The Effect of Information Scent on Searching Information Visualizations of Large Tree Structures

Peter Pirolli
Xerox PARC
Palo Alto, CA 94304, USA
Pirolli@parc.xerox.com

Stuart K. Card
Xerox PARC
Palo Alto, CA 94304, USA
card@parc.xerox.com

Mija M. Van Der Wege
Xerox PARC
Palo Alto, CA 94304, USA
mija@psych.stanford.com

ABSTRACT

Focus + context information visualizations have sought to amplify human cognition by increasing the amount of information immediately available to the user. We study how the focus + context distortion of the Hyperbolic Tree browser affects information foraging behavior in a task similar to the CHI '97 Browse Off. In comparison to a more conventional browser, Hyperbolic users searched more nodes, searched at a faster rate, and showed more learning. However, the performance of the Hyperbolic was found to be highly affected by "information scent", proximal cues to the value of distal information. Strong information scent made hyperbolic search faster than with a conventional browser. Conversely, weak scent put the hyperbolic tree at a disadvantage. There appears to be two countervailing processes affecting visual attention in these displays: strong information scent expands the spotlight of attention whereas crowding of targets in the compressed region of the Hyperbolic narrows it. The results suggest design improvements.

Keywords: Focus + context, hyperbolic tree, Information visualization, information foraging

1. INTRODUCTION

A key aim of information visualization research is to discover and develop ways of amplifying human cognition. One way to amplify cognition is to increase the amount of information that can be placed into users' attention. Focus + context techniques [1] are one class of information visualization techniques aimed at increasing the amount of information that is displayed to a user. Although there has been a great deal of research on designing and implementing focus + context techniques, there has been very little analysis of their underlying assumptions, and little empirical study of their impact on cognition and attention.

1.1 Focus + context should accelerate browsing

The Hyperbolic browser [2], presented in Figure 1, is an example of a focus + context technique. This browser is used to display large hierarchical tree structures. As can be seen in Figure 1, more display space is assigned to one part of the hierarchy (focus) than others (context). This is achieved by laying out the hierarchy in a uniform way on an imaginary hyperbolic plane and then mapping this to the Euclidian space of the display region. The hierarchy can be mouse-dragged through the central display region to bring new parts into the focus, or nodes can be mouse-clicked to bring them to the center of the focus.

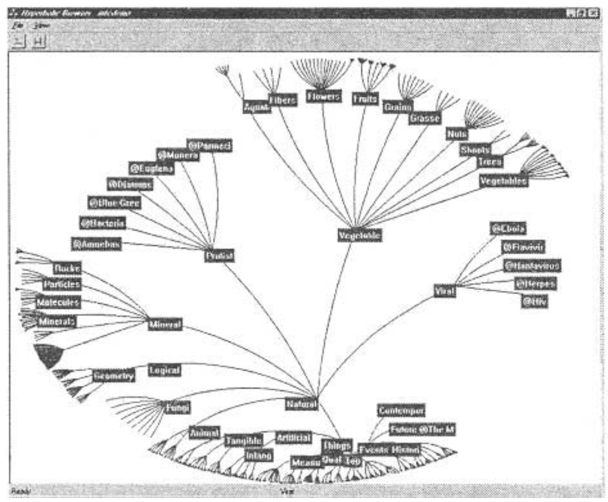


Figure 1. The Hyperbolic browser

1.2 Information scent and visual search

The tree displayed in Figure 1 contains about 10,000 nodes. A standard two-dimensional tree-layout algorithm would not be able to give a detailed presentation of all the tree on a standard display screen. Like many focus + context techniques, the Hyperbolic browser uses distortion to get all the information (the tree structure) into the display space. Conceptually, the Hyperbolic visualization would distort a rectilinear mesh in Euclidean space into something like the mesh presented in Figure 2a. Information in the focus is "stretched" to occupy more pixels, while information in the context is "squeezed" to occupy fewer pixels. Thus the user can view detailed information in the focus while having all information present on the display.

Because more of the tree structure is accessible on the display, the Hyperbolic browser is expected to accelerate users' browsing performance over conventional tree browsers. It is expected that distortion effects, such as Figure 2a, would enable users to visually search more of the information structure per unit time and would enable users to move through greater distances in the tree structure on each mouse-drag or mouse-click.

1.3 Visual search studies suggest a more complicated story

Results from the field of visual search and attention, however, suggest that focus + context display distortions, such Figure 2a, may actually have a complicated effect on the efficiency of visual search. Studies [3] indicate that the efficiency of visual search can sometimes be affected by the density of information on the display. To use a common metaphor, we may imagine that visual attention is like a spotlight. The size of this spotlight is usually called the useful field of view (UFOV). Information within the boundaries of the UFOV has a higher propensity for being perceived and processed than information outside the UFOV. Figure 2b, illustrates the situation in which the attentional spotlight is unaffected by the density of information on the display. The UFOV remains the same size regardless of where the spotlight is aimed. In contrast, Figure 2c illustrates the situation where the attentional spotlight is affected by the information display density. In this case, the UFOV decreases in size with density. Furthermore, one expects visual search over the entire display to be less efficient in Figure 2c than in Figure 2b.

We propose that the notion of *information scent*, developed in information foraging theory [4] can be used to predict the circumstances under which the attentional spotlight will be affected by the density of information in a visualization. Information scent is provided by the proximal cues perceived by the user that indicate the value, cost of access, and location of distal information content. In the context of foraging for information on the World Wide Web, for example, information scent is often provided by the snippets of text and graphics that surround links to other pages. The proximal cues provided by these snippets give indications of the value, cost, and location of the distal content on the linked page. Computational modeling of human information foraging [4] suggests that users' browsing choices are based on the evaluation of information scent.

With respect to focus + context techniques, we hypothesize that when the information scent of some target items "pops out" from the information scent of background items, then the attentional spotlight will be less affected by the density of those background items (Figure 2b). When the information scent of target and background items is roughly the same, then the attentional spotlight will be affected by the density of background items (Figure 2c). This prediction has some relation to "pop-out effects" in studies of visual search. Such pop-out effects have been studied extensively in the context of preattentive vs. attentive visual search [5].[1] The typical task in such studies involves finding some target item displayed amongst a set of distractor items. In such research, preattentive visual search appears to pick out target information at a rate that is largely unaffected by the number of distractors in the visual display. This occurs when the target and distractors may be visually discriminated on the basis of what have become known as preattentive features, such as color. For instance, a red X, can be found in a field of black Xs at a rate that is largely unaffected by the number of black Xs on the display. Subjectively, the red X seems to "pop out" from the display. On the other hand, attentive visual search for a target occurs at a rate that is affected by the number of distractor items. Subjectively, there is no "pop-out" effect in attentive search

