcomes, process and information requirements [Byström & Järvelin, 1995]. As such, they have the uncertainty characteristics of wicked problems [Rittel & Webber, 1973].

Part of the challenge of designing systems to support this kind of work is to provide information interaction functionality in a way which is sympathetic to the intrinsic uncertainty of the task. There is a need for systems that are uncertainty-tolerant. How this is achieved depends upon the actions supported and how uncertainty impacts on them. Consequently, there is a need to understand how information intensive work gives rise to particular kinds of information interaction and how uncertainty impacts on these.

In this paper, we report a study which focuses in particular on uncertainty as it applies to information gathering, and the ad-hoc information collections that people create as part of research and writing tasks. As users proceed through cycles of information acquisition, assimilation and use, they create and maintain personal and informal collections of information. O'Hara et al [2002] and Sellen, Murphy and Shaw [2002] use the analogy of a 'holding pattern' to describe such a collection. The metaphor seems apt, bringing to mind a transient collection managed by a central agency. An information holding pattern is an ad-hoc, intermediate and temporary formation created for some use, after which its value has generally expired.

The study was performed with a test-bed system called NewsHarvester. NewsHarvester integrates news archive search, document display, text gathering, note-taking and copy-writing within one application, with equal prominence given to each. This integration was motivated by the idea that knowledge work (ideally) forms a flow of iterative and interdependent activities, and that by reducing separation opportunities should arise for interrelating actions across tools and for better supporting natural continuity. This is intentionally contrary to an implicit design assumption of independent activity behind most of today's knowledge work applications (e.g. digital libraries, browsers, word-processor etc.).

The study had two aims, both motivated by findings from a previous qualitative field study of journalists who were researching and writing news reports for a national newspaper [Attfield & Dowell, 2003]. The field study showed that, given the dynamic,

constructive nature of the journalists' task and its changing context, relevance judgements that underpin information gathering actions from archive documents were uncertain and subject to change. Consequently, a source document may be consulted several times as a task unfolds. The first aim of the current study was to evaluate a set of claims about information gathering behaviour within this kind of task, including the need to refer back to documents from which information had already been gathered, and some explanations for why this might be so.

The second aim of the study was to evaluate 'drag-and-link' as one way of supporting referring back in this way. Drag-and-link is a feature whereby content copied from a source is automatically appended with a hyperlink back to the originating document [Karger & Jones, 2006]. The team working on the Keeping Found Things Found project at the University of Washington implemented drag-and-link in their Universal Labeler system [Jones et al, 2005; Karger & Jones, 2006]. The Universal Labeler is designed to address the fragmentation of users' information across multiple locations and organisational schemas (e.g. in the head, on paper, in bookmarks and emails etc.). A similar motivation underlies the implementation of drag-and-link within Microsoft OneNote¹.

In the current study we link the integration of information from multiple sources supported by drag-and-link with uncertainty tolerance. Where information is gathered into a personal collection using drag-and-link the collection acts as an index to further source content with easily traversable links. This allows the originating context to be reviewed conveniently and without unduly disrupting task flow. One way of understanding the role of such links is that they provide convenient 'channel factors' [Ross & Nisbett, 1991, p10] supporting information re-access. Channel factors are aspects of a situation which can appear as minor facilitators (or barriers) for an action, but which in fact have a strong influence.

Drag-and-link is one approach to a problem for which there may be other solutions. For example, an interface might maintain a representation of documents read during an assignment from which any can be selected and reviewed. Users might additionally annotate these with contextual information about relevance to an assignment. To some extent the choice of drag-and-link delimits the scope of our study, which is to evaluate one way of supporting the uncertain knowledge worker; further studies might consider or even compare alternative methods. However, drag-and-link has the benefit of offering an economy of representation by augmenting an existing interface object (i.e. a collection of notes), and economy of interaction by exploiting existing gathering behaviours (i.e. drag-and-drop).

The structure of the remainder of the paper is as follows: - In the next section we review literature relating to some key concerns of the paper. These are: information seeking uncertainty, information gathering, and the use of source documents during writing tasks. Then in section 3, we give an overview of findings from the newsroom field study that provide the motivation for using 'drag-and-link' to support news research and writing. In section 4 we describe NewsHarvester and how it implements drag-and-link. Then, in section 5 we detail the objectives of the study in terms of the claims it tested, and in section 6 we describe the method of the evaluation. We report the evaluation results in section 7; in section 8 we summarise and discuss our findings; and we finish with some concluding comments in section 9.

¹ <http://office.microsoft.com/en-gb/FX010858031033.aspx>

2 Background literature

2.1 Information seeking uncertainty

The idea that information seekers are frequently uncertain about the information they want has been a persistent theme within user-centered information seeking research in Information Science. In a seminal paper about negotiations between information seekers and library intermediaries, Taylor [1968] defined different levels of information need within a process of moving from an actual but perhaps unrecognised need ('visceral') to an expression of a need which could be presented to an information system ('compromised'). Belkin, Oddy and Brooks' ASK hypothesis [1982a, 1982b] echoed this idea by stating that "an information need arises from a recognized anomaly in the user's state of knowledge concerning some topic or situation and that, in general, the user is unable to specify precisely what is needed to resolve that anomaly" [Belkin, Oddy & Brookes, 1982a, p.62]. They argued that supporting uncertainty was a fundamental problem for Information Retrieval (IR) and a point where system design should begin. A related idea arose within Bates' influential Berrypicking model [1989]. Bates pointed out that as people search, so they gather information and that their needs evolve as information triggers shifts in thinking and new lines of enquiry.

Kuhlthau performed a series of studies of information-seeking (primarily of school and university students) [Kuhlthau, 1988a; 1988b; 1988c; 1989] which gave rise to her ISP model [Kuhlthau, 1993]. The ISP model describes cognition, behaviour and affect during searching in the context of wider tasks. Central to her model, is the idea that information needs begin as vague and unclear, and interaction with information services is consequently difficult. Later, as the user develops a greater focus, so searching becomes more focused and precise.

Following Kuhlthau's work, there has been growing acceptance that formulation, in respect of a wider task, has a significant effect on information behaviour. For example, Byström and Järvelin [1995] described formulation as creating a solution space and determining information requirements. Vakkari [1999] argued that after formulation the information seeker has a problem that might be solved, and knows more clearly what information is relevant.

Several studies have explored the progressive reduction of uncertainty as it applies to specific information retrieval behaviours. Tang and Solomon [1998], for example, observed how relevance judgements change with exposure to information. Similar observations were made by Yang in a study of undergraduate students seeking information with a hypermedia database in order to write a class assignment [Yang, 1997]. The system (Perseus) permitted users to create a collection of items for later use. In relation to one participant's exploratory information-seeking at an early stage of the task, Yang commented:

It seemed that he [Eric] had no specific goal or coordinated plan in mind. He appeared to be exploring the database in hopes of hitting on something that might trigger an insight or idea.

[Yang, 1997, p.83]

Similar results to these have been reported by Spink et al. [2002].

Adopting and refining Kuhlthau's ISP model, Vakkari et al. [Vakkari, 2000a; Vakkari, 2000b; Vakkari & Hakala, 2000; Vakkari & Pennanen, 2001; summarised in Vakkari, 2001] observed the impact of formulating a task focus on the search tactics, search terms, relevance judgments and sources used by a group of students writing research proposals for their masters theses. Over a four month period they found that levels of formulation could be systematically related to the use of increasingly more specific search terms and more discriminating relevance judgments.

