Revisiting Read Wear:
Analysis, Design, and Evaluation of a Footprints Scrollbar

Jason Alexander[1], Andy Cockburn[1], Stephen Fitchett[1], Carl Gutwin[2], Saul Greenberg[3]

[1] Dept. of Computer Science
University of Canterbury
Christchurch, New Zealand
**{jason, andy, saf75}
@cosc.canterbury.ac.nz**

[2] Dept. of Computer Science
University of Saskatchewan
Saskatoon, Canada
gutwin@cs.usask.ca

[3] Dept. of Computer Science
University of Calgary
Alberta, Canada
saul@cpsc.ucalgary.ca

ABSTRACT

In this paper, we show that people frequently return to previously-visited regions within their documents, and that scrollbars can be enhanced to ease this task. We analysed 120 days of activity logs from Microsoft Word and Adobe Reader. Our analysis shows that region revisitation is a common activity that can be supported with relatively short recency lists. This establishes an empirical foundation for the design of an enhanced scrollbar containing *scrollbar marks* that help people return to previously visited document regions. Two controlled experiments show that scrollbar marks decrease revisitation time, and that a large number of marks can be used effectively. We then design an enhanced *Footprints* scrollbar that supports revisitation with several features, including scrollbar marks and mark thumbnails. Two further experiments show that the Footprints scrollbar was frequently used and strongly preferred over traditional scrollbars.

Author Keywords.
Document revisitation, read wear, scrolling.

ACM Classification Keywords.
H.5. Information interfaces and presentation: User interfaces (*Interaction styles*, *screen design*).

INTRODUCTION

Scrollbars are a familiar widget common to graphical user interfaces, and have become a standard idiom for view navigation within documents. They are compact and conceptually simple, yet powerful, providing useful information about the viewport's location and extent in a larger information space. In particular, the scroll thumb's spatial cue can help *revisitation* – returning to previously visited document regions. For example, a user may know that "moving the thumb roughly four-fifths of the way down will bring me to the Results section".

However, rapid and effective revisitation depends on people

having attended to the thumb's location during previous visits, remembering it, and reproducing it accurately. One or more of these activities can fail. In their comparison of navigation in paper and electronic documents O'Hara and Sellen [25] observed that people use fingers to mark document regions in paper documents for easy return, and that equivalent methods are lacking when scrolling. They conclude that scrolling is "irritatingly slow and distracting" and that people need "quicker, more effortless navigation" [25], p.341.

One way to improve support for revisitation is by augmenting the scrollbar region with *scrollbar marks*. Such marks are not a new idea. *Attribute-mapped scrollbars*, patented in 1990 [31], used coloured marks in the scrollbar to draw attention to salient properties; scrollbars are well-suited to showing this information, as they provide an overview of the entire document. Hill et al. [18] used a similar approach to denote the *read wear* that occurs with use – the marks portray how often sections of the document have been read. However, their focus was on showing history of use, not on supporting revisitation.

Read-wear marks on the scrollbar offer a lightweight method for improving document revisitation compared to traditional scrolling: a mark shows a person where they have been, and provides a navigation cue to help them quickly return to that spot. However, the idea has not caught on, and we are unaware of any system currently using it (although scrollbar marks are becoming common in IDEs for marking code errors or comments).

We believe the poor adoption of read-wear scrollbars is due to a lack of knowledge about how revisitation occurs in the real world, how best to design a read-wear scrollbar, and the potential benefits and harms of using it in realistic systems. Consequently, we conducted several investigations to address these knowledge gaps.

1. *How do people revisit document locations, and are current tools well used?* We analysed logs from a 120-day longitudinal study of document navigation in Microsoft Word and Adobe Reader, and draw three main conclusions: that region revisitation is frequent, that current revisitation tools other than the normal scrollbar

are used very little, and that short recency lists can contain most locations that people will revisit.

2. *Can scrollbar marks improve revisitation, and how many marks should be used?* We conducted two experiments to test the performance of adding marks into the scrollbar. One study demonstrates that marks can decrease revisitation time; the other shows that people can successfully use a large number of marks.

3. *How can a realistic interface support revisitation?* We designed our Footprints Scrollbar (Figure 5) to maximize support for revisitation, based on principles derived from the log analysis and the two studies.

4. *Is this support used effectively and is it preferred?* We evaluated the Footprints Scrollbar in two experiments: a controlled lab study and a realistic usage study. The Footprints Scrollbar was frequently and successfully used, improved users' revisitation time, and was strongly preferred to traditional scrollbars.

Overall, our investigations show that revisitation is worth supporting, and that read wear – long known but little studied – is a valuable and usable technique for improving revisitation. To our knowledge these studies are the first to analyse real world revisitation in documents and to use the results in the design, implementation, and evaluation of a read-wear scrollbar.

RELATED WORK

Revisitation and its support

Zipf's Law [33], the Pareto Principle [22] and the "80-20 Rule" all demonstrate that many facets of human behaviour are highly repetitive: although we have extensive vocabularies, wardrobes, or choices at our disposal, most of the time we reuse a small set of favoured options. Researchers have also demonstrated that many aspects of computer use are also highly repetitive, such as command use [15, 16], menu selections [11, 14], and web page visits [2, 12, 29].

These observations have stimulated research into interface techniques for supporting reuse and revisitation. Some strategies require manual intervention while others automatically observe actions and update the interface.

Manual revisitation tools require an explicit user action to mark information as interesting. For example, bookmark tools in web browsers and Microsoft Word let people explicitly create iconic labels as shortcuts for returning to particular pages or positions in a document. The *Bookmark Scrollbar* [24] is similar, but bookmarks are placed within a standard scrollbar. However, bookmarks have problems that limit their use in practice [1]. First, they depend on people knowing in advance that the information will be required in the future. Second, people must believe that adding a bookmark will yield benefits that exceed the manipulation costs of creating and managing them. Third, people's propensity for adopting satisficing strategies [26] (where

immediate suboptimal strategies are favoured over more efficient long-term ones) suggests that people will often fail to place bookmarks, even when they can foresee the long term advantage [1]. This is why our research interest lies primarily in automatic revisitation support.