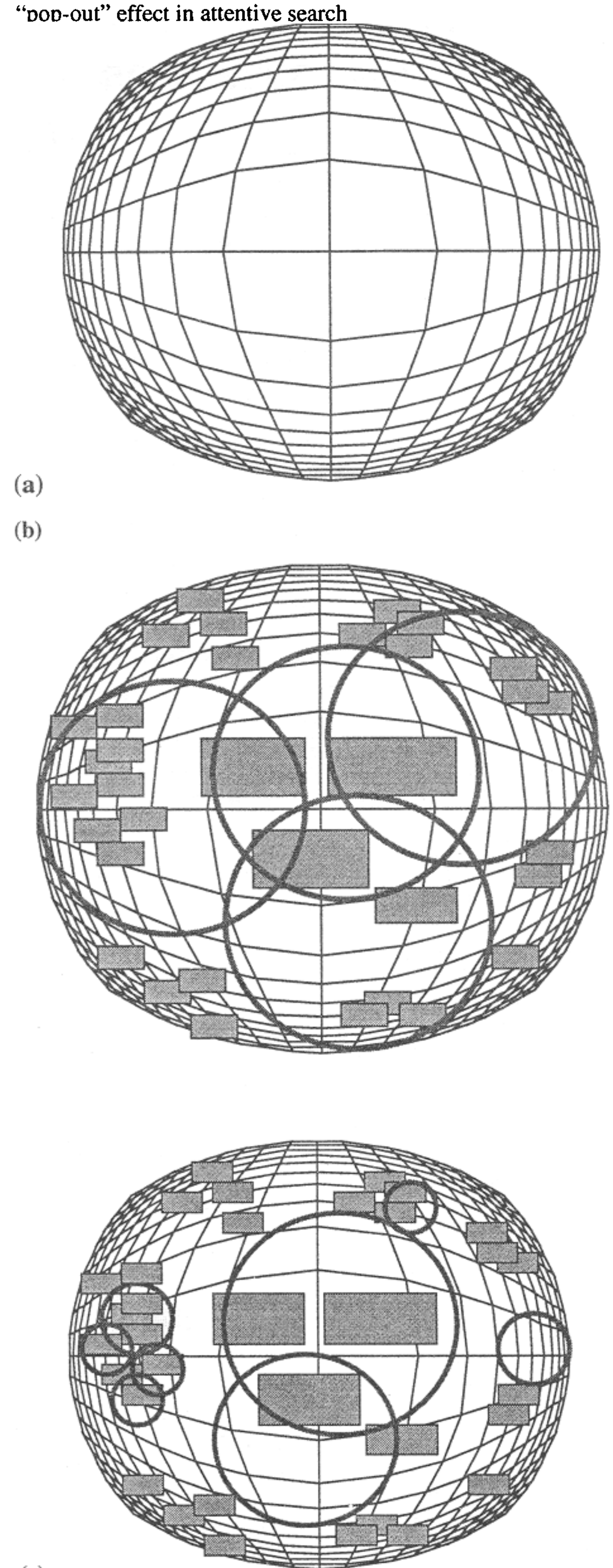


Figure 2. (a) a focus + context distortion decreases the density of information in the focus but increases density in the context, (b) the size of the attentional spotlight (circles) may be unaffected by density, or (c) it may decrease with information density.

[1] Although see [6] for the strong case that this "standard understanding" is overly simplistic.

Models of preattentive vs. attentive search, however, do not directly address the size of attentional spotlight (the UFOV). Such research also concentrates on pop-out or non-pop-out effects due to visual features, including motion and texture. We hypothesize a direct relation between the notion of information scent and the UFOV. Furthermore, our notion of information scent is not tied exclusively to visual features. In the tasks studied in this paper, information scent is mainly determined by linguistic information.

1.4 Overview

We next present a brief review of a public demonstration experiment of different browsers conducted at the CHI '97 conference. This public test provided the tasks used in our studies. We then discuss a study that obtained user judgements about the information scent of the tasks. Two experiments using these tasks are then presented. These experiments are aimed at understanding how use of the Hyperbolic browser differs from the use of a more conventional browser, the Microsoft Explorer browser, which is probably the most commonly used browser in use. We were specifically interested in understanding how browser use changed with changes in the information scent of tasks.

2. THE GREAT CHI '97 BROWSE-OFF

Our interest in studying the Hyperbolic browser began after a public competition among browsers. The CHI '97 meeting in Atlanta, GA presented a panel called The Great CHI '97 Browse-off [7]. The aims of the panel were partly entertainment, partly evaluative, and partly to spur on further research on the evaluation of browsers. As summarized later by the organizers:

> The Great CHI'97 Browse-Off provided attendees of CHI'97 in Atlanta with an opportunity to see six leading structure visualization and browsing technologies for an entertaining yet informative "live" comparison. Users of each system competed "head-to-head" in a series of races designed to simulate the stressful conditions under which real world browsing often takes place. Expert and (for two systems) novice operators used the visualization and browsing tools to complete a set of generic retrieval tasks as quickly and accurately as possible within a large hierarchical data set. Attendees were able to see for themselves which techniques worked well or poorly for various classes of retrieval problems.
>
> Every entry had something special to offer (two heroic contestants using only a DOS command-line shell became an instant audience favorite by staying close to the leaders through the first two rounds), but the top performance was turned in by Ramana Rao using the Hyperbolic Tree (tm) technology from Inxight [Software, Inc.]. The Hyperbolic Tree proved itself to be extremely responsive, graphically efficient, and devastatingly effective in the hands of a skilled operator using novel techniques like "fanning" the data in a focus-plus-context display.
>
> -K. Mullett and D. Schiano, BayCHI Meeting Announcement, August 1997
>
> (http://www.BayCHI.org/meetings/archive/0897.html)

The Hyperbolic browser was the clear winner among a field of research prototypes and off-the-shelf systems.[2] We wanted to understand how a successful browsing system worked. To do this, we performed a more thorough characterization of the tasks used at the CHI '97 competition and studied browsers under controlled laboratory conditions using more sensitive instrumentation, including an eye tracker.

For the purposes of our studies, we wanted to contrast performance on the Hyperbolic system against another system. We chose the runner-up in the Great CHI '97 Browse-off: the Windows Explorer. This is a widely used standard application used to view the file system in the Microsoft Windows operating system (Figure 3) and is a variant on what has become the conventional way of showing file hierarchies, based on the Apple Macintosh Hierarchical File system. The Explorer has two views. The first is a left-to-right tree-like layout. In this tree view, folders may be opened or closed to view or hide subfolders and files. In the second view, files and folders in the currently selected folder are listed. Clicking on folders in either the tree view or file view changes both views. Folders (but not files) can be viewed in either view.

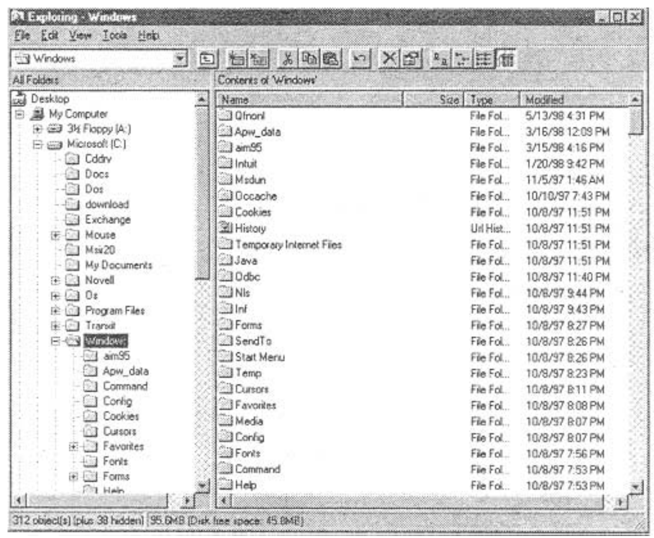


Figure 3. The Microsoft Explorer browser

3. MEASUREMENT OF INFORMATION SCENT

The Great CHI '97 Browse-off provided entrants with a tree data structure, which was to be displayed in the browsers. This tree was compiled in an ad-hoc manner from a variety of on-line sources. It was intended to reflect an ontological hierarchy. A portion of this hierarchy can be seen in Figure 1, which is a Hyperbolic browser.

A set of N = 128 tasks were compiled from those used at the CHI Browse-off and at a follow-up Browse-off held at a meeting of BayCHI, the San Francisco Bay Area division of SIGCHI. Not all were used in the events. The organizers had divided these tasks into four types. *Simple retrieval tasks* required finding a leaf node in the tree; e.g., "Find Lake Victoria." *Complex retrieval tasks* also involved finding leaf nodes, but involved some ambiguity and lack of familiarity; e.g., "Which

[2] One of the authors (Pirolli) was on the Hyperbolic Tree team at the competition.

army is lead by a Generalissimo?" *Local relational tasks* involved examination of several nodes that were reasonably close together in the tree structures; e.g., "Which religion has the most holidays in this list?" *Complex relational tasks* required examination of several nodes in disparate parts of the tree; e.g., "Which Greek deity has the same name as a space mission?"