Attfield, Blandford and Dowell [2003] used Sharples' [1996] idea of writing as a type of design activity as a point of departure for understanding information need uncertainty in relation to theories and findings from the Design Studies literature. For example, design problems are frequently radically under-specified and uncertain with this uncertainty resolved through iterations of analysis and synthesis [Lawson, 1997; Schön, 1983]. Attfield, Blandford and Dowell noted that where information seeking is embedded within a wider task, so a reciprocal relationship occurs such that information seeking is shaped by the needs of the task, and yet the evolving task is shaped by the information found.

We have discussed research relating to information seeking uncertainty, and how this relates to uncertain and evolving relevance judgements (which underpin information gathering decisions). We now turn to research related to the ways in which knowledge workers gather and use source information.

2.2 Information gathering and Use

A number of researchers have reported studies of the way knowledge workers gather and use information. For example, Sellen, Murphy and Shaw [2003] reported a diary study of Web activity of knowledge workers (including two journalists) over a two-day period. Among other findings, they discuss 'finding' and 'gathering'. Finding corresponded with locating specific facts which might be maintained for reference on a temporary basis. 'Gathering' involved locating and storing information to address questions which were difficult to specify. Of all the activities they reported, information gathering was most frequent, and consumed most time with 40% of instances extending over multiple sessions. Once gathered, information was stored in a number of ways, although printing was preferred since this allowed documents to be close to hand and provide task context.

Jones, Dumais and Bruce [2002] similarly reported information gathering or 'keeping' behaviours of a group of researchers, information professionals and managers. In order of frequency of use, they observed: sending emails to oneself, sending emails to others, printing, saving documents as files, pasting URLs into documents, putting documents onto personal websites, creating bookmarks, writing paper notes, creating toolbar links, and using the note facility in Microsoft Outlook. Jones, Dumais and Bruce [2002] also explored factors influencing method choice, which included: portability; the subsequent number of access points (a rationale, for example, for sending oneself an email, which is often accessible from multiple locations); ease of (re-)capturing information context (i.e. the intended purpose of the information); how well a method prompted further action; and ease of maintaining organisational schemas. Hyams and Sellen [2003] similarly looked at how a group of knowledge workers gathered information specifically from the Web and observed similar keeping methods including: copying and pasting; saving documents as files; printing; bookmarks; archiving in email; hand-written notes; and saving in personalised Web folders.

Implicit within an information storage strategy is the idea of facilitating relocation. Where a lot of information is involved, users may promote easier relocation by creating an organisational schema [Jones, Dumais & Bruce, 2002]. Hyams and Sellen [2003] found that personal information collections tended to be organised by project or by topic, and gathered information was rarely used outside of the boundaries of a project.

Whilst information may be stored (and consequently accessed) in various ways during an assignment, the need for fluid access and interchange between cycles of analysis and synthesis during intense periods of work may place particular demands on the way information is managed. O'Hara et al [2002] report a study of participants interacting with multiple source documents during everyday writing tasks. They focused on the movement of attention between documents and composition, how spatial layout is managed to support cognition, the role of annotation and mark-up, and the simultaneous use of paper and electronic documents.

They observed periods of frequent attentional shifts between source documents and composition during writing. These occurred, for example, when looking for a specific fact, comparing information, paraphrasing text or copying—including local experiments to test the effect of source material on the flow of writing. One participant glanced back and forth 13 times between two printouts before writing a single idea. During these periods participants might use a finger or annotations as spatial markers.

O'Hara also observed participants spreading documents out on their desks to enable visual availability without sacrificing visibility of the composition. This layout could change many times as different source documents became the focus of attention. On occasion, participants used paper and electronic versions of the same document concurrently. For example, they might use a word-processor's search facility to locate specific information, but might also browse a paper version where they had more 'implicit awareness' of something useful. O'Hara et al conclude by discussing the limitations of traditional single monitor display arrangements for supporting concurrent visibility.

This research shows that the creation of personal information collections is an important component of knowledge work. As they encounter useful information, so knowledge workers orchestrate the means for re-encountering it later, perhaps multiple times, and with ease. Information is marshalled into bespoke, project-specific collections, in ways that depend upon the affordances of the method and requirements of the task at hand. When the time comes for cognitively intense parts of an activity, however, additional demands are placed on information availability. To promote task fluidity, users manipulate the information they need so as to minimize access costs and interruption.

This, combined with the fact that users are often uncertain in advance about what information they need presents a dilemma for which we explore drag-and-link as a solution. We do this using news research and writing as an example information-intensive task domain. First, though, we briefly summarise the field study of the information behaviours of news journalists working in a national newspaper. This study extends the notion of uncertainty to differentiate between intrinsic and extrinsic factors, and provides the primary point of reference for the system and task that we used for our evaluation.

3 A newsroom study

The aim of the newsroom study [Attfield, 2005; Attfield & Dowell, 2003] was to understand the information behaviour of newspaper journalists in relation to electronic news cuttings services, and to explain these in terms of constraints on the task and the nature of the resources that the journalists had available—an idea taken from Cognitive Systems Engineering (CSE) [Rasmussen, Pejtersen & Goodstein, 1994]. Data were gathered primarily through semi-structured interviews conducted with twenty-five journalists.

Working within a newsroom, the journalists' assignment deadlines were short, typically within the same day. The essence of what was to be reported was a particular 'angle'. This would be negotiated early in an assignment, and took the form of a statement (perhaps a conjecture) about an event or state of affairs. Where the angle was a conjecture, the journalist's task would include validating it.

A Chief Sub-editor explained how all news stories are driven by an angle, and how the angle makes 'sense' out of facts:

... Essentially there is an *angle* to all news and features; it is really a working hypothesis that translates the gathered *facts*, which may include some speculation, into a coherent account.

For each assignment, a journalist would compile a temporary and informal collection of information (i.e. information 'holding patterns') such as a word-processor document or annotated printouts.

The study divided behaviours into information-seeking, information-gathering and information-reviewing. Information-seeking included checking archived news reports to see if an angle was original, seeking background information, seeking evidence for an angle, and making comparisons with similar events. Where information seeking provided useful information it would be gathered into the temporary collection by drag-and-drop or printing. Given the focus of the current paper we will say a little more about these:

- **Drag-and-drop** was used to transfer text from a browser window to a word-processor document using a split-screen arrangement. Dragging and dropping tended to be preferred where time was short and only a small amount of information was wanted. A number of journalists offered terms to describe this document, including: 'holding document', 'work-paste-pad', 'information basket' and 'kind of database'.
- **Printing** tended to be used where more time was available, where a large amount of information was considered useful, or where mobility was important. Reading from printouts was also considered more comfortable, and some said that printouts supported greater ease of reference when writing than computer based text files. Once printed, documents would usually be highlighted or annotated to draw attention to particular parts.

A distinction between gathering by dragging and dropping into a document and gathering by printing is that printing allows relevance judgements to be more liberal and coarse. Drag-and-drop requires greater precision. Printing, however, took more time and caused interruptions to the work (as a result of sending a document to a

printer and having to collect it). Where the journalists' deadlines were short, this could be prohibitive.

In each case, the goal was to make some information easily available at some later time, particularly during writing. When writing, the journalists might again split their screens between a 'holding document' (if they had one) and a file for writing their copy, and printouts would be arranged on the desk for easy access.