Automatic revisitation tools have been developed to support revisitation both between and within documents. Familiar interface controls such as history lists, 'Back/Forward' buttons, and 'Recent Document' menus facilitate navigation *between* documents. These automatic revisitation tools minimally intrude on users' activities: they silently record actions, populate a data structure or visualisation, and provide support when called on. Their primary disadvantages are that people may not understand the algorithm for recording or presenting events, and the event set may overwhelm or fail to match the interests of the user. For example, people often misunderstand the behaviour of the web 'Back' button, causing frustration when items cannot be revisited [12], and we have observed similar problems with missing items in 'Recent Documents' menus, and with user misunderstanding of the temporal order of Alt-Tab window switching.

There are also several widely-deployed examples of automatic revisitation tools for navigating *within* documents. For example, the web browser's 'Back' button works as normal when navigating through internal page links, and Adobe Reader's 'Previous/Next View' feature steps through a linear history of scroll positions and zoom levels. Visual Studio also has a 'previous/next' list of lines that the I-beam cursor has visited. These history lists leave no visible trace in the scrollbar, however, so people cannot visually scan potential target regions without displaying additional windows or menus.

Hill *et al*'s [18] original *Read Wear* system showed a histogram overview of the reading history of an entire document within the scrollbar. Each horizontal line of pixels in the scrollbar encoded information such as the number of edits, or length of time reading. Similar scrollbar marks are used by several code editors, but the marks are used to highlight semantic information such as compilation errors rather than to support revisitation.

The concept of document read wear inspired several researchers to examine a variety of techniques for recording and visualizing activity beyond the scrollbar. These include Wexelblat and Maes' [30] 'Footprints' system, which provided maps, trails, annotations and signposts for information foraging, and Skopik and Gutwin's [27] 'visit wear' marks for revisitation in fisheye visualizations.

To our knowledge, the idea of combining visitation histories with overview visualizations in the scrollbar has not been pursued since Hill *et al's* initial investigation.

Scrolling Enhancements

Many researchers have investigated ways to improve scrolling. Four general approaches have been: to improve the control device (e.g. isometric joystick [32] or scroll-wheel [19]); to improve the control-feedback loop (e.g. speed-dependent automatic zooming [13, 21] or OrthoZoom [5], both of which use zooming to speed long distance movement); to improve spatial memory of document regions (e.g. Space-Filling Thumbnails [10] or Stationary Scrolling [28]); and to augment the scrollbar with meaningful information (e.g. Edit/Read Wear [9, 18], bookmarks [24], or the location of search terms [6-8]).

Our research bridges the third and fourth approaches, using an augmented scrollbar to reinforce spatial memory and aid revisitation. First, however, we report the findings of our log analysis of behaviour in document revisitation.

LOG ANALYSIS OF REVISITATION

How do people currently revisit document locations? Knowing actual revisitation patterns is critical for the design of a revisitation system. To answer this question, we analysed logs of all navigational actions in Microsoft Word 2003 and Adobe Reader 7 as used by 14 participants (all computer scientists) during their normal work over 120 days. The logging software and initial data analysis are reported in [3, 4]; the analysis of revisitation discussed below is new. The logs precisely record the location and extent of the scroll thumb every 200ms, describing the user's viewport independent of zoom level. Note that in large documents (where the scroll thumb is the minimum allowable size) the logged viewport exceeds the actual one. Table 1 shows summary statistics of the documents used during the study.

	Microsoft Word	Adobe Reader
1. Total documents	2342	1706
2. Number of documents	167±122, 37-384	122±183, 5-718
3. Document page length	6.3±4.5, 1-160,	38±35.7, 1-1743
4. Number of days used	41.6±21.6,16-81	25.6±15.1, 4-53
5. Mean docs opened/day	3.7±1.6, 1.8-7.3	4.1±5.6, 1.3-23.1
6. No navigation in document	928 (39.6%)	334 (19.6%)
7. Document scrolled	1414 (60.4%)	1372 (80.4%)
8. No revisits	969 (41.4%)	998 (58.4%)
9. >1 revisit (% of scrolled docs)	445 (31.5%)	374 (27.3%)
10. Mean max. daily revisits	188.8±114.0	32.3±28.3

Table 1: Document usage and revisitation with Microsoft Word and Adobe Reader. Rows 2-5 show mean, standard deviation (as ±), and the min-max range across users.

Defining Visits and Revisits

When navigating within documents, people frequently stop over a region without any intention to do so, caused by mechanical demands such as clutching the mouse or scroll-wheel, or by cognitive and perceptual issues such as determining whether the document is correctly positioned. We therefore need to define a specific pause length as denoting an *intentional visit*, and to prune visits of shorter duration. The analyses that follow were all repeated with one, two and three second stationary pauses used to define a

'visit', with little difference to the results. We report results only for two-second pauses.

When revisiting document locations, people are unlikely to arrive at precisely the same viewport as the one previously seen. We therefore define a revisit as returning to anywhere within the bounds of an earlier visit and remaining there for more than two seconds. Furthermore, revisits are only logged when the person visits another position outside the current view prior to returning. This condition ensures that actions such as slowly advancing line by line are not logged as revisits.

Patterns of Revisitation

A large proportion of documents were opened and closed without any navigation: 39.6% and 19.6% for Word and Reader. This is probably due to people launching the application without an associated document (e.g., via the Start Menu), or due to opening the wrong document and immediately closing it. Removing these navigation-free instances from the analysis shows that 32% of Word and 27% of Reader documents include at least one revisit (row 9, Table 1). Row 10 shows the mean of participants' maximum daily revisits: 188.8 (Word) and 32.3 (Reader), indicating that revisitation is a frequent activity in some documents.

Use of existing revisitation tools

None of the participants used Word's bookmarking function or Reader's 'Next View' function during the 120 days. A single participant used Reader's 'Previous View' function in two separate documents (0.3% of his documents). During interviews following the log analysis, all but two of the participants stated they were not aware of the Previous/Next view features.