Following initial pilot studies using the Browse-off tasks, we felt that some tasks seemed to involve less familiarity and greater ambiguity (in terms of knowing *a priori* its location in the ontological hierarchy) than others. Specifically, a task like "What's the highest rank you can achieve in Freemasonry?" seemed less familiar and more ambiguous than "Find a hammer?". Moreover, it seemed that these properties had a large influence on performance. We felt it was necessary to control for these factors in our experiments. Consequently, we developed some normative data about these tasks. These normative data were later used to operationalize the notion of information scent in our experiments.

3.1 Method

Our normative data gave us information on how much the participants knew about each term and how well they could determine the location of an item in the tree by looking at the labels on the upper branches.

Participants. $N = 48$ Stanford University students and members of BayCHI were paid to answer our survey.

Materials. The questionnaire contained two questions for each of the 128 Browse-off tasks (a term, such as "Ebola virus" to be found in the tree). Along the left side of each page of the questionnaire was a tree diagram depicting the top four levels of the Browse-off tree data (The actual terms to be found were farther down in the tree). Beside the fourth-level nodes were identification codes.

Procedure. For each of the 128 tasks, the instructions asked participants (1) to rate their familiarity with the term on a 7-point scale, and (2) to identify their top choices of categories for locating the answer to tasks (using the identification codes on the diagram). For example:

1. Find Lake Victoria.

Rate your familiarity with this subject:

1	2	3	4	5	6	7
Not Famili			Moderat Famili			Very Famili

Choices:

1. ____________
2. ____________(optional)
3. ____________(optional)

3.2 Results and Discussion

For each task term, we calculated mean familiarity scores. We also found the locations participants identified as the answers—the "modal answers." To measure how well these places were identified for our participants, we calculated the percentage of people who agreed with each modal answer (the percentage modal scores). The familiarity scores and the percentage modal scores for the 128 tasks had substantial Spearman rank correlation, $\rho = 0.51$.

From these data, we developed an *information scent* score:

> *Information scent* = the proportion of participants who correctly identified the location of the task answer from looking at upper branches in the tree.

This, of course, is an instantiation for our particular situation of the more general notion that information scent is the use of proximal cues to lead the way to distal information. A priori, tasks with high scent scores should lead people to the correct answer locations better than low scent scores. This gave us another way of classifying the tasks in the Browse-off contest.

4. EXPERIMENT 1. EFFECT OF BROWSER DESIGN AND INFORMATION SCENT ON TREE SEARCH.

The first experiment was an exploratory experiment to help us understand better how the design of the visualization and interaction components of a browser effect performance. We wanted to examine differences in the use of the Hyperbolic and Explorer browsers and understand their actual affect on user performance. We also wanted to study the interaction of the browsers with information scent. In particular, we wanted to collect data relevant to the analysis of visual search on the Hyperbolic browser under different information scent conditions so as to learn more about focus + context visualizations. In regards to this goal, we used an eye tracker to collect eye fixations during browsing.

As discussed above, we expected that information scent might interact with attention in a focus + context visualization. Specifically, in high information scent conditions we expected the useful field of view (UFOV) to be relatively independent of the density of information on the display. In low information scent conditions, the UFOV would decrease with information density. This hypothesis leads us to expect visual search on the Hyperbolic browser to be relatively more efficient on high information scent tasks than on low information scent tasks.

4.1 Method

Participants. $N = 8$ participants were recruited from the Stanford University Psychology Graduate program, and from Xerox PARC. Some recruits were eliminated due to problems with eye-tracking. The Stanford students were paid $50 for their participation.

Apparatus. We used the Hyperbolic and Explorer browsers described above. An ISCAN RK-426PC eye tracker was used to record eye fixations and saccades.

Materials. For the test portions of Experiment 1, we selected 56 of the 128 Browse-off tasks. These 56 were divided into two test lists, with each list containing seven tasks of each type: simple retrieval, complex retrieval, local relational, and global relational. To the extent possible, we matched tasks on their scent scores across lists and across task types. These scent scores were the ones obtained from the survey discussed above.

We drew an additional 56 tasks to use as practice tasks and also divided these into two lists. Tasks on the two practice lists were also matched for their scent scores. For the purposes of this matching we collected tasks into seven levels of scent scores centered around scent = 0.00, 0.10, 0.15, 0.25, 0.30, 0.35, and 0.40.

Procedure. The participants proceeded through (a) a familiarization phase, (b) a practice phase, (c) a test phase, and (d) a retest phase. During the familiarization phase participants read on-screen instructions that described the browser's basic functions. They were then invited to become familiar with the browser by exploring a hierarchy unrelated to the tasks in the experiment.

After the participant expressed a degree of comfort with the browser, the practice phase began. During the practice phase participants were presented with one list of practice tasks with one browser and the other list with the other browser. Each of the two lists was a randomized block of 28 tasks. Experimental tasks were counterbalanced so that half the participants began with Explorer and half began with the Hyperbolic browser.

After the practice tasks, the participants' eyes were tracked, using the ISCAN eye movement monitoring system. A brief session was devoted to calibrating the tracking system along a nine-point grid (the four corners of the screen, the midpoint of each side, and the center point). Following every set of 14 questions, the eye tracking was verified by having the participant re-trace the calibration grid. If the eye-tracking had drifted from the grid, the system was recalibrated. The subject took a break every 20-30 minutes (after completing each set of 14 questions).

The test phase was conducted in the same way as the practice phase, except the two test lists of 28 tasks each were used instead of the practice lists. For each participant, one test list was presented with one browser, then the second list of test 28 items with the other browser. List order and browser order were counterbalanced across participants, and the test items in each list were presented in a random order.

The retest phase occurred 1-3 weeks after the initial test phase. This phase consisted of additional practice and test phases identical to the first. Each subject saw the same items with the same browsers in the same order.

4.2 Results and Discussion

Preliminary analysis. An analysis of variance was conducted on the performance times recorded for the test and retest phase tasks completed by participants. Exploratory analyses showed that performance times had lognormal distributions. Consequently, we performed logarithmic transformations on the raw performance times prior to conducting statistical analyses.

A preliminary analysis of variance was conducted based on the Browser (Explorer, Hyperbolic) × Question Type (Local Retrieval, Global Retrieval, Local Relational, Global Relational) × Scent (seven levels) × Test Session (two levels) factorial design. Table 1 presents the mean time to complete tasks of different Question Types in the different browsers. There were significant differences among the task times for the different question types, [$F(3, 832) = 117.13$, $MSE = 0.20$, $p < 0.001$]. More detailed post-hoc tests indicated that this difference in question types was due to a difference between the Retrieval Tasks (Simple and Complex) vs. the Comparison Tasks (Local Relational and Global Relational) and their relationship to Information Scent, [$t(832) = 7.62$, $MSE = 0.015$, $p < 0.0001$]. The subsequent analyses of performance times reported below were conducted using an analysis of variance that collapsed the Question Type factor into just two levels: Retrieval and Comparison.

Table 1. Mean performance times in Experiment 1 by task type and browser.

Question Type	Browser: Explorer (sec)	Browser: Hyperbolic (sec)
Retrieval Tasks		
Simple	35.55	34.37
Complex	41.55	42.02
All retrieval	38.55	38.20
Comparison Tasks		
Local	42.78	41.91
Global	71.07	73.19
All comparison	56.93	57.55
All questions	47.74	47.87

Browsers. There was no significant difference in the overall task times between the Hyperbolic and Explorer browsers [$F(1, 832) = 0.14$, $MSE = 0.27$]. Our failure to find an overall difference between the two browsers was somewhat surprising. At the Great CHI '97 Browse-off, the Hyperbolic browser had appeared clearly superior to the Explorer in terms of task performance times. The victory was repeated in a separate contest at the BayCHI meeting. Furthermore, our participants expressed a preference for the Hyperbolic over the Explorer browser.