These were the information behaviours identified, but to understand the impact of uncertainty we return to the idea of the angle. Information-gathering depends on the ability to judge the value of information at the time that it is encountered. The task is defined by an angle to be pursued, but this is only a broad characterization. As the assignment progresses, so the journalist's ideas about what information to include is shaped and reshaped. As one journalist put it:

...the ideas will take shape all the time... at the point that they change all the time. It is only really when you have to sit down and actually write it that I would have to decide what way to go into the story... I am preparing it... I am preparing all the time.

Further, the angle can change. As the journalist encounters new information, so they might come to consider the angle unsustainable. And as new information comes in, so the situation being reported can change and editors can change their minds about the angle they want.

The significance of this for information behaviour lies in the effect it has on relevance judgements, which necessarily evolve and change. Much gathered information may not be used, the journalist may need to return to previously read documents to extract new information, and they may need to do more searching. And, once a report is written, they may review their collection for additional information to include. In this way the process, and hence the tools that support it, need to be dynamic and flexible.

4 NewsHarvester design

NewsHarvester² was designed specifically with the aim of being tolerant to the information gathering uncertainty we have described in the context of newswriting tasks. The user-interface is shown in figure 1 as it might look during use. The system incorporates tools for searching a database of news reports and browsing the results (Search Form), viewing full-text documents (Document View Form), and gathering text extracts by dragging them into an integrated text editor (Collection Space/Copy Editor).

Given the need for frequent cross-referencing between source documents and the composition, and, we assumed, searches and source documents, NewsHarvester presents all of the associated spaces concurrently with roughly equal prominence. Within the text editor, extracts can be retained and optionally annotated, edited, or incorporated into a new piece of writing. With the drag-and-link feature, when an extract is dragged into the text editor, it is automatically suffixed with a hyperlink. When clicked, this will navigate the document display back to the document from which the extract had been taken.

² NewsHarvester was developed in collaboration with Microsoft Research, Cambridge, England.

Search Form (left)

Users can construct queries using Boolean syntax with implicit AND. Results are displayed as a list of linked headlines below the query field. When the user clicks a headline in the results list, the full text is displayed in the document view form. For the study reported here NewsHarvester was connected to the Media News Archive at the European Journalism Centre in Maastricht. This provided access to about 10,000 articles originating from newspaper and television sources which are trimmed to about 200 words.

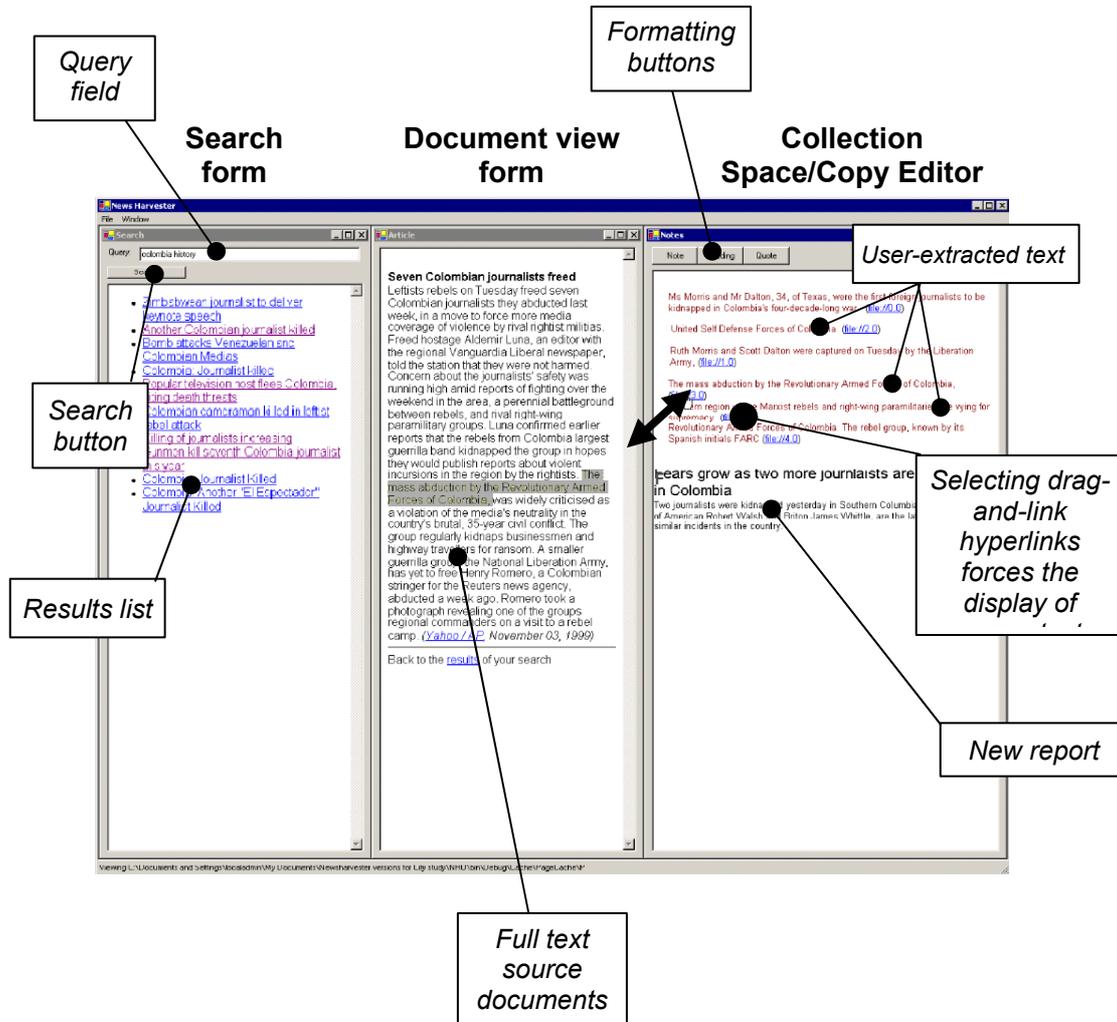


Figure 1. The NewsHarvester interface

In the document view form the report that has been selected is shown. The headline is shown in bold and keyword highlighting is used within the text. If the user finds some text they want to gather, they can highlight it and drag it into the collection space/copy editor.

Collection Space/Copy Editor (right)

The collection space/copy editor acts as the temporary collection space and also as a text editor. Here the user can annotate extracts with contextual information and write their finished report.

During the early design stages, a number of methods for structuring the collection space were considered, including spatial hypertext [e.g. Marshall, Shipman & Coombs, 1994; Shipman et al., 2004; Buchanan et al., 2002]. In a spatial hypertext, information elements are added and freely arranged within a 2D workspace. Although this approach affords flexible structuring, it also places high demands on screen real-estate. It became clear that, for NewsHarvester, the collection space would become cluttered quickly. Consequently, we implemented the collection space as a text editor in which text could be dropped, typed and organised. For space reasons, this text editor also doubled as the place where users would construct their article text.

Figure 1 shows NewsHarvester after a series of text extracts have been added to the collection space and writing the report has begun. Text that has been dropped into the collection space/copy editor is distinguished by being shown indented and in a specific font and colour. Within the collection space/copy editor extracts are automatically suffixed with a hyperlink, which, when clicked, navigate the document view from the document from which the extract was taken. When this happens, the extract in the document view form is highlighted (see figure 1).