Word's 'Split window' feature is another tool that can support revisitation; it allows users to simultaneously view more than one document region, eliminating the need for repetitive scrolling between two regions. Seven of the fourteen participants used split windows at least once, with a total of 42 invocations. One participant accounted for more than half (24) of these events.

*Recency and frequency list coverage by length*

Hill *et al.*'s [18] read wear system displayed the amount of time spent on every line, which is useful for returning to frequently used locations, but not for recent locations. This raises the question of whether recency or frequency is a more effective basis for revisitation support.

One key question for either design is that of how many items should be displayed in the scrollbar: e.g., the two most recently/frequently visited, or top twenty? Increasing the number of marked items increases the number of positions available for revisitation, but it also increases the user's search load when scanning for a target. We therefore

analysed the logs (using methodology from previous studies of revisitation [15,29]) to determine the proportion of revisited items covered by a recency or frequency list of length n.

Figure 1a shows the results for recency list analysis. For example, a one-item recency list (e.g., a simple 'Back' button) would allow users to reach 19% and 28% of previous locations with Word and Reader. Longer lists rapidly increase the proportion of accessible locations, with ten-item recency lists covering 81% of revisitations with Word and 84% with Reader.

Frequency list analysis (Figure 1b) shows a similarly steep slope, with 10 items allowing access to slightly fewer items than the recency list, at 78% and 83% of revisited locations for Word and Reader respectively. The values for a one-item frequency list are also lower than recency lists at 16% and 13% respectively, meaning that fewer regions would be accessible with a simple 'Back' button than with recency lists.
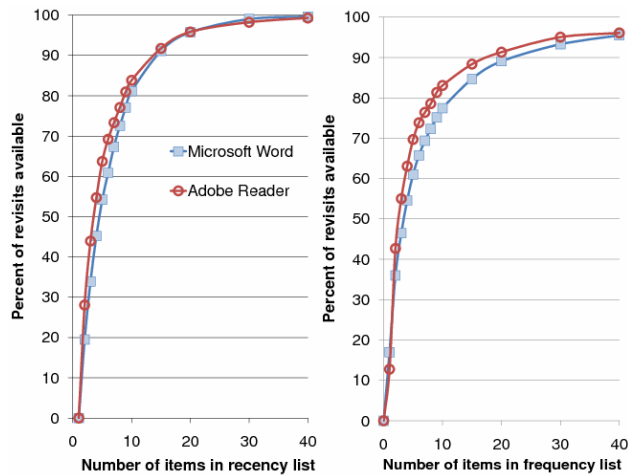
(a) Recency list.   (b) Frequency list.
Figure 1: Percent of revisited items available through theoretical recency and frequency lists of different sizes.

Summary of Log Analysis
The log analysis showed: that people frequently revisit document locations; that current revisitation tools in Word and Reader are seldom used; that relatively short revisitation lists provide good coverage of the locations that users revisit; and that recency-based lists provide better coverage than frequency based lists. This last point is intuitively reasonable – users' interests in document regions change over time; for instance, the morning's work at regions x and y may be unimportant during the afternoon's work at regions w and z.

EXPERIMENT ONE: CAN RECENCY MARKS HELP?
The log analysis suggests that a recency-based revisitation mechanism is worth pursuing. We designed a simple marking scrollbar that displayed recency-based scroll marks on it (Figure 2). We then conducted an initial experiment to determine whether people would use the recency marks (or ignore them), and whether using them would aid their revisitation performance.

Tasks involved searching for locations in a plain text document (Joyce's *Ulysses*, obtained from gutenberg.org) and periodically revisiting them. The experiment also investigated how performance was influenced by the number of marks in the scrollbar and by the number of revisits to the same location (the latter providing insight into how well users learn locations). Users completed the tasks using a simple document-viewing application that allowed navigation only with the scrollbar. Two versions were created, whose interfaces differed only in scrollbar type: a *standard scrollbar,* or a *marking scrollbar* that showed red marks for visited locations. Middle-clicking on a mark immediately scrolled the view to its associated location. In this experiment, marks were only placed when targets were successfully visited or revisited; pausing elsewhere in the document did not place a mark. Consequently, this is a best-case study for marking scrollbars; it considers questions about whether people will use the marks, whether the marks improve performance, whether the marks distract from other tasks, and how well users learn document locations both with and without marks.
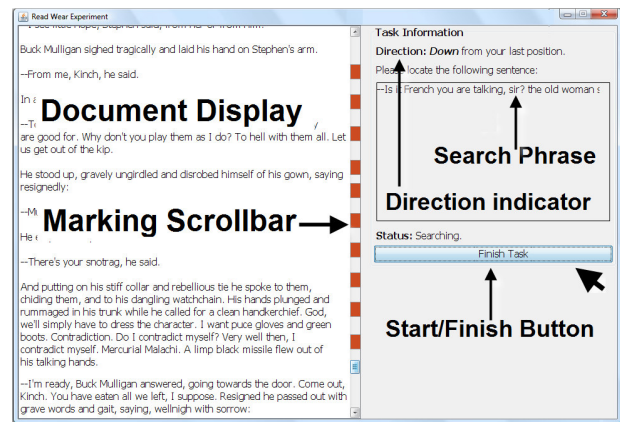

Figure 2. Experiment one interface. Marking scrollbar (left) and task cueing interface (right). Image altered for clarity.

Participants and Apparatus
Twelve volunteer university students (seven female) participated in the experiment. Their mean age was 24 years (s.d. 5 years), and they were all experienced users of window-based software (>10 hours/wk). The experiment lasted approximately 30 minutes.

A Java-based system was built for the study. A single window displayed both the experimental interface (1485×800 pixels, Figure 2 left side) and a task cueing pane (Figure 2, right). The task-cuing pane displayed an initial direction to the next target (up or down), the text of the sentence to be located, and a start/finish task button. The

experiment ran on a standard Windows PC with a 22" LCD monitor.

Experimental Design and Procedure

The task-time dependent measure is analysed using a 2×2×4 repeated measures analysis of variance (ANOVA), with the three factors: *interface type* (standard and marking scrollbars), *positions visited* (5, 10) and *revisit iteration* (1st, 2nd, 3rd, or 4th revisit).