The first factor we investigated to account for this difference of outcomes was individual differences of browser operator. Recall that at the Great CHI '97 Browse-off each browser was operated by a person who was an expert at using that browser. It could be the case that the performances seen at the Great CHI '97 Browse-off were mainly due to differences among the individual experts rather than due to differences among the browsers.

To test this, we ranked participants' performance in our experiment with the Explorer, then ranked the participants by their performance on the Hyperbolic browser. The correlation in the two rankings, by Spearman rank correlation was $\rho = 0.78$, which is significant, $p < 0.01$. This indicates that individual performances can overwhelm browser design for the overall task. One possible reason this might be true is that a number of the tasks involved finding information in non-obvious places and remembering where it was, thereby giving a role to individual factors, such as participant preparation, ability to remember locations, and performance tricks with the browsers.

We see a similar indication when we examine the sums of squares (*SS*) of the different factors and their contribution to the total *SS* in the analysis of variance reported above. In Experiment 1, the participant effects $SS = 8.58$ and the Browser

SS = 0.0363 (in an experiment where the total sums of squares was SS = 345.67, and the error sums of squares was SS = 167.55). That means (for Experiment 1) that individual differences factors contributed more to the performance times than differences between browsers, although neither is large (because of the amount of variance in other factors such as task differences and learning). The contest participants were more highly trained than our experimental participants, potentially magnifying individual differences.

Overall, therefore, there were no net differences between the browsers. Performance with both browsers improved with practice [$F(1, 860) = 95.68$, $p < 0.001$]. There were no other main effects or interactions. However, there were interesting differences between the browsers that are hidden in the average result. These were revealed when we investigated a second factor, information scent.

Information scent. The picture changes in interesting ways when we look at the interaction of information scent with the design of the browsers. Overall, information scent, as computed from by our technique, reduces task time[$F(6, 860) = 8.25$, $p < 0.001$]. More specifically, Figure 4 shows that information scent has little effect on time for complex comparisons, but strongly reduces time for retrieval tasks [$F(6, 860) = 12.16$, $p < 0.001$]. (Although there seems to be a small slope for Comparison Tasks in Figure 4, a fit of the logarithmic values gives a line with a tiny slope), we focused further analyses on retrieval tasks.

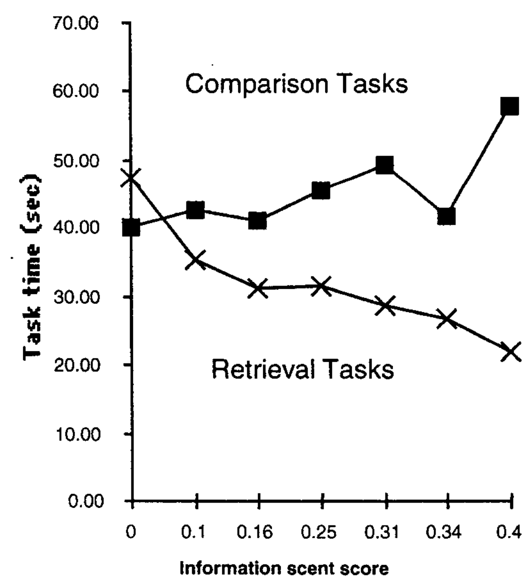


Figure 4. Performance times in Experiment 1 as a function of question type and information scent.

In retrieval tasks, information scent seems to act on the two browsers differently. As Figure 5 shows, both browsers are faster when there is higher information higher scent. But the Hyperbolic browser seems to be faster than the Explorer at high scent and slower at low scent. The analysis of variance of this interaction between scent and browser is marginally significant [$F(6, 860) = 2.04$, $p < 0.06$]. For the analysis, Figure 5 collapses all tasks with scent scores less than or equal to 0.16 into low scent scores and all tasks with scent score greater than 0.16 into high scent scores.

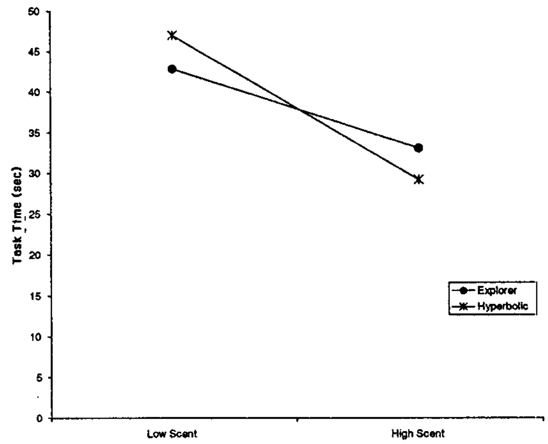


Figure 5. Mean performance times for retrieval tasks as a function of browser and information scent.

Effect of Information Scent on Eye Movements. We can see more of the difference between the browsers by looking at how information scent interacts with eye movements.

Preliminary analysis. The ISCAN eyetracker software segments eye-movement data into fixations and saccades. For the data available for each task trial, we further analyzed the data to determine the number of fixations. A number of hardware and calibration problems lead to missing data for some task trials. Consequently, we had to run less complete analyses of variance. Exploratory data analysis revealed that the number of fixations showed lognormal distributions. Consequently, an analysis of variance was conducted on logarithm-transformed fixations. This analysis of variance was conducted based on the Browser × Question Type × Scent × Test Session factorial design. Analyses were performed using all of the four question types.

Fixations. As shown in Table 2, on average participants using the Hyperbolic browser made more fixations per task than with the Explorer browser [$F(1, 787) = 16.67$, $MSE = 0.26$, $p < 0.001$], and the average duration of each fixation was shorter [$F(1,787) = 259.24$, $MSE = 0.013$, $p < 0.001$]. As Figure 6 shows, low information scent increased the number of fixations for the Hyperbolic browser more than for the Explorer, $F(1, 787) = 5.31$, $p < 0.05$. In fact, there is no reliable statistical difference between the low and high scent comparisons for the Explorer. This suggests that visual search becomes relatively more difficult for the Hyperbolic compared to the Explorer as information scent decreases. There was also an effect of Question Type, $F(3, 787) = 6.98$, $p < 0.001$.

Table 2. Eye fixations using browsers.

	Hyperbolic	Explorer
Fixations/Task	265	227
Mean Time/Fixation (msec)	72	93

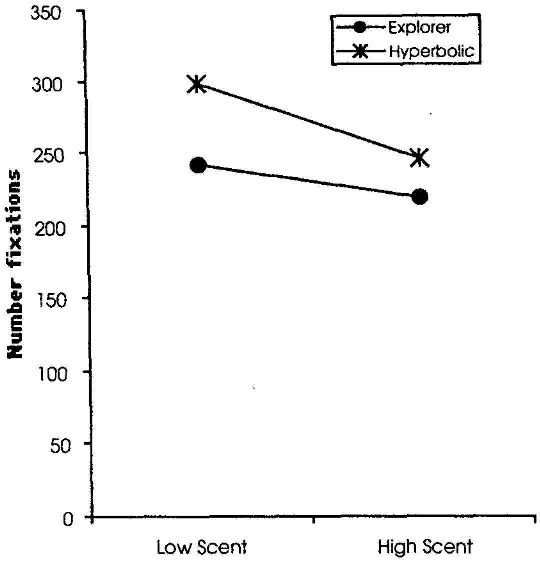


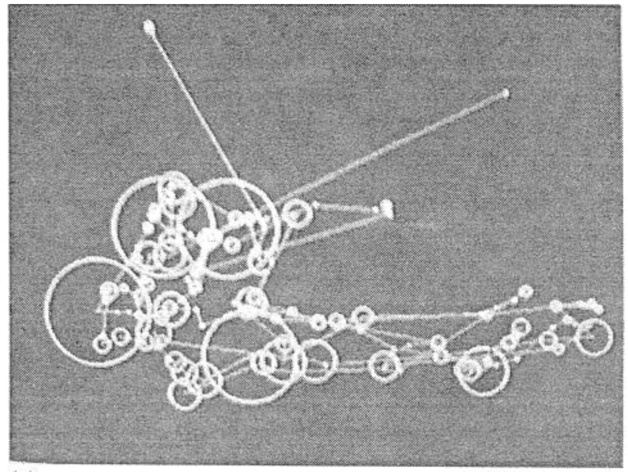
Figure 6. Number of fixations in Experiment 1 as a function of browser and information scent.