When an extract has been dropped into the Collection Space/Copy Editor and a link created, the user can add text around or within the extract. If an extract is split, its link is duplicated so that its separate parts can be used to return to the originating document. This happens for any further splits. In order to provide some structuring tools, users can also force selected text to appear as a heading.

5 Objectives of the study

The study was framed as the evaluation of a set of non-comparative and comparative claims.

5.1 Five non-comparative claims

These were concerned with aspects of user-behaviour when performing a news researching and writing task (NC1 to NC5 in table 1). By evaluating these, the study would test and elaborate findings from the newsroom study which underpin the choice of drag-and-link as a solution. In particular, the study sought to validate the finding that, when researching and writing a news report, journalists do need to relocate information they had not anticipated needing, and some reasons why this might be so.

NC1	When researching and writing a news report, journalists often want to refer to a given archived news report multiple times.
NC2	Non-comparative claim 1 can, in-part be explained by the goal of relocating information that had been read but not previously been identified as useful.
NC3	Non-comparative claims 1 and 2 can in-part be explained by the goal of including specific items of information in a report.
NC4	Non-comparative claims 1 and 2 can in part be explained by goal of understanding the context of information which had previously been identified as useful.
NC5	Non-comparative claims 1 and 2 can in part be explained by the goal of identifying additional information that could be included in a near-complete report.

Table 1. The non-comparative claims evaluated by the study

5.2 Six comparative claims

The newsroom study had shown journalists using two methods for collecting information: on-screen drag-and-drop, and printing (with additional highlighting or annotation). Whilst drag-and-drop was described as quicker and easier (and hence used when time was short), printing was preferred when there was more time and is more forgiving of the kind of uncertainty observed. This is because specific judgments about what information is useful do not have to be made when a document is initially encountered. Drag-and-link was incorporated into NewsHarvester as a third option that might combine ease of gathering with uncertainty tolerance.

The comparative claims (shown in table 2) were concerned with how drag-and-link compared with these two more traditional methods. To test these claims the study was performed with three variations of NewsHarvester, each supporting one of the three methods. Each of the comparative claims is made in terms of a variable (e.g. the perceived user-cost of gathering information). They are also composite, each making two sub-claims about the way drag-and-link compares with each of the other methods (a and b). Hence, claim C1a is that: the user-cost of gathering information with drag-and-link is less than with printing. Each sub-claim was tested independently.

Claim	Variable	Relational Statement				
		Printing	a	Drag-n-link	b	Drag-n-drop
C1	The perceived user-cost of gathering information		>		≤	
C2	The perceived user-cost of relocating gathered information		≥		≤	
C3	The perceived user-cost of relocating non-gathered information from documents that contain gathered information		≥		<	
C4	The perceived affordance of a dynamic and flexible way of researching and writing.		<		>	
C5	User enjoyment		<		>	
C6	Quality of final report		<		>	

Table 2. The comparative claims evaluated by the study

C1, C2, and C3 are concerned with users' perceptions of the ease of three actions: gathering information, relocating gathered information, and relocating non-gathered information from documents that contain gathered information. This third claim may appear convoluted, but it is essential to the evaluation. Supporting easy relocation of documents from which information has already been gathered was a principal motivation for drag-and-link.

Whilst these three actions were significant to the evaluation, alone they may not tell the whole story. For example, drag-and-link might offer net gains for all these actions, and yet impact in some negative but unpredicted way. To address this possibility, comparative claims C4 and C5 focus on more general user experience variables.

C1 to C5 are concerned with the quality of user experience as distinct from issues of task performance. Beyond this we also included C6 in order to address the possibility that drag-and-link may offer gains in terms of user-performance when using the system.

6 Method

The participants were fifteen students recruited from the Department of Journalism at City University in London. Thirteen were studying towards an M.A. in International Journalism and 2 were studying for a BA in Journalism. All had previously worked as professional journalists with experience ranging from 3 months to 10 years. Many different nationalities were represented but all of those who had English as a foreign language had obtained a score of 7.0 or above in the IELTS English language test. Participants were paid £30.

Each of the three variations of NewsHarvester (drag-and-link, drag-and-drop and printing) corresponded with a condition in a three condition, repeated measures, cross-over design. For the printing condition, a printer was placed on the desk beside the computer and highlighter pens and ordinary pens were provided. This was intended to provide a 'best case' condition for the printing variation.

Given the repeated measures design, three news research and writing assignments were devised. To ensure authenticity and balance across the tasks, they were designed in collaboration with an Executive Editor at The Times. Each task involved writing a report about a 'breaking' news event for which additional background information would be required. To avoid participants finding real reports on the events, all were imaginary. However, to ensure authenticity and the availability of relevant background material, most events involved real people, organizations and places, and in all cases piloting was performed to ensure that they were credible in terms of the news context.

Each task was explicitly designed to require roughly equal amounts of background material. For example, for one task participants were told of two journalists who had been kidnapped by a guerrilla group in Columbia. The task was to write a piece to go alongside a main article ('a sidebar') with the angle: 'This is the latest in a series of journalist kidnappings in Columbia'. Each assignment had a 50 minute deadline, and each condition was completed in a single session with no more than two sessions performed in any one day.

At the beginning of each session, the system to be used was explained to the participant and they were handed a page of instructions and a newswire and given some editorial instructions in the form of a required angle and word count. Thirty minutes into each task, the participants were interrupted and handed a second newswire with new information. In the light of this new information they were given a new angle and word count to work to. In the case of the kidnapped journalists example, the new information was that the journalists had been found dead and the new angle was to be 'This is the latest in a spate of journalist killings in Colombia'.

To control for potential order effects between conditions (system designs), the order was balanced using a single Latin square, with each participant randomly assigned to one of three sequence groups (see table 3). To balance for unforeseen interactions between conditions and tasks (i.e. that a design might be more suited to one task), the task sequence was kept the same for all groups. This ensured that each task was performed the same number of times with each design.

	<i>Task 1</i>	<i>Task 2</i>	<i>Task 3</i>
Group 1	printing	drag-n-drop	drag-n-link
Group 2	drag-n-link	printing	drag-n-drop
Group 3	drag-n-drop	drag-n-link	printing

Table 3. The three condition sequences

Non-comparative claim 1 was assessed using a protocol of document consultations produced for each session. This was taken from recorded screen movies in combination with a manual log of consultations during printing sessions. For the purposes of the protocol, a single consultation (or re-consultation) was defined as a period of reading which could be interrupted by any other activity except for the reading of another document. At such a point, the initial consultation was deemed to have ended and a new consultation begun. During printing sessions, participants might quickly flick through one or two documents before settling on one to read. A participant viewing a document for a period in the order of only a second was not regarded as a consultation.

Non-comparative claims NC2 to NC5 and comparative claims C1 to C5 were assessed by measuring participants' levels of agreement with items on a post-task questionnaire administered immediately following each session. Think-aloud protocols were not taken since we considered that this might unduly compromise task-flow.