Participants were given a brief introduction to each interface and the experimental method before completing three practice tasks with their first interface. They then completed all tasks with one interface before proceeding to the next, with the order counterbalanced. Participants completed NASA Task Load Index (TLX) worksheets and provided comments after completing each interface.

Tasks involved using the experimental interface to locate sentences displayed one at a time in the cueing interface. Participants were informed that all target sentences were the first in a paragraph and that tasks were completed by placing the target sentence anywhere within the viewport and clicking the "Finish Task" button. Successfully completing one task automatically cued the next.

Two sets of ten tasks were generated: the first set within the first twenty pages of the document, and the second set within the second twenty pages. All participants used the first set with their first interface. The targets were generated to have similar locations across sets (within one paragraph of each other). Targets were spaced nearly evenly through the 20 page regions, with at least one window of text between them. Equivalent targets in both sets were revisited the same number of times. Participants were unaware of these constraints on target placement.

The set of targets were visited as follows. First, participants found five consecutive targets for the first time, progressing downwards through the document: t_1, t_2, t_3, t_4, t_5. Next, they entered the first of two revisitation phases, revisiting items in the order t_2, t_4, t_2, t_1, t_2, t_4, t_2. They then visited the five remaining targets for the first time (t_6–t_{10}), resulting in ten marks in the scrollbar, followed by a second revisitation phrase with target order t_3, t_9, t_3, t_6, t_3, t_9, t_3. In total, each participant completed 48 experimental tasks: 2 targets visited 5 times each (t_2 and t_3), two visited three times (t_4 and t_9), two visited twice (t_1 and t_6), and four visited once (t_5, t_7, t_8 and t_{10}), giving 24 tasks. These were then repeated for the second interface.

Tasks were automatically completed after 90 seconds to reduce the impact of situations where participants became lost. Six tasks were discarded due to exceeding the time limit: two with marks, and four with traditional scrollbars.

Results

Figure 3 summarises these results. The marking scrollbar allowed significantly faster revisitation (mean 11.3s, s.d. 11.0s) than traditional scrollbars (20.4s, s.d. 14.8s): $F_{1,11}=33.8$, p<.001. As anticipated, there were also significant main effects for *revisit iteration* ($F_{3,33}=10.8$, p<.001), as indicated in Figure 3, and for *positions visited* ($F_{1,11}=21.2$, p<.005), with mean acquisition times increasing from 13.1s with five visited locations to 18.6s with ten. There was a significant *interface type* × *positions visited* interaction ($F_{1,11}=5.3$, p<.05), with the marking scrollbar showing greater benefits with more positions visited. There was no interaction between *interface type* and *revisit iteration*: $F_{3,33}<1$.
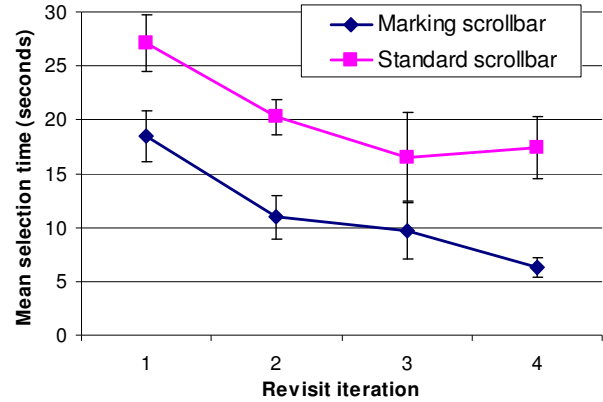

Figure 3. Mean acquisition times in experiment one for the two scrollbar types across revisit iteration.

Subjective measures and participant comments supported these positive results for marking scrollbars. NASA-TLX worksheet results (ratings from 1=low to 5=high) showed lower mean workload and higher mean performance ratings for the marking scrollbar in all categories, although only the overall 'Effort' measure showed a significant effect (see Table 2).

NASA-TLX	Standard	Marking	Sig?
Mental load	2.9 (1.1)	2.4 (0.9)	0.08
Physical load	2.1 (1.4)	1.7 (0.7)	0.14
Temporal load	2.4 (1.0)	2.3 (1.1)	0.32
Performance	3.3 (1.2)	3.7 (0.9)	0.25
Frustration	2.5 (1.2)	2.2 (1.1)	0.31
Effort	2.9 (0.9)	2.0 (0.9)	**0.02**

Table 2: Mean (st. dev.) NASA-TLX responses for standard and marking scrollbars in experiments one. Wilcoxon Signed Ranks significance test.

Participant comments were positive: "The scrollbar mark is extremely nice". One participant noted that the marks helped reduce their spatial search space: "I would usually only be off by one mark if I didn't choose correctly the first time, so it was easy to correct my mistakes."

Discussion

These results show that accurately placed scrollbar marks helped participants return to document locations, and that the support was appreciated. Also, as more places were visited, participants gained more from the marks.

There are three questions that are yet to be addressed in the investigation: first, how is performance influenced by large numbers of visited locations (and therefore marks); second, what is the performance of the marking scrollbar when marks are automatically placed wherever the user pauses, rather than the idealized placement on targets tested here; and third, how does the scrollbar work with naturalistic revisitation rather than artificially-organized targets? Experiment two addresses the first of these issues: how is performance influenced by large numbers of visited locations.

EXPERIMENT TWO: MARKING OVERLOAD
The log data showed that 10 marks covers 80% of revisitations, and that 30 marks covers close to 100%. Yet as Figure 1 reveals, each additional mark covers a smaller proportion of revisitation targets, therefore offering progressively lower utility while increasing the number of distracters. This study, therefore, investigates how performance is influenced by the number of marks.

Experimental Method
Apparatus and participants. The experiment used the same marking interface, cuing interface, and apparatus as experiment one. In addition, the same people from the first experiment participated in the second, and advanced to the second experiment after a short break.