Scent-finding and scent-following. To understand more about how the browser designs affected users' eye movements, more detailed analyses of eye-scan paths were conducted for a high scent and a low scent Simple Retrieval tasks using each of the two browsers. These detailed analyses were done by hand from videotapes recorded during the test sessions of Experiment 1. The videotapes recorded the users' screens, as well as the point-of-regard as determined by the eye tracker (i.e., where the users' eyes were gazing on the screen). Because of the difficulty of such hand coding, we were only able to analyze a small subset of tasks in this manner. The first Simple Retrieval task is "Find the Ebola Virus" which has a Scent Score of 0.44. The second task is "Find the Library of Congress," which has a Scent Score of 0.12.

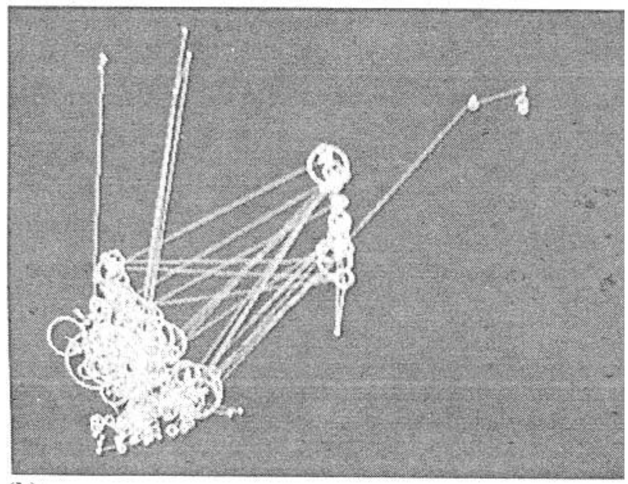
Figure 7 shows typical scan patterns for the Hyperbolic and Explorer browsers. The Hyperbolic browser uses more of the screen. The Explorer browser involves concentration on two smaller regions of the screen corresponding to the tree view and the folder view.

Watching the recordings of eye-movements recorded by the eye-tracker gives the impression that there are at least two modes of visual search activity. The first we call *scent-following* [8]. This kind of activity seems to be very directional, as if the eye focus were following cues up a gradient towards a maximum reward. The activity is very reminiscent of an organism following a stimulus gradient (scent) towards a reward [9]. The other mode of visual search seems to be a non-directional *scent finding* activity, aimed at finding directional cues. This activity is very reminiscent of organisms who have been alerted to a scent (e.g., a puff of pheromones), but must acquire additional cues to identify the direction of the reward [9].

Figure 8 shows a typical pattern of simple scent following. The x-axis measures time in seconds. The y-axis shows the depth of the node in the tree. The line indicates the level of the node on which the participants eyes are fixated as a function of time. Attention to a node was assumed to be indicated by an eye fixation or mouse-click on a node. The curve in Figure 8 moves monotonically and rapidly upward, indicating that the user progressed deeper into the tree with no backtracking. The short plateau on each level indicates that the participant had multiple fixations at the 2nd, 3rd, and 4th levels.



(a)



(b)

Figure 7. Typical eye scan and fixation patterns for (a) the Hyperbolic browser and (b) the Explorer browser.

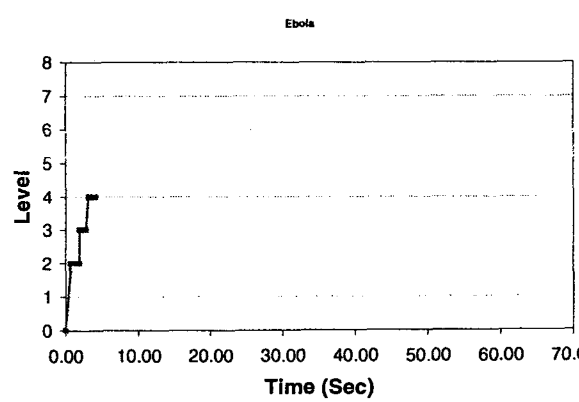


Figure 8. Scent-following in the Hyperbolic browser.

Figure 9 shows the eye-movements for the same high-scent search plotted for a participant using the Explorer browser. The pattern is similar. Triangles indicate mouse clicks. We have also plotted as "Level –1" whenever the user looked back to reread the question. The pattern in Figures 8 and 9 seems to be more or less common across the two browsers, with the Explorer browser being somewhat slower.

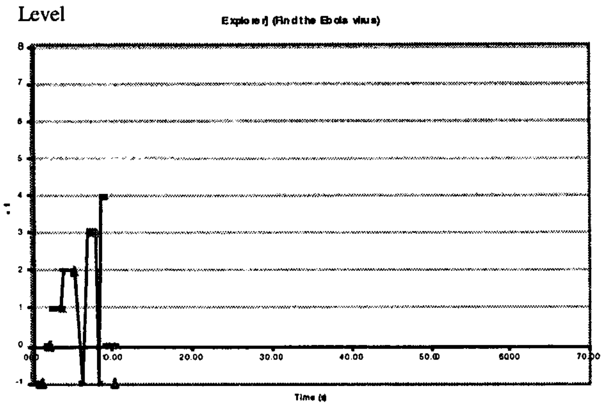


Figure 9. Scent following in the Explorer on the Ebola task.

Figure 10 displays what a difficult case looks like for the Hyperbolic browser. We have added a square symbol for the use of the mouse for dragging the display. The case starts out with scent following, which we see in the series of mouse clicks. Notice that the eye movements indicate that the user is typically looking one or two levels ahead. When the scent following fails, this subject begins dragging the display to reveal different places for examination. This is typical of scent finding activity.

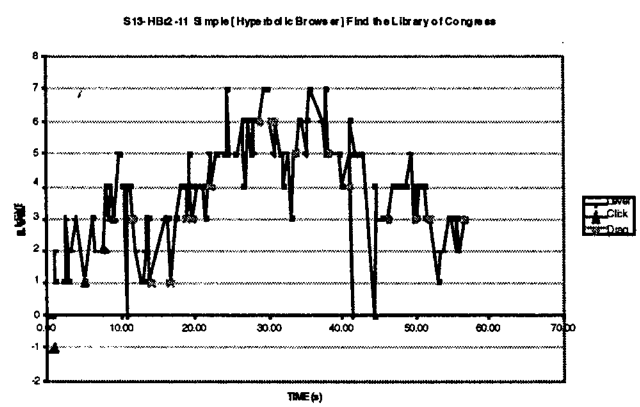


Figure 10. Scent finding and scent following on the Hyperbolic browser for the Library of Congress task.