A single questionnaire was used for all sessions, with a rating given for each item on each condition. Questions were presented as statements with an accompanying visual analogue scale (VAS) on which to indicate agreement. The scale ends were calibrated with "strongly disagree" (scored 0) and "strongly agree" (scored 10). Given the importance of between-condition differences for the study, in the second and third trials participants were encouraged to review their previous ratings in order to accurately reflect their comparative judgement (participants frequently adjusted previous ratings following their experience of a new version of the system)

In order to assess comparative claim C6, a sub-editor at the Guardian was asked to perform a blind quality rating of the final 45 reports. To simplify comparisons, the reports were grouped by task (3 groups of 15) and comparatively rated within their groups. Each group contained 1 report from each participant and equal numbers from each condition. The rater ranked the reports using a five point scale, making judgements according to 'the extent and relevance of the background content and the quality of its delivery in the context of the brief'. Judgments were indicated by sorting printouts into piles. All ratings could be reviewed and adjusted throughout the ranking of reports from each group.

The document consultation logs and screen movies were also used for an exploratory analysis of document re-consultations and dragging behaviour. The results of this analysis are reported at the end of the results section.

Results

6.1 Non-comparative claims

Non-comparative claim 1 was evaluated using the log of document re-consultations (irrespective of condition). Given difficulties in quantifying 'often', a hypothesis was

not formulated. However, out of a total of 45 assignments, there were only 5 in which participants did not re-consult documents and the maximum number of re-consultations in any one session was 15. The mean number of first-time document consultations across all conditions was 14.2, and the mean number of re-consultations was 6.2 (30% of total consultations). Hence the data supports non-comparative claim 1.

Whilst the current study does not allow us to assess the effect of the interruption on the number of first-time consultations and re-consultations (since all tasks featured the interruption), of the first-time consultations, 42% occurred before the interruption and 58% after. Of the re-consultations, 30% happened before the interruption and 70% after. Hence, it can be concluded that needing to refer back to documents is not only a result of an interruption that changes the task slightly.

For each of non-comparative claims 2, 3, 4 and 5 participants rated questionnaire statements at the end of each of the three sessions, providing three ratings per participant overall. Since condition differences were not important for the non-comparative claims, aggregated ratings were calculated as the mean response from each participant to each statement across the three conditions. The VAS statements and mean responses are shown in table 4.

Non-comparative claim	Statement	Mean VAS score (1 to 10)
2	During the task, I found I wanted to re-consult source documents to find information I had remembered reading but did not necessarily consider useful at the time.	7.9
3	This happened because later I found I wanted to include specific items of information in my report.	8.5
4	This happened because later I wanted to better understand the context of information which I had identified as useful.	8
5	This happened because later I wanted to re-consult source documents to check if there was anything else I could add to my report.	8.6

Table 4. Statements used in the post-trial questionnaire to test non-comparative claims 2, 3, 4 and 5, and the mean responses obtained (on a scale of 1 to 10).

Since the poles of the VAS scales were calibrated as “strongly disagree” (0) and “strongly agree” (10), a mid-point score of 5 was taken to signify indifference. The data on all statements were normality distributed (Shapiro-Wilk, $p > 0.05$). One-tailed, one sample t-tests showed that, for all statements, responses were greater than a hypothetical mean of 5 with a very high level of significance ($p < 0.0001$ in all cases). These results show strong support for non-comparative claims 2 to 5.

Comparisons between the results for each statement and each of the others showed that no explanation was rated significantly higher than any other (two-tailed t-test, $p > 0.05$).

6.2 Comparative claims

Claims C1 to C5

Due to a ceiling effect, many of the ratings participants gave for the questionnaire items used to assess claims C1 to C5 were non-normally distributed (Shapiro Wilks,

$\alpha = 0.05$). Figure 2 shows the statements set against median responses for each condition (bars are labelled A to O for ease of reference).

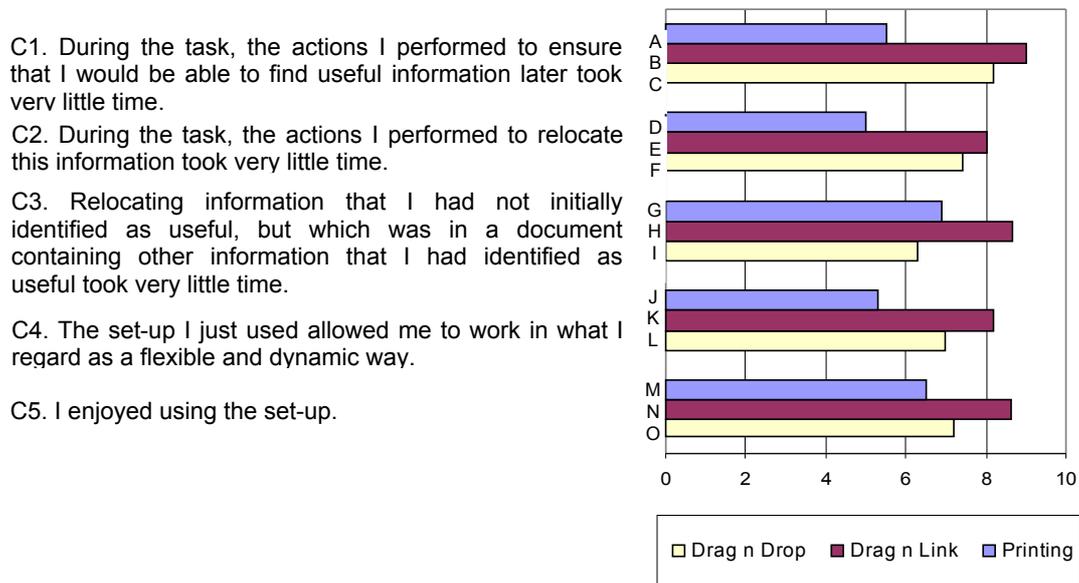


Figure 2. Statements and participants' levels of agreement corresponding to comparative claims C1 to C5

Each sub-claim was tested for statistical significance as an independent comparison. Where a sub-claim stated the superiority of drag-and-link over a reference method, difference testing was used (Wilcoxon's). Failing to show that two populations are significantly different, however, is not a guarantee that they are equivalent [Senn, 1997], or indeed that one is non-inferior to the other. Where a sub-claim stated the non-inferiority of drag-and-link, a non-inferiority test [Senn, 1997] was used (a variety of equivalence testing). Accordingly, an experimental treatment is considered non-inferior to a reference treatment on a dependent measure, so long as the lower confidence limit for the difference between conditions is above the lower bound of a region of practical equivalence. For the current study, the region of practical equivalence was taken as $\pm 15\%$ ³ of the mean of the reference sample. Confidence intervals were calculated at the 95% level using the W statistic.

Table 5 shows the comparative claims C1 to C5 (repeated from table 2) with median ratings and test results added. Where each user-interaction variable relates to a condition, a capital letter is shown to indicate how it corresponds with the results in figure 2. For each superiority sub-claim, obtained p values are shown. For non-inferiority sub-claims, a value (diff) is shown corresponding to the difference between the lower confidence limit and the lower limit of the region of practical equivalence. Where this value is positive (all cases), the data supports the sub-claim (at $p < 0.05$) i.e. drag-and-link is at least as good as the reference system on that measure. In each case the corresponding W value is shown.

³ Conventionally, 20% is usually taken as the region of practical equivalence. 15% was used in the current study, placing slightly greater demand on the data to demonstrate equivalence.