Design. Revisitation time is analysed using a 4×4 repeated measures ANOVA with two factors: *positions visited* (5, 10, 20, 30) and *revisit iteration* (1st, 2nd, 3rd, 4th).

Tasks. Participants completed three preparation tasks and then searched and revisited locations within the first 45 pages of James Joyce's *Dubliners* using a similar procedure to that used in experiment one (marks were automatically placed on targets when acquired).

Results and Discussion
Figure 4 summarises the results. As anticipated, there are significant main effects for both *number of positions visited* ($F_{3,33}$=13.4, p<0.001) and *revisit iteration* ($F_{3,33}$=22.4, p<0.001). There is also a significant interaction between the factors ($F_{9,99}$=1.99, p<0.05), which is best explained by the large absolute task time reduction between iterations 1 and 4 with 30 marks compared to the smaller reductions with fewer marks. This should be expected, since learning and remembering thirty marks is clearly more demanding than five.

By the fourth iteration, performance appears to be reaching an asymptote for all conditions, with little performance difference between 10, 20 and 30 marks (Figure 4, far right). This suggests that participants were not overloaded by 30 marks.

Additional evidence suggesting that users were successfully "deciding" about marks rather than linearly searching them

is provided by regression analysis of *acquisition time* with *positions visited*, which gives a logarithmic best fit with R^2=0.98. The Hick-Hyman Law [17, 20] shows that logarithmic functions relate decision time to number of items, while linear functions apply to serial candidate searches.
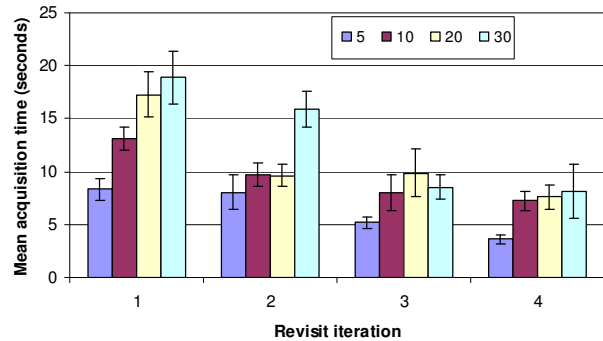

Figure 4. Mean revisit times with 5, 10, 20 and 30 scrollbar marks on the 1st to 4th revisit in experiment two.

These results suggest a tradeoff between the number of marks and performance. Although the participants quickly learned many marks and their associated regions (performing well by the 4th revisit), acquisition times on their first revisit increased steeply with the number of marks (Figure 4, far left). Nevertheless, these results show that large mark sets are feasible, and that designing to cover nearly all revisits is a possibility. Since ten marks cover more than 80% of revisited locations, however, (see Figure 1), we use 10 marks in our remaining studies.

THE FOOTPRINTS SCROLLBAR
The results of the first two studies, and our experiences with the initial version of the marking scrollbar, led us to design a new version – called the *Footprints Scrollbar* (Figure 5) – that contains a number of more advanced features for supporting revisitation.

The Footprints scrollbar supports six related methods for revisiting regions. First, coloured marks are placed in the scrollbar to provide spatial cues to previously visited areas. Marks gradually fade from 'hot' colours (reds, oranges) through to 'cold' one (greens, blues) to denote their increasing age in the recency set. Second, middle clicking on a mark causes a rapid, animated scroll transition to the associated view [23]. Third, when the user moves the cursor over the scrollbar, small thumbnail images quickly fade into view alongside each mark, giving a visual overview of the associated document regions. Moving the cursor over a thumbnail expands that thumbnail for better visual inspection (an example is shown at the bottom of Figure 5). The thumbnails fade out when the cursor moves away from the scrollbar. Clicking a thumbnail also moves to that view. Fourth, back/forward functions are invoked by the 'left' and 'right' keyboard keys; this allows users to rapidly move through the mark history and its corresponding region views. Fifth, depressing the Shift key and rotating the
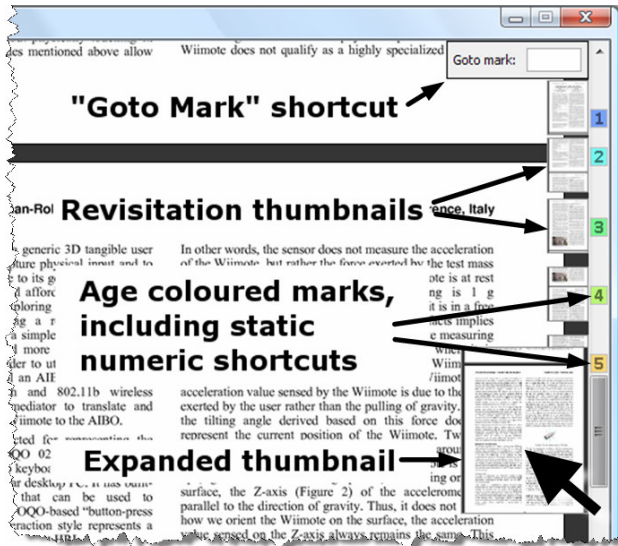
Figure 5: The Footprints scrollbar.

scrollwheel moves the document position to the closest mark in the direction of rotation (according to distance, not recency). Finally, marks are overlaid with the numbers 1–10: users can trigger rapid movement to the selected position through the numeric keypad or by typing a number into a 'go-to' box. Number assignment is arbitrary, with the shortcut remaining constant for the life of the mark. Details of these features follow.

Marking Algorithms and Behaviour

Each mark is 16 pixels high (equal to the default minimum size for a scroll thumb in most widget sets), and marks never overlap one another. Marks are only placed in the scrollbar when the scrolled region remains static for more than two seconds. Consequently, both continuous scrolling and scrolling with short pauses for device clutching or visual inspection have no impact on the marks and the recency list.

Scrollbar marks are produced from a recency list data structure that removes duplicates [15, 29]. Whenever the scroll view remains static for more than two seconds, that region is inserted at the end of the recency list, and any earlier entry for the same location is removed from the list. The current implementation has a 1:1 correspondence between marks and list entries, but this could be relaxed to allow more items on the recency list than are visualized in the scrollbar: e.g., providing 10 visible marks, but allowing a larger number of previously visited regions, accessible via the Back/Forward keys.