Figure 11 is the same difficult case for the Explorer. We have added to the diagram circles, which indicate the manipulation of controls and the scrolling bar as another location for the eye. We see that superimposed on the search behavior there is a amount of behavior devoted to the manipulation of controls, especially as the display grows, to scrolling. Scrolling involves taking the eyes off of the primary display and added control manipulation. We have also indicated when the user is using the tree and when the user is using the list view of the explorer. This user uses both displays, with the tree view being used to explore into deeper levels. Each node is more expensive to explore in the Microsoft Explorer because it involves more control manipulation, so the user looks at fewer nodes. The descriptive statistics are in accord for the 16 tasks (4 participants x 2 task types (low scent, high scent) x 2 browser types (Hyperbolic, Explorer)) whose eye movements we examined by hand. If we look at scent-following moves (moves down the hierarchy on the current path) we find that the Hyperbolic users move down at a rate of 1.2 forward links/move as opposed to Explorer users who move at rate of 1.0 links/move. The Hyperbolic users explore more of the hierarchy (14.5 total distinct paths from root to leaves) than Explorer users (11.9 total distinct paths).

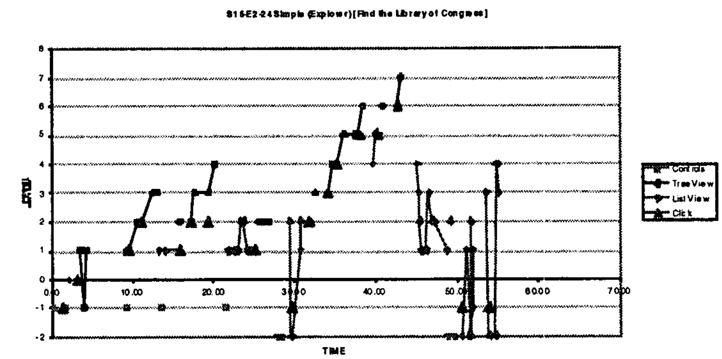


Figure 11. Scent-finding and scent-following on the Explorer for the Library of Congress task.

Figure 12 superimposes all 8 individual cases of the high-scent task. The x-axis has been enlarged relative to Figures 8 to 11 to make the individual paths visible. It is clear from Figure 12 that search in the high-scent case with the hyperbolic browser is much faster. In fact, fitting a regression through the points (plotted as a straight line in the figure) shows that the Hyperbolic browser requires only 0.92 sec/level compared to 1.75 sec/level or 53% as long.

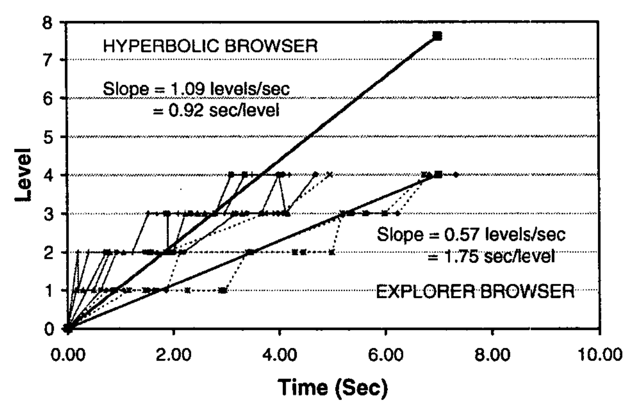


Figure 12. High-scent search

4.3 Summary.

Over all the question types, there was no difference in the time required by the Hyperbolic and Explorer browsers. It is possible that the Hyperbolic browser won its tournaments because of more skilled users (which of course includes the possibility that the users were able to exploit advanced techniques available through the browser). The two browsers differ however in the way they can take advantage of information scent. When we examine users' eye movements, under high-scent conditions and for simple retrieval tasks, the Hyperbolic browser can traverse levels almost twice as fast as the Explorer can. But it is slower than the Explorer under low-scent conditions. Additionally, participants using the Hyperbolic browser use more fixations to do the task, but their fixations are shorter.

5. EXPERIMENT 2

The aim of Experiment 2 was to provide more rigorous statistical analysis of differences in visual search between browsers on different information scent tasks. The hand-coded analyses presented in Figures 8 to 11 provided descriptive data that were highly suggestive, but these analyses were too time consuming to carry out on enough tasks to perform statistical analysis. We decided to pursue a semi-automatic technique for coding users' visual search over the information visualizations.

This approach required developing instrumented versions of each browser. These instrumented browsers would provide logs of the display states of the visualization and the mouse-clicks and mouse-drags of the user. The space and time coordinates of the browser logs would then be synchronized with the space and time coordinates of the eye tracking logs. This synchronization was done by an analyst with the use of a playback simulator that integrated the two logs and allowed the analyst to coordinate time and space scaling parameters associated with each log. Once the logs were synchronized, they could be analyzed automatically to determine such things as the interface objects (e.g., tree nodes) being fixated by the eye.

In Experiment 1 we found that it was retrieval questions that produced browser effects and an interaction of browser with information scent effects. Consequently, in Experiment 2, we used only retrieval questions, and we selected tasks that were at the extremes of high and low information scent. Our analyses focused on performance time differences and differences in aspects of visual search such as number of nodes visited, number of paths explored, and the range of paths explored.

5.1 Method

Participants. Eight participants were recruited from Xerox PARC and Stanford University. Four participants (Experts) were experienced in the use of both browser systems and the hierarchical tree structure. The other four participants (Novices) were unfamiliar with the Hyperbolic browser and the dataset, though they were most likely familiar with the Explorer browser or a similar type of browser.

Materials. A subset of the tasks used in Experiment 1 were selected for study in Experiment 2. Of the original 56 test tasks, 8 were selected. All selected tasks were from the Simple Retrieval category. Half were of low information scent (0-.10), the other half were of high information scent (.35-.40). These were divided into two test lists of four questions each matched for level of information scent (two low scent and two high scent). Eight equivalent practice items were also selected and divided into two practice lists based on scent.

Apparatus. An instrumented version of the Hyperbolic browser was developed using source code obtained from the Inxight corporation. We were unable to develop a way of directly instrumenting the Microsoft Explorer, so we instrumented a prototype browser fragment called the VFM (Figure 13), also provided by Inxight. VFM contains a set of windows that operate in the same way as the Microsoft Explorer. One limitation of this instrumented prototype is that the users could not double click in the list view window (the right hand window) as they could with the original Explorer program. Otherwise the two browsers were equivalent for purposes of this experiment. The instrumented versions of the Hyperbolic and VFM browsers produced logs that contained records of the location of every window and object on the screen, all mouse actions, and all keyboard actions. Each display update and each action was time stamped. The ISCAN RK-426PC eye tracker was used to record eye fixations and saccades.

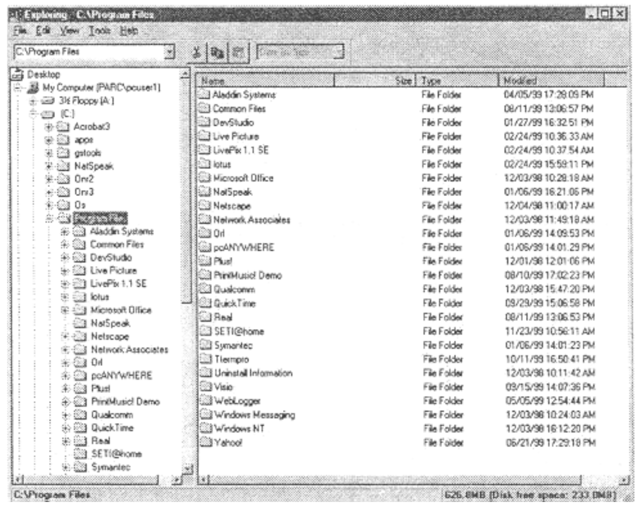


Figure 13. VFM lookalike to the Microsoft Explorer browser.

Procedure. As in Experiment 1, the participants proceeded through (a) a familiarization phase, (b) a practice phase, (c) a test phase, and (d) a retest phase. The familiarization phase was the same as in Experiment 1, although the Experts did not spend much time during it.

The practice phase was similar, but participants only saw four questions with each browser. They questions they saw were different with each browser, and the order of presentation of the two browsers was counterbalanced across subjects.