Claim	User Interaction Variable	Relational Statement				
		printing	a	drag-n-link	b	drag-n-drop
C1	The perceived user-cost of gathering information	5.5 (A)	> p < 0.01 W = 112	9 (B)	≤ diff = 0.68 W = 63.5	8.2 (C)
C2	The perceived user-cost of relocating gathered information	5 (D)	(**) ≥ diff = 2.3 W = 104	8 (E)	(*) ≤ diff = 1.31 W = 74	7.4 (F)
C3	The perceived user-cost of relocating non-gathered information from documents that contain gathered information	6.9 (G)	(*) ≥ diff = 0.88 W = 70	8.7 (H)	< p < 0.01 W = 96	6.3 (I)
C4	The perceived affordance of a dynamic and flexible way of researching and writing.	5.3 (J)	< p < 0.01 W = 91	8.2 (K)	> p < 0.01 W = 86.5	7 (L)
C5	User enjoyment	6.5 (M)	< p < 0.01 W = 96	8.6 (N)	> p < 0.01 W = 100.5	7.2 (O)

Table 5. Results of statistical tests for comparative claims C1 to C5

* Drag-n-link is significantly better than the reference condition at $p < 0.05$

** Drag-n-link is significantly better than the reference condition at $p < 0.01$

Throughout the comparative usability claims, where the superiority of drag-and-link was predicted over printing or drag-and-drop (sub-claims C1a, C3b, C4a, C4b, C5a and C5b), the results support the claim at a high level of significance. This is the case for assessments of perceived user-cost of specific kinds of action, as in claims C1 to C3, and also for the more abstract user-experience variables (C4 and C5). For claims where predictions were made about the non-inferiority of drag-and-link with respect to printing or drag-and-drop (C1b, C2a, C2b and C3a) this was also demonstrated. Thus comparative claims C1 to C5 are supported by the questionnaire data.

The statement ratings in figure 2, however, also suggest that drag-and-link may have been rated significantly higher than reference conditions where only non-inferiority was predicted. In other words, rather than being rated as being 'as good as' the reference conditions, drag-and-link might have been rated unexpectedly and systematically higher. This was a possibility for all the non-inferiority claims. Compare B with C (claim C1b), D with E (claim C2a), E with F (claim C2b), and G with H (claim C3a). Difference tests were used to explore whether these were real differences (two-tailed Wilcoxon). Significant differences were found in three of the four sub-claims (indicated in table 5 with asterisks). Specifically:

- C2a - users rated their perception of the costs of relocating gathered information with drag-and-link as significantly lower than with printing ($p < 0.01$)

- C2b - users rated their perception of the costs of relocating gathered information with drag-and-link as significantly lower than with drag-and-drop ($p < 0.05$)
- C3a - users rated their perception of the costs of relocating information that had not been gathered, within documents containing gathered information as significantly lower with drag-and-link than with printing ($p < 0.01$)

Considering a potential halo effect

As discussed above, drag-and-link was rated significantly more positively than the reference conditions even where this was unexpected (i.e. comparative claims 2 and 3). This may suggest that responses to the questionnaire items were subject to a confounding halo effect [Saal et al, 1980]. According to this interpretation, having developed a preference for the drag-and-link version of NewsHarvester, participants rated it more favourably irrespective of what individual questionnaire items were asking.

An analysis was performed to assess whether a halo effect could be eliminated by measuring the strength of correlation between each participants' responses to different items within each version of the system (condition). Low correlations would show absence of a halo effect, whilst high correlations would be consistent with it (if not conclusive). Further, as a direct measure of user-enjoyment, responses to the comparative claim 5 provided a baseline for detecting a halo effect.

The results of the correlational analysis (see table 6) show that C4 correlates strongly with C5 and this is highly significant. C1, C2 and C3 do not have strong correlations with C5, but they correlate significantly with each other. C3, however does not have strong correlations with C1 and C2, even though these are significant. Hence, C3 can be considered independent of C1 and C2 to some extent.

Question	C1	C2	C3	C4	C5
C1	1.0	.675(**)	.372(*)	.008	-.031
C2		1.0	.312(*)	.326(*)	.212
C3			1.0	.176	.108
C4				1.0	.735(**)
C5					1.0

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Table 6. Summary of correlations between comparative claim questions.

These relationships are confirmed by principal components analysis. Using the latent root criterion for the selection of factors, a two factor model accounts for 74% of the variance in the data. The first factor is made up of C1, C2 and C3 and the second of C4 and C5. However, using the scree criterion results in a three factor model accounting for 89% of the variance in the data. In this model, the first factor is C4 and C5, the second is C1 and C2 and the third is C3. This confirms that, though C3 does seem to have some relationship with C1 and C2, it can reasonably be considered to have been answered independently from the other two claims.

Having found that the questions are producing this pattern, it may be that the data on one particular system is dominating these correlations. However, a further analysis of correlations between the questions only divided in to the separate systems produces a broadly similar picture with high correlations between C1 and C2 and between C4 and C5 in all three conditions, and weaker correlations between C3 and any of the other questions.

Claim C6

The results of the report rating exercise and statistical tests are shown in table 7. Ratings are on a scale of 1 (poor) to 5 (good).

Claim	User Interaction Variable	Relational Statement				
		printing	a	drag-n-link	b	drag-n-drop
C6	Quality of final report	2.73	< p > 0.05 W = 22	3.07	> p > 0.05 W = 24	2.78

Table 7. Mean ratings of final reports for each condition according to independent rating, and results of statistical tests for comparative claim C6

The results show that the mean rating for reports written under the drag-and-link condition were higher than for both the printing and drag-and-drop conditions. However, one-tailed Wilcoxon's tests showed that these differences were not significant ($p > 0.05$).

6.3 Exploratory analysis

Document re-consultations

Comparative claim C3 was concerned with the perceived user-cost of relocating non-gathered information from documents containing gathered information. Although the document consultation log did not provide direct evidence for this claim, it was possible to compare re-consultation frequencies for the different conditions. On the assumption that, for a given level of motivation, a task is performed more frequently the easier it is [Ross & Nisbett, 1991, p10], this frequency was taken as an inverse function of user-cost.

A precondition for the re-consultation of a document is that it should already have been consulted. Since the number of documents consulted varied from session to session, so did the number of opportunities for re-consultation. Hence, the comparison was made on the basis of the number of re-consultations per document consulted in a session. These data were normally distributed for all conditions (Shapiro-Wilk, $p > 0.05$). The mean re-consultation rates and standard deviations for each condition are shown in table 8.

These results show that the mean re-consultation rate for drag-and-link is higher than for printing and almost twice that of drag-and-drop. A two tailed t-test showed that the difference between drag-and-link and drag-and-drop was highly significant ($p < 0.01$). However, a similar comparison between drag-and-link and printing was non-significant ($p > 0.05$).

	Mean re-consultation rate	Std dev
Printing	0.45	0.30
Drag-and-link	0.61	0.34
Drag-and-drop	0.34	0.25

Table 8. Document re-consultation rates by condition

Dragging behaviour

Using the screen movies, an analysis was performed to see whether there were differences between the drag-and-link and drag-and-drop conditions in terms of how many times participants performed drag operations, how many words they dragged each time, and from this, the average number of words per drag operation. One goal of this analysis was to see whether participants dragged less information in the drag-and-link condition (knowing that they could easily return to documents). If this were the case, it might offer an explanation for the fact that participants rated drag-and-link as providing easier access to gathered material than drag-and-drop (claim C2b) since finding information within a smaller collection would be easier.