Marks are numbered from 1 to n, where n is the configurable maximum (with a default of 10). To reduce visual distraction associated with scrollbar changes, the numbers are as stable as possible, with each mark after n receiving the mark number of the least-recent member of the recency list. Clicking the mark, typing its shortcut number, or dragging the thumb to the mark are all equivalent methods of revisitation.

The two-second timeout has an important impact on the behaviour of the interface: it determines when marks are placed, and determines the semantics of the back/forward keys. To help users predict and comprehend this behaviour, the scroll-thumb depicts the approaching timeout by gradually filling with colour, similar to a progress bar (see right). Once filled, the timeout expires, a mark is left in the scrollbar, the region is inserted at the tail of the recency list, and the window view is captured for use in the associated thumbnail.

To precisely describe this behaviour, Table 3 uses comma separated letters $a, b, c...$ to represent visited document regions; the symbol $\perp$ denotes a 2 second static location; subscripts $1, 2, 3...$ denote shortcut digits on marks; $\Leftarrow$ and $\Rightarrow$ denote the back and forward keys; and curly braces $\{...\}$ enclose the comma-separated content of the recency list, with the most recent item at the right hand end of the list. An underlined item in the recency list represents the region displayed at the end of each action sequence.

Actions	Region sequence	Recency list and marks
1. Visit three regions	$a\perp, \ b\perp, c\perp$	$\{a_1, b_2, c_3 \}$
2. Scan 3 regions then pause	$d, e, f\perp$	$\{a_1, b_2, c_3, f_4\}$
3. 'Back' three times (no pauses)	$\Leftarrow, \ \Leftarrow, \ \Leftarrow$	$\{\underline{a_1}, b_2, c_3, f_4\}$
4. 'Forward twice (still no pauses)	$\Rightarrow, \Rightarrow$	$\{a_1, b_2, \underline{c_3}, f_4\}$
5. Pause	$\perp$	$\{a_1, b_2, f_4, \underline{c_3}\}$
6. 'Back' and pause	$\Leftarrow_\perp$	$\{a_1, b_2, c_3, \underline{f_4}\}$
7. 'Back' and pause	$\Leftarrow_\perp$	$\{ a_1, b_2, f_4, \underline{c_3}\}$

Table 3: Navigational behaviour of the Footprints Scrollbar.

The timeout allows users to dynamically traverse the recency list with the 'Back' and 'Forward' shortcut keys: for example, quickly flipping between regions (e.g. rows 3 and 4 of Table 3). However, when editing two document regions (e.g. 'Results' and 'Abstract') users are likely to spend relatively long periods in both areas, and they can move between them with subsequent presses of 'Back' (e.g. rows 6 and 7). Experiment four investigates the usability of this model, and how users chose to interact with the system; first, we report on an experiment that tested the performance of the new system.

EXPERIMENT THREE: FOOTPRINTS EVALUATION

Log analysis showed that people *could* benefit from revisitation support. The experimental results above suggest that people *can* benefit when using even a simple marking scrollbar, if marks are correctly placed. Correct mark placement is an artificial ideal, however, and realistic mark placement needs to be evaluated.

This experiment compared the Footprints scrollbar with a standard scrollbar in terms of people's performance and preferences. Tasks involved finding and revisiting document regions, with the Footprints system automatically placing marks whenever the user paused for more than two

seconds. Unlike experiment one (which controlled the number of revisits to each position), this experiment controlled revisits according to mark position within the recency list. This allowed us to inspect how performance of the Footprints scrollbar was affected by the differing ages of different marks. The disadvantage of this method is that it artificially made less-recent revisits more likely than our logs show them to be.

Experimental Method
Twelve participants (two female) took part in the experiment. Tasks involved finding and re-finding document regions that were cued by displaying an image of the target region and an initial direction (up or down). Participants began tasks by pressing a "Start Task" button, and completed them by scrolling the target region into the middle two-thirds of the screen and clicking a "Finish Task" button. A red status message was displayed if the target was not correctly positioned, and the task continued. Training with each interface was similar to experiment one, and was followed with eight sample tasks to familiarize participants with the procedure.

All participants used the Footprints scrollbar and the standard scrollbar (in counterbalanced order) with ten- and forty-page documents (two papers from the CHI proceedings and two instruction manuals). Eleven target regions were generated for each document by evenly distributing preliminary locations, then randomly adjusting these locations by between 0 and 5%.

The eleven targets were initially presented consecutively from the top of the document to the bottom. Participants then revisited targets according to their *ideal* position on the recency list: three times for each position 2–11. The ideal positions assume that marks *only* fall on targets. However, marks were placed whenever the user paused, so the ideal recency list is likely to differ from the user's actual one. Therefore, latter list positions may not have been visible in the scrollbar when needed, due to being displaced by other marks. Analysis of the study logs showed that the location in 8th position on the ideal list was in the user's visible list 90% of the time, the 9th 72%, the 10th 52%, and the 11th 15%.

Results
Task time data are analysed using a 2×2×10 RM-ANOVA for factors *interface* (Footprint, standard), *document length* (10 and 40 pages) and *recency list position* (2–11).

There was a significant main effect for *interface type* ($F_{1,11}$=8.24, p<0.05), with Footprints (7.4s) faster than standard scrollbars (8.8s). *Document length* ($F_{1,11}$=69.6, p<0.001) and *recency list position* ($F_{1,11}$=18.8, p<0.001) both showed expected main effects. Importantly, there were significant *type×length* ($F_{1,11}$=15.1, p<0.01) and *type×position* ($F_{10,110}$=2.9, p<0.01) interactions. Figure 6

shows the *type×length* interaction – both interfaces performed similarly with 10 page documents (means of 6.9s vs. 6.6s), but Footprints performed better with 40 page documents (mean 7.9s vs. 10.9s). The *type×position* interaction is due to the Footprints scrollbar out-performing standard scrollbars in all *but* the 11th position on the recency list. On trials in the 11th list position, the Footprints system provided a corresponding mark only 15% of the time (due to marks falling off the end of the recency list), so users who tried to use marks would have been misled. Means on the 11th trial were 12.0s and 10.5s for Footprints and standard scrollbars.
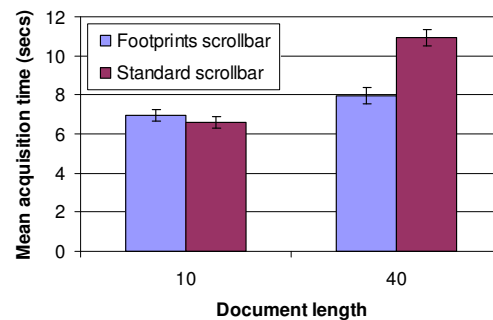

Figure 6: Experiment three results, by document length.