The eye-tracking system calibration was also the same as in Experiment 1. Subjects were calibrated to a nine-point grid, and the eye-tracking accuracy was verified after each set of four questions.

The test phase was again the same as in Experiment 1, except the two test lists were composed of four tasks each. For each participant, one test list was presented with one browser, then the second list with the other browser. List order and browser order were counterbalanced across participants, and the test items in each list were presented in a random order.

The retest phase occurred approximately one hour after the initial test phase. This phase consisted of just an additional test phase identical to the first. Each subject saw the same items with the same browsers in the same order as in the first test phase.

5.2 Results and Discussion

Performance time. In contrast to Experiment 1, the Hyperbolic browser obtained faster performance times than the VFM browser by about 62% (See Table 3), $F(1, 113) = 19.99$, $MSE =$

42.18 $p < 0.001$]. Actually, this result is consistent across the two experiments. When we use the data of Experiment 1 to examine the same 8 tasks used in Experiment 2, the Hyperbolic browser is faster although not significantly so [$F(1, 113) = 1.17$, $MSE = 0.25$], and it is faster when all the Simple Retrieval Task data combined (including Experiment 1, Experiment 2, and a third experiment not reported here) [$F(1,317) = 16.67$, $MSE = 0.31$, $p < 0.001$]. High scent tasks are faster then Low scent tasks (Table 3). There was no significant effect due to expertise [$F(1,113) = 1.88$].

As would be expected, practice improves performance for both browsers, but only for the low scent tasks (because the high-scent tasks are largely limited by the speed at which users act) (see Figure 14). But practice seems to help the Hyperbolic browser in low scent tasks more. This effect was marginally significant [$F(1, 113) = 3.12, p = 0.08$].

Table 3. Reaction Times in seconds for Experiment 2.

	Hyperbolic	VFM
Experiment 2	26.98	43.74
8 Simple Retrieval Tasks		
Experiment 1	33.64	37.01
Retrieval Tasks of Expt. 2		
Overall – Expt. 1 and 2	30.13	40.24
Experiment 2		
High Scent Tasks	16.12	28.76
Low Scent Tasks	82.95	101.79

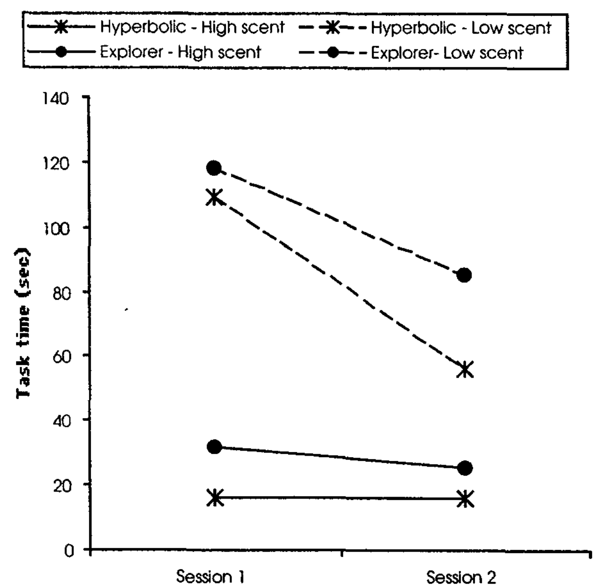


Figure 14. Performance times in Experiment 2 as a function of browser, information scent, and practice session.

Fixations. On these 8 Simple Retrieval tasks, participants had about the same number of fixations with each browser [$F(1, 108) = 0.01$, $MSE = 39.91$]. As can be seen from Figure 15, Low scent tasks required many more fixations than high scent tasks [$F(1, 108) = 94.87, p < 0.001$].

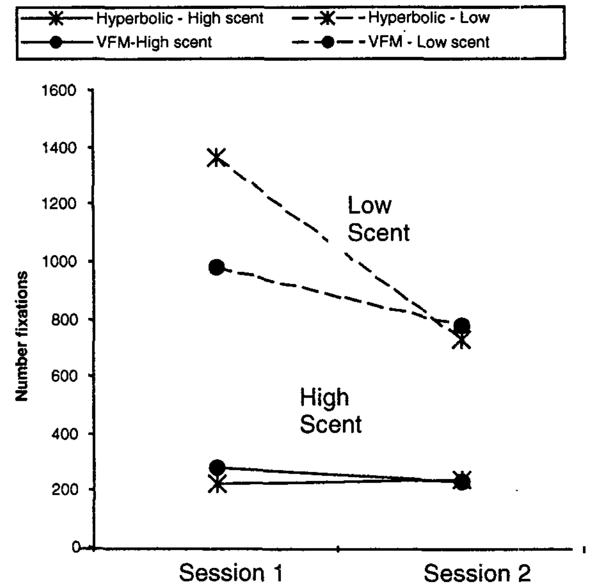


Figure 15. Number of fixations in Experiment 2 as a function of browser, information scent, and practice session.

As with performance times, practice had a larger effect on the Hyperbolic browser than on the VFM. Thus it is not surprising that experts require fewer fixations than novices $F(1, 108) = 5.25, p < 0.05$).

Number of nodes visited. Participants visited more nodes in the tree with the Hyperbolic browser (Figure 16) [$F(1, 114) = 77.89$, $MSE = 224$, $p < 0.001$]. Low scent tasks caused them to increase the number of nodes visited much more than was the case for the VFM [$F(1, 114) = 40.19$, $p < 0.001$].

This interaction of browser with information scent in Figure 16 supports the analyses in Figures 14 and 15. Users of the Hyperbolic appear to be more adversely affected by low information scent tasks than users of the VFM. In low scent tasks, the Hyperbolic users engage in more costly visual scent-finding search.

There is a pattern in these results. The Hyperbolic browser allows the user to access a target faster if the user knows where it is or at least the path that it is on (that is, if there is strong information scent)—about twice as fast as for the Explorer-style browser. If the user must engage in visual search, it is possible to search more nodes/sec. Practice or expertise has a strong effect on performance when scent is low. This might indicate that indicate that the Hyperbolic browser allows the user to learn more of the structure of the tree and cut down the search. Or it may mean that when the search space gets very large other factors begin to make the Hyperbolic tree less effective. This brings us back to our analysis of how the size of attentional spotlight may be altered by information scent in a focus + context display.

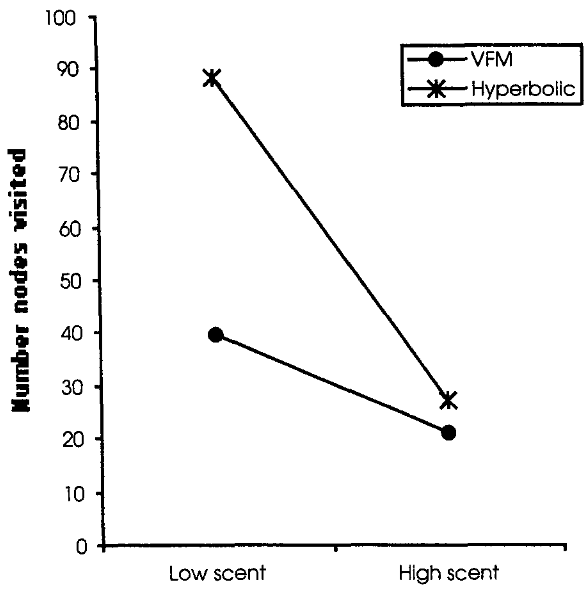


Figure 16. Number of nodes visited Experiment 2 as a function of browser and information scent.