A sample of ten videos from each of these two conditions (2/3) were reviewed to obtain a count of drag operations, a count of the number of words dragged in each operation, and the mean words copied per drag operation. This analysis showed no systematic differences on any of these measures between conditions.

7 Summary and discussion

This study had two primary aims. The first was to evaluate a set of claims (non-comparative) concerning the need to refer back to source documents during a news research and writing task. The second was to evaluate claims (comparative) about drag-and-link in comparison with printing and drag-and-drop as a way of supporting information gathering and relocation. Beyond this we also performed some exploratory analysis of document re-consultations and dragging behaviour.

Non-comparative claims

On the first issue, we found that participants wanted to refer back to documents multiple times and this need could occur independently of the mid-task change we imposed. This indicates that changes in relevance criteria can arise as a result of intrinsic as well as extrinsic factors. We also found strong support for three motivations for relocating documents. These were:

- Targeted information re-finding (i.e. specific items) for inclusion in a composition
- Reviewing information context
- Opportunistic information re-finding (i.e. non-specific items) for inclusion in a composition

With respect to the second finding, we did not unpack which aspects of information context are important, and how reviewing them helps; this is a question for further study. Our hypothesis is that users want to review issues such as the originating sources of information (e.g. to judge authority), date of publication, and other

information that is being reported within a document, so as to assess how they might make use of the information they have.

We can also see that people sometimes know what they are looking for and sometimes they don't. The idea of opportunistic re-finding corresponds closely with O'Hara et al's [2002] observation of users browsing documents with an 'implicit awareness' of something useful. They have a feeling there may be useful information to be found, but they need prompting.

In research and writing tasks, like the one used in this study, there may be motivations for re-encountering documents beyond the ones we have considered here. But in themselves, they are strong indicators of the need to support low-cost document review as we have attempted to do—to enable people to interact fluidly with the documents they have seen and perhaps read in detail. This is also consistent with findings by O'Hara et al [2002]. This need points to uncertainty as intrinsic to this kind of task, and, as many have argued before, systems need to be designed to accommodate this (for example, Kuhlthau [1993], Kuhlthau and Tama [2001]). The important question, of course, is how?

Drag-and-link is presented as one possible solution. Whilst drag-and-link is limited by the fact that it depends upon users initially gathering some information from source documents, it has the advantage of incidentally allowing ad-hoc collections to act as indexes to source materials. And since these collections are actively created by users, evolve over time, and can be made highly visible at the interface, we might assume that users can develop a high degree of familiarity with their content and structure. Consequently, they may be very usable as indexes.

Comparative claims

Testing of the comparative claims through the post-task questionnaire showed very positive preferences for drag-and-link. Participants felt that that drag-and-link supported information gathering in a way that is as easy as drag-and-drop, and easier than printing.

In terms of subsequent ease of access to gathered information, the questionnaire results suggested an advantage over both printing and drag-and-drop. An advantage over drag-and-drop, however, is difficult to explain. One possibility was that a drag-and-link collection could be smaller (due to lower re-consultation costs). But the exploratory analysis of dragging behaviour showed that this was not the case. The possibility of the results being confounded by a halo effect was also eliminated by the correlational analysis. However, this did show a lack of independence between the question about re-access and the question about gathering. And since the basis for concluding an advantage in terms of access to gathered information was only small and marginally significant (the median difference was 0.6 on a 0 to 10 rating scale), we will not draw solid conclusions here. The re-access advantage of drag-and-link over printing, however, was large and highly significant. That drag-and-link within a high-visibility, multiple work-space, integrated system showed advantages over paper, at least within the context of the task we studied, is a very promising result given the frequent preference that users have for paper.

With respect to perceived user-costs of relocating documents containing gathered information, drag-and-link was rated more favourably than drag-and-drop as expected, but also more favourably than printing. This difference was marginally significant, but not negligible in terms of magnitude. This result might be explained in terms of a difference between sorting through printouts piled or scattered on a desk

and doing the same with links embedded within an onscreen text file. Once again, drag-and-link appears to show an advantage over printing as an information gathering method (as well as drag-and-drop in this case). Providing a more general measure of document relocation, the analysis of document re-consultations also showed a significant advantage over drag-and-drop, but (despite a mean difference in the right direction) the difference between drag-and-link and printing was non-significant.

The results of the more abstract usability measures showed a very positive response to drag-and-link in comparison with the reference conditions. With a high level of significance, participants thought that the drag-and-link version of NewHarvester afforded a more flexible and dynamic way of working and they enjoyed using it more. Also, the correlation we found between responses to C4 and C5 demonstrates the important relationship between these two factors i.e. the extent to which participants enjoyed using the systems was highly correlated with the extent to which they regarded it as affording a 'dynamic and flexible way of researching and writing'.

Finally, the assessment of the quality of news reports written with the three system configurations showed a slightly more positive rating for reports written with drag-and-link, although the differences were not significant. Two interpretations are possible: the first is that drag-and-link does not result in quality improvements over drag-and-drop and printing. The second interpretation is that the results indicate an effect, but that this was not strong enough for the unequivocal results in the context of the study paradigm we used. A factor which argues in favour of this is that, whilst ratings for all participants were distributed across the full rating scale, they were fairly consistent for each participant. The mean difference between participants' highest and lowest scores accounted for only 1/3 of the rating scale. Hence, any effect of drag-and-link on quality may have been submerged under participant differences.

However, what we can conclude on the basis of the current study is that the principal effect of drag-and-link over printing and standard drag-and-drop occurs in terms of improving the user-experience of performing research and writing tasks of the type studied here.

8 Conclusions

Previous research has shown how relevance criteria can change during complex information tasks [viz. Tang and Solomon, 1998; Yang, 1997; Spink et al. 2002; Vakkari, 2001]. In research and writing tasks, changes in relevance criteria can trigger the need to find previously seen but non-gathered information. Sometimes the target is recalled, sometimes it is not.

Information gathering is performed to make useful information more accessible, but a task-based personal information collection can prove insufficient. Information gathering can only be motivated by relevance criteria as they stand at the time of an information encounter. Also, whilst a single piece of information from a document may be considered important, the document context in which it was originally delivered may also be important for its interpretation. The current study showed that, in information gathering, enabling users to easily review this context is important.

In summary, the study demonstrates the value of the drag-and-link solution in terms of user-experience during hybrid information seeking and authoring tasks. The work domain we have focused on is journalism, but the apparent ubiquity of personal, ad hoc information gathering, rapid switching between source documents and

composition in writing tasks, and uncertainty concerning relevance judgements suggests much broader applicability.

To design systems which are sympathetic to information uncertainty, it is necessary to support actions in ways which not only involve low costs to the user, but also represent reversible commitments—that they are tolerant of uncertainty and so not cast in stone. In the case of information gathering, the actions are to select and later relocate information encountered during information seeking. Uncertainty tolerance, in this case, entails ensuring that users can easily re-encounter source information where the relevance judgements that underpin information gathering change. Further, information selected now may not be the information that is wanted later, but it is likely to be sourced from documents with which users have already engaged in some depth.

Drag-and-link is one way of integrating information gathering/authoring tools with information seeking tools. The result is that users can develop a more interactive and flexible information holding pattern that is well suited to the vagaries of information intensive work. As such it demonstrates one way in which tools can be integrated such that actions that occur between them provide leverage for functionality to positively assist the user.