All but one participant preferred the Footprints scrollbar; the remaining participant could not choose between them. Mean NASA-TLX worksheet results, were uniformly better for Footprints: significantly so for 'Physical Load', 'Performance' and 'Effort'. Participant comments from experiment three are discussed below, together with those of experiment four.

EXPERIMENT FOUR: OBSERVATIONAL STUDY
Tasks in experiment three were tightly constrained for experimental control, but they artificially induced revisitation and artificially exaggerated temporally distant revisitations. To inspect more realistic tasks as well, our final study used a structured interview process to observe participants' more natural interaction within a document of their choosing.

Eight participants (one female), all graduate students or faculty from Computer Science, took part in the experiment. They were asked to identify a favourite research paper, which we loaded into the system.

The experiment began with a two-minute introduction to the system's features, explaining marks, thumbnails, digit shortcut keys, and the back/forward arrow keys. Participants then completed 25 tasks using whichever methods they chose. Tasks involved describing the paper to the experimenter in response to a consistent set of questions: "show me what you think is the best part of the paper", "who is on the reference list?", "where is related work summarised?", "what's the first paper referenced in the Introduction", "where was that paper published", etc.

Participants then completed a questionnaire on the effectiveness of various system components. Automatic logs captured all user actions with the interface.

Results

Questionnaire responses (summarized in Table 4) show that participants found marks and thumbnails very helpful for visualizing and navigating to previously visited locations, with Likert scale (1=disagree; 5=agree) means ranging between 4.5 and 4.9 (rows 1–3). These positive ratings were achieved without substantial distraction (mean distraction ratings of 2.3 for marks and 1.5 for shortcut numbers). Overall, comments were highly positive, including "revisitation is tremendously useful and would probably only improve as the document increases in size", "really useful", and "it's additive: no interference with any other widget". Five of the participants stated that they would want all of the supported features in their desktop interfaces; three stated they would want some of the features, with all wanting the thumbnails, region markers, and middle-click shortcuts.

Question	Mean (s.d.)
1. Marks were helpful to see where I'd been.	4.8 (0.5)
2. Marks were useful for revisiting locations	4.9 (0.4)
3. Thumbnails were useful for revisiting locations	4.5 (1.1)
4. I understood mark placement	4.9 (0.4)
5. Marks were distracting	2.3 (1.2)
6. Colours were useful	2.1 (1.0)
7. Back/Forward keys were useful	2.8 (1.3)
8. I understood the Back/Forward keys	4.5 (1.1)
9. Shortcut numbers were useful	2.5 (1.3)
10. Shortcut numbers were distracting	1.5 (0.8)

Table 4: Mean responses to five point Likert scale questions in experiment four.

Discussion of Experiments Three and Four

Experiments three and four show the value of the Footprints scrollbar for revisitation in both a controlled, and a more naturalistic situation, especially for longer documents. This success validates our empirical recommendation that revisitation should be supported (log activity analysis), and our initial evaluations of scroll marks in artificial situations. People also preferred the features of the Footprints scrollbar.

Despite these successes, the Footprints scrollbar is still an early design. Although the overall system was praised, participants identified several areas for improvement, all of which can be incorporated in the next design iteration.

Mark colours. Nearly all participants commented that colour poorly communicated mark recency. Worse, three participants observed that their memory for items was harmed by colour changes – they might remember the 'red' mark in a region, only to be confused by later colour reconfiguration. Furthermore, two participants noted that coloured marks increased the difficulty of visually acquiring the scroll thumb. We therefore recommend that future implementations use stable mark colours (not denoting mark age) with smaller marks and higher levels of opacity to ensure they do not interfere with thumb acquisition.

Digit shortcuts. Only one of the participants used digit shortcut keys for navigation, and another one commented that he used the digit marks 'to map locations'. Generally, though, the digit marks were criticized by a few participants as either unnecessary or mildly confusing.

Back/forward key shortcuts. Most participants stated that they understood the behaviour of this mechanism (mean 4.5) and that they might use the keys in other documents, but *none* of the participants actually used them during the tasks. One mentioned a conceptual clash between the 'Forward' key and 'forward=down' in the document, despite understanding the behaviour. Another stated "using back/forward arrows is something that just didn't occur to me." This comment echoes the findings of the log analysis – current recency tools, such as Reader's 'previous/next view' or Word's bookmarks are not sufficiently ready-to-hand, and hence they go unused despite their potential utility.

Lack of control over mark placement. Some participants felt pressured by the thumb-filling animation, with one commenting that it "made me rush before it dropped a mark" and another stating that "it would be nice to somehow stop the dropping of the marks". The questionnaires asked participants to comment on whether the two second marking timeout was too short or too long, with several responding that it was "sometimes too short and sometimes too long". Lightweight controls for manually adding and removing marks in the scrollbar could solve this problem, but it is unclear whether people would use such controls.

CONCLUSIONS

Revisitation has been comprehensively investigated in domains such as web navigation and command use. Somewhat surprisingly, however, region revisitation within documents has been largely overlooked. Our log analysis demonstrates that users frequently revisit document regions and that short revisitation lists can theoretically provide access to most locations that users return to. We used these findings to motivate and inform the design of a system, based on Hill *et al.'s* read wear, that augments the scrollbar with marks that aid revisitation; a series of evaluations demonstrate that the system can improve user performance.