Rate of downward tree search (scent-following). In our hand-analysis of fixation paths in Experiment 1, we found that Hyperbolic users were moving from the root of the tree down to the leaves at a faster rate than the Explorer users. In Experiment 2, the average number of levels traversed down the tree in a single move was: Hyperbolic mean = 1.3 levels and VFM mean = 1.1 levels, which is a significant difference [$F(1, 106) = 0.92$, $MSE = 0.10$, $p < 0.01$]. These values are also very close to those obtained in Experiment 1. Users of the Hyperbolic are able to visually search the tree in bigger jumps than users of the VFM because they can see ahead. This is one reason why searches are faster in the Hyperbolic tree if information scent is strong.

Eye movements in the Context area of the Hyperbolic Browser. In our discussion of Figure 2, we hypothesized that the size of the attentional spotlight (the UFOV) might be affected by the density of information on the visual display, under low information scent conditions. Under high information scent conditions, we hypothesized that the UFOV is relatively less affected by information density. More specifically we hypothesized that search in the Context (peripheral) area of the focus + context Hyperbolic display would be affected by information scent, as suggested in Figure 17: under high information scent conditions users might have a larger UFOV while searching the Context area, and consequently make longer fixation-to-fixation movements than in the low scent conditions.

To test this hypothesis we computed the radius from the center of the display out to every fixation in our data set (see Figure 17). (This radius, and all fixation movement distances, were computed from the ISCAN eye tracker coordinate system.) We then defined the Context area as fixations that occurred 0.8 of the radial distance from the center to the border. We then selected fixation-to-fixation movements that terminated in this peripheral Context area. As hypothesized by Figure 17, the high scent fixation movements were longer than the low scent movements: low scent movements had a median length of 7.21 (eye tracker distance units), and high scent movements had a median length of 8.94 (eye tracker distance units). This was about a 25% increase in the length of fixation-to-fixation movements with increased information scent, which was statistically significant [$F(1, 7) = 10.72$, $p < 0.02$]. This suggests that the attentional spotlight narrows with display density when there is low information scent, and broadens when there is high information scent.

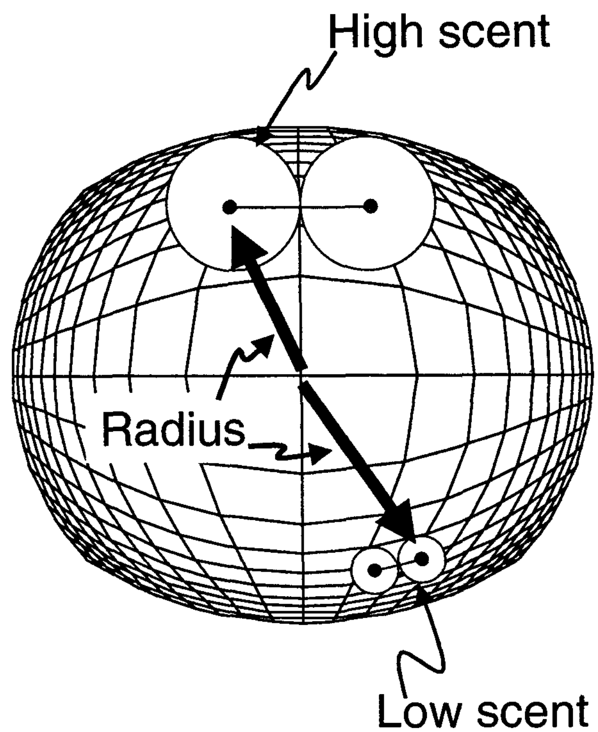


Figure 17. It is hypothesized that in the Context area of the display (a large radial distance from the center), low information scent conditions will have a smaller UFOV and shorter fixation-to-fixation path lengths than high information scent conditions.

5.3 Summary

In Experiment 2, we used a more restricted set of tasks than Experiment 1 and developed a semi-automatic method of analyzing visual search over browser interfaces. In Experiment 2, we found that:

- The Hyperbolic browser yielded better overall search performance than the VFM
- There was some evidence that the Hyperbolic users were learning more of the tree structure than the VFM users. This was indicated by a marginally superior learning effect with the Hyperbolic on low information scent tasks.
- Hyperbolic users examined more of the tree nodes at a faster rate than VFM users
- It appears that visual attention can be adversely affected by focus + context distortion techniques, under conditions of low information scent, as outlined in the discussion surrounding Figure 2.

6. CONCLUSION

In the case of the Hyperbolic Tree browser, there are a number of intuitive design improvements that make sense in light of our finding about the role of information scent. Providing landmarks is often proposed as a way of aiding navigation. In practice, Hyperbolic displays of datasets have landmarks that are colored differently than other items. It is also possible to use different colors for different subtrees. Both of these intuitive design improvements have the effects of improving the "pop out" effect of information scent, or making the information scent more discriminable.

Focus + context techniques attempt to deliver more information into the span of human attention. One important subclass of such methods uses distortion of the display to achieve this effect. The Hyperbolic Tree browser, as an instance of these methods, showed many superior aspects in comparison to the more conventional Explorer and VFM browsers. On retrieval tasks, we found that the Hyperbolic yielded better performance times and more learning of the data tree. Overall, it appears that users can examine more nodes at a faster rate with the Hyperbolic. However, we also found that the Hyperbolic browser was greatly affected by information scent, and this lead us to reexamine some of its underlying design assumptions in light of research on visual attention and visual search. The Hyperbolic browser, like many other information visualizations, seems to assume that "squeezing" more information into the display "squeezes" more information into the mind. The studies reported here suggest that this simple assumption is probably wrong. Visual attention and visual search interact in complex ways with the density of information on the display as well as the information scent or "pop out" of information from the display. Strong information scent made hyperbolic search faster than with a conventional browser. Conversely, weak scent put the hyperbolic tree at a disadvantage. There appears to be two countervailing processes affecting visual attention in these displays: strong information scent expands the spotlight of attention, whereas crowding of targets in the compressed region of the Hyperbolic narrows it. Further empirical studies of information visualizations, informed by basic research on visual search and visual attention may provide more complex formal models on which new design principles may emerge.

Acknowledgments: This work is supported by an Office of Naval Research grant No. N00014-96-C-0097. We would like to thank Paul Whitmore for his assistance in carrying out Experiment 1, Gwyneth Card for her assistance in coding the eye path data in Experiment 1, Ahna Girshick for developing the instrumented Hyperbolic and VFM browsers. We especially thank Ramana Rao and the Inxight Corporation for their cooperation in providing the Hyperbolic and VFM source code.

REFERENCES

1. Card, S.K., J. Mackinlay, and B. Schneiderman, *Information Visualization.* 1999, Los Altos, CA: Morgan Kaufmann.
2. Lamping, J., R. Rao, and P. Pirolli, *A focus + context technique based on hyperbolic geometry for visualizing large hierarchies*, in *CHI '95, ACM Conference on Human Factors in Computing Systems.* 1995, ACM: New York.
3. Drury, C.G. and M.R. Clement, *The effect of area, density, and number of background characters on visual search.* Human Factors, 1978. **20**: p. 597-602.
4. Pirolli, P. and S.K. Card, *Information foraging.* Psychological Review, 1999.
5. Ware, C., *Information visualization: Perception for design.* 1999, San Francisco, CA: Morgan Kaufmann.
6. Wolfe, J.M., *What can 1 million trials tell us about visual search?* Psychological Science, 1998. **9**: p. 33-39.
7. Mullet, K., C. Fry, and D. Schiano. *On your marks, get set, browse!* in *Human Factors in Computing Systems, CHI '97 (Extended Abstracts).* 1997. Atlanta, GA: ACM.
8. Pirolli, P. *Computational models of information scent-following in a very large browsable text collection.* in *Conference on Human Factors in Computing Systems, CHI '97.* 1997. Atlanta, GA: Association for Computing Machinery.
9. Bell, W.J., *Searching behavior: The behavioral ecology of finding resources.* 1991, London: Chapman and Hall. 358.