Acknowledgements

This work was in part funded under a CASE studentship funded by the EPSRC and Microsoft Research Laboratories, Cambridge, and in part under the project Making Sense of Information (EPSRC Grant EP/D056268). We would also like to thank Microsoft Research Laboratories Cambridge for contributing to development work on the NewsHarvester prototype, Richard Dixon of The Times for helping to design the news-writing tasks, and Rick Williams of The Guardian for rating the results.

References

Attfield, S., 2005. Information seeking, gathering and review: Journalism as a case study for the design of search and authoring systems, PhD Thesis, University of London.

Attfield S., Blandford A. & Dowell J., 2003. Information seeking in the context of writing: a design psychology interpretation of the 'problematic situation'. *Journal of Documentation*. 59(4). pp. 430-453.

Attfield, S., Dowell, J., 2003. Information seeking and use by newspaper journalists, *Journal of Documentation*. 59(2), pp. 187-204

Bates, M., 1989. The design of browsing and berrypicking techniques for the on-line search interface. *Online Review*, 13(5), pp. 407-431.

Belkin, N., Oddy, R. & Brooks, H., 1982a. ASK for information retrieval: part I. *Journal of Documentation*. 33(2), pp. 61-71.

Belkin, N., Oddy, R. & Brooks, H., 1982b. ASK for information retrieval: part II. Results of a design study. *Journal of Documentation*, 38(3), pp. 145-164.

Buchanan, G., Blandford, A., Jones, M. & Thimbleby, H., 2002. Spatial Hypertext as a Reader Tool in Digital Libraries. In Börner, K. & Chen, C. (Eds.) *Visual Interfaces to Digital Libraries*, 13-24, Springer Verlag, LNCS Series, Volume 2539.

Byström, K. & Järvelin, K., 1995. Task complexity affects information seeking and use. *Information Processing & Management*, 31(2), pp. 191-213.

Hyams, J. and Sellen, A., 2003. How knowledge workers gather information from the Web: Implications for peer-to-peer file sharing tools. *Proceedings of the British HCI Conference*, Bath, UK, Sept. 2003.

Jones, W., Dumais, S. and Bruce, H., 2002. Once found, what then?: A study of 'keeping' behaviors in the personal use of web information. *Proceedings of ASIST 2002*, Philadelphia, Pennsylvania.

Jones, W., Munat, C., Bruce, H., and Foxley, A., 2005. The Universal Labeler: Plan the Project and Let Your Information Follow. *Proceedings of the 68th Annual Meeting of the American Society for Information Science and Technology (ASIST)*.

Karger, D., and Jones, 2006. Data Unification in Personal information Management. *Communications of the ACM*, 49(1).

Kuhlthau, C., 1988a. Developing a model of the library research process: investigations of cognitive and affective aspects. *Reference Quarterly*, 28, pp. 232-242.

Kuhlthau, C., 1988b. Perceptions of the information search process of users in libraries: a study of changes from high school through college. *Information Processing and Management*, 24, pp. 419-427.

Kuhlthau, C., 1988c. Longitudinal case studies of the information search process of users in libraries. *Library and Information Science Research*, 10, pp. 251-304.

Kuhlthau, C., 1993. A principle of uncertainty for information seeking. *Journal of Documentation*, 49(4), p.39-55.

Lawson, B., 1997. *How designers think: The design process demystified*, 3rd ed. Oxford, Butterworth-Heinemann.

Marshall, C., Shipman, F. and Coombs, J., 1994. Spatial Hypertext supporting emergent structure. *Proceedings of the 1994 ACM European Conference on Hypermedia Technology*, Edinburgh, UK, ACM Press, NY, pp.13-23.

O'Hara, K., Taylor, A, Newman, W. & Sellen, A., 2002. Understanding the Materiality of Writing from Multiple Sources. *International Journal of Human-Computer Studies*. 56(3), pp. 269-305.

Rasmussen, J. Pejtersen, A.M. & Goodstein, L.P., 1994. *Cognitive Systems Engineering*. New York, Wiley.

Rittel, H., and Webber, M., 1973. Dilemmas in a General Theory of Planning. *Policy Sciences*, Vol. 4, Elsevier Scientific Publishing Company, Inc., Amsterdam, pp. 31-39.

Ross, L., and Nisbett, R.E., 1991. *The Person and the Situation*. Temple University Press, Philadelphia.

Saal, F., Downey, R. & Lahey, M., 1980. Rating the ratings: Assessing the psychometric quality of rating data. *Psychological Bulletin*, 88(2). pp.413-428.

Schön, D.A., 1983. *The reflective practitioner: How professionals think in action*. New York, Basic Books.

Sellen, A., Murphy, R. & Shaw, K., 2002. How Knowledge Workers Use the Web. *Proceedings of CHI 2002*, Minneapolis, MN. New York: ACM Press, pp. 227-234.

Senn, S., 1997. *Statistical issues in drug development*, Chichester, UK, Wiley

Sharples, M., 1996. An account of writing as creative design. IN: Levy, C.M., & Ransdell, S. eds. *The science of writing*. New Jersey, Lawrence Erlbaum.

Shipman, F., Hsieh, H., Moore, J., Zacchi, A., 2004. Supporting Personal Collections across Digital Libraries in Spatial Hypertext, *Proceedings of the ACM and IEEE Joint Conference on Digital Libraries*, 2004, pp. 358-367.

Spink, A. Wilson, T.D. Ford, N. Foster, A. & Ellis, D., 2002. Information seeking and mediated searching study, Part 3, successive searching. *Journal of The American Society for Information Science and Technology*, 53(9), pp. 716-727.

Tang, R. & Solomon, P., 1998. Toward an understanding of the dynamics of relevance judgment: An analysis of one person's search behaviour. *Information Processing and Management*, 34(2-3), pp. 237-256

Taylor, R., 1968. Question-negotiation and information seeking in libraries. *College and Research Libraries*. 29. pp. 178-194.

Vakkari, P., 1999. Task complexity, problem structure and information actions: Integrating studies on information seeking and retrieval. *Information Processing and Management*, 35, pp. 819-837.

Vakkari, P., 2000a. Cognition and changes of search terms and tactics during task performance: a longitudinal study. *Proceedings of the RIAO 2000 conference*, Paris, CID. pp. 894-907

Vakkari, P., 2000b. Relevance and contributory information types of searched documents in task performance. *Proceedings of SIGIR 2000 Conference*, Athens, New York, ACM, pp. 2-9.

Vakkari, P., 2001. A theory of the task-based information retrieval process: a summary and generalisation of a longitudinal study. *Journal of Documentation*, 57(1), pp. 44-60.

Vakkari, P. & Hakala, N., 2000. Changes in relevance criteria and problem stages in task performance. *Journal of Documentation*, 56(5), pp. 540-562.

Vakkari, P. & Pennanen, M., 2001. Sources, relevance and contributory information of documents in writing research proposal: a longitudinal case study. IN: Höglund, L. & Wilson, T.D. eds. *The new review of information behaviour research: studies of*

information seeking in context (Proceedings of ISIC 2000), Cambridge, Taylor Graham. pp. 217-232.

Yang, S., 1997. Information seeking as problem-solving using a qualitative approach to uncover the novice learner's information-seeking processes in a Perseus hypertext system. *Library & Information Science Research*, 19(1), pp. 71-92.