Our Footprints scrollbar works within the current ecology of graphical user interfaces – it augments the familiar scrollbar rather than replacing it, and it occupies the same location on the screen. Except for shortcut keys (which can be easily modified) the input to this scrollbar does not compete with input actions that control other parts of a document viewer or GUI. While improvements were

suggested through experimental feedback, these can be easily incorporated in future system refinements.

Our future work with the Footprints scrollbar will progress in two main directions. First, we plan to revise the design of the scrollbar according to the user responses discussed above, and produce a version that can be incorporated into real-world document readers. Second, we will carry out longer-term evaluations of the device, to provide additional information about how revisitation and revisitation support work in real use.

REFERENCES

1. Abrams, D., Baecker, R. and Chignell, M. Information Archiving with Bookmarks: Personal Web Space Construction and Organization. in *Proc. CHI'98*, (1998), 41-48.
2. Adar, E., Teevan, J. and Dumais, S. Large Scale Analysis of Web Revisitation Patterns. in *Proc. CHI'08.* (2008), 1197-1206.
3. Alexander, J. and Cockburn, A. An empirical characterisation of electronic document navigation. in *Proc. Graphics Interface* (2008), 123-130.
4. Alexander, J., Cockburn, A. and Lobb, R. AppMonitor: A Tool for Recording User Actions in Unmodified Windows Applications. *Behavior Research Methods 40*, 2 (2008), 413-421.
5. Appert, C. and Fekete, J. OrthoZoom Scroller: 1D Mutli-scale navigation. *Proc. CHI'06.* (2006), 21-30.
6. Beaudouin-Lafon, M. Instrumental Interaction: An Interaction Model for Designing Post-WIMP User Interfaces. in *Proc. CHI'00.* (2000), 446--453.
7. Björk, S. The ScrollSearcher Technique: Using Scrollbars to Explore Search Results. in *INTERACT'01*, (2001), 745-746.
8. Byrd, D. A Scrollbar-based Visualization for Document Navigation. in *Proc. Digital Libraries*, (1999), 122-129.
9. Chimera, R. ValueBars: An Information Visualization and Navigation Tool. *Proc. CHI'92.* (1992), 293-294.
10. Cockburn, A., Gutwin, C. and Alexander, J. Faster Document Navigation with Space-Filling Thumbnails. in *Proc. CHI'06.* (2006), 1-10.
11. Cockburn, A., Gutwin, C. and Greenberg, S. A Predictive Model of Menu Performance. in *Proc. CHI'07.* (2007), 627-636.
12. Cockburn, A. and McKenzie, B. What Do Web Users Do? An Empirical Analysis of Web Use. *Int. Journal of Human-Computer Studies 54*, 6 (2001), 903--922.
13. Cockburn, A., Savage, J. and Wallace, A. Tuning and Testing Scrolling Interfaces that Automatically Zoom. in *Proc. CHI'05.* (2005), 71-80.
14. Findlater, L. and McGrenere, J. A comparison of static, adaptive, and adaptable menus. in *Proc. CHI'04* (2004), 89-96.
15. Greenberg, S. and Witten, I. Supporting Command Reuse: Empirical Foundations and Principles. *Int. Journal of Man-Machine Studies 39* (1993), 353--390.
16. Greenberg, S. and Witten, I. Supporting Command Reuse: Mechanisms for Reuse. *Int. Journal of Man-Machine Studies 39* (1993), 391--425.
17. Hick, W. On the rate of gain of information. *Quarterly J. Experimental Psychology 4* (1952), 11-36.
18. Hill, W., Hollan, J., Wroblewski, D. & McCandless, T. Edit Wear and Read Wear. *Proc. CHI'92.* (1992) 3-9.
19. Hinckley, K., Cutrell, E., Bathiche, S. and Muss, T. Quantitative Analysis of Scrolling Techniques. in *Proc. CHI'02.* (2002), 65-72.
20. Hyman, R. Stimulus information as a determinant of reaction time. *J. of Exp. Psych. 45* (1953), 188-196.
21. Igarashi, T. and Hinckley, K. Speed-dependent Automatic Zooming for Browsing Large Documents. in *Proc. UIST'00.* (2000), 139-148.
22. Juran, J. *Quality Control Handbook*. McGraw-Hill, New York, 1951.
23. Klein, C. and Bederson, B. Benefits of Animated Scrolling. *Ext. Abstracts CHI'05.* (2005), 1965-1968.
24. Laakso, S.A., Laakso, K.-P. and Saura, A.J. Improved Scroll Bars. *Ext. Abstracts CHI'2000*, (2000), 97–98.
25. O'Hara, K. and Sellen, A. A comparison of reading paper and on-line documents. in *Proc. CHI'97.* (1997), 335-342.
26. Simon, H. Satisficing. in Eatwell, J., Millgate, M. and Newman, P. eds. *The New Palgrave: A Dictionary of Economics*, Stockton Press, New York, 1987, 243-245.
27. Skopik, A. and Gutwin, C. Improving Revisitation in Fisheye Views with Visit Wear. in *Proc. CHI'05.* (2005), 771-780.
28. Smith, R. and Taivalsaari, A. Generalized and stationary scrolling. in *Proc. UIST'99.* (1999), 1-9.
29. Tauscher, L. and Greenberg, S. How People Revisit Web Pages: Empirical Findings and Implications for the Design of History Systems. *Int. Journal of Human Computer Studies, 47*, 1 (1997), 97-138.
30. Wexelblat, A. and Maes, P. Footprints: History-Rich Tools for Information Foraging. in *Proc. CHI'99.* (1999), 270--277.
31. Wroblewski, D., Hill, W. and McCandless, T. Attribute-Mapped Scroll Bars, US, 1990.
32. Zhai, S., Smith, B. and Selker, T. Improving Browsing Performance: A Study of Four Input Devices for Scrolling and Pointing Tasks. in *Proc. INTERACT'97.* (1997), 286-292.
33. Zipf, G. *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. Addison-Wesley, Reading, Mass., 1949.