

Exploring Blog Spaces: A Study of Blog Reading Experiences using Dynamic Contextual Displays

Sven Laqua
University College London
Department of Computer Science
Gower Street
+44 (0)20 7679 0351
s.laqua@cs.ucl.ac.uk

M. Angela Sasse
University College London
Department of Computer Science
Gower Street
+44 (0)20 7679 7212
a.sasse@cs.ucl.ac.uk

ABSTRACT

In this paper we report on an eye-tracking experiment conducted with 60 participants to gain an understanding of how people interact with blog environments. We compared a standard blog interface with a novel contextual blog interface, which dynamically adjusts its contextual navigation to a selected article. We measured task performance and interaction behaviour for explorative tasks and goal-oriented search tasks. We further collected subjective feedback to evaluate user preferences. We found that participants using the contextual blog interface completed search tasks 19% faster and made 80% fewer errors. Moreover, participants using the contextual blog interface interacted more with the provided information during the exploration tasks. We did not find significant differences in user preference overall between both blog interfaces. However, a more detailed analysis of our results suggests significant demographic differences for performance, behavioural and subjective measures.

Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation]: User Interfaces - *Graphical User interfaces, Evaluation/methodology.*

General Terms

Design, Experimentation, Human Factors.

Keywords

Blogging, Contextualization, Contextual User Interfaces, Focus + Context Interfaces, Eye Tracking, Cognitive Style.

1. INTRODUCTION

The World Wide Web is increasingly about social interaction and collaboration. Blogging is a key activity in this Social Web enabling collective contributions of any type of information. Blogs have empowered millions of users to share their knowledge and experiences. But meaningful blogging experiences are as much about accessing information (reading) as they are about contributing information (writing).

The blogosphere (entirety of all blogs) faces the general problem of imbalance between ease of information contribution and meaningful information seeking. Millions of individual authors create millions of small and unique blog sites, and

compete for attention¹ in this messy space. Every contribution to this universal conversation - the actual content of a blog post - is wrapped into an individual visual design and a tailored structure of information through means of categories or tags. The dynamic nature of blogs quickly buries older content in archives or at best category lists reflecting the individual mind sets of their authors. In a sense, blogs are much like streams of individual thoughts. Finding useful information can be hard and time-consuming often with a negative impact on the interaction experience.

The main problem with information spaces as dynamic as the blogosphere is information discovery [2], and information-seeking going beyond undirected browsing is problematic. Commonly, users' desire to explore a variety of information sources to feel confident in their judgment on complex problems [3]. Users also find it harder to formulate clearly what the problem is [3] - something essential for effective search engine usage. With increasing task complexity, these two issues, (1) perceived quality of answers being bound to personal needs, and (2) people wanting to feel they make the choice, hold growing significance.

2. RELATED WORK

Previous research has shown that users' generally prefer orienteering strategies over keyword-based search [17]. The richness of contextual information provided through orienteering behaviour helps the user to understand where they are and where they want to go. By gradually narrowing down the space they need to explore, users can also circumvent the need to "articulate exactly what they were looking for" at the beginning. The actual sensemaking process is continuously shaping over time whilst dealing with the retrieved information.

However, readers' information goals when accessing content wrapped into layers of information structure, navigation and visual design often diverge from the authors' goals. This problem is particularly apparent in environments, as dynamic and diverse as the blogosphere.

RSS feeds are an interesting example of new technology designed to cope with the above mentioned information problems. RSS is particularly used in the blogging context to provide personalized information delivery. But, when looking at how phenomena like banner blindness [16] affect the perception of overloaded web sites and blogs it becomes clear that traditional interfaces represent a burden to the user [11].

Search is commonly regarded as "universal" solution to these problems [14], and research has shown that user behavior for blog search is similar to that in general web search [13]. Rather than redesigning the user interface to provide a seamless and

¹ This paper focuses on information-centric, rather than on personal blogs.

contextual *information experience*, current efforts are focused around *search-based interaction*. However, traditional search only works well in a limited subset of information tasks, and recent research suggests that even “the perfect search engine is not enough” [17].

Search result clustering and contextual search queries aim to add gradual improvements to traditional search techniques by serving more meaningful results to simple information queries, either by grouping results semantically [20] or by trying to understand the user’s task context [8].

Contextual user interfaces aim to provide *improved orientation* on information spaces and facilitate more *explorative interaction strategies*. The most important user interface (UI) techniques to display information within context (also: focus + context) are Fisheye views [6] and Degree-of-Interest trees (DOI trees) [4]. Both techniques have found a range of applications in scientific and mainstream scenarios.

However, most of the existing focus + context implementations only work on linear information spaces or require well-defined information structures. In addition, many implementations try to display as much context as possible. This richness of contextual information can help to understand large structures, relationships and to get a general overview.

But, displaying too many links simultaneously for very large information spaces can create visual noise. Too many concurrently displayed elements create information overload within the context itself. In contrast, presenting less information in a context more relevant to the individual user might provide easier access and a more pleasant information experience.

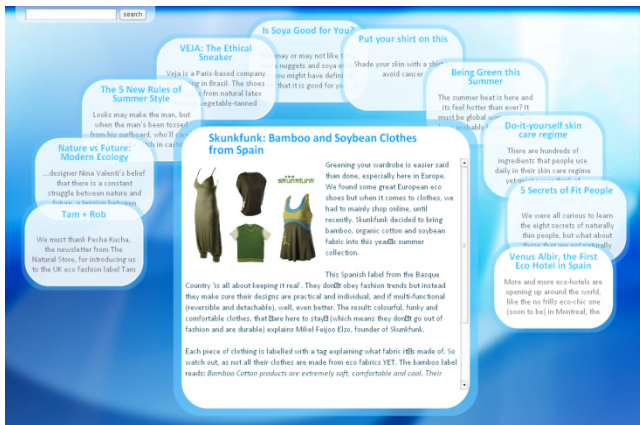


Figure 1. Contextual Blog Interface (FMI)

2.1 Focus-Metaphor Interface (FMI)

Users’ interaction with a focus-metaphor interface (FMI) [9, 10] enables seamless exploration of the underlying information spaces. This approach combines a contextual navigation with the actual display of information (see Figure 1) and particularly facilitates orienteering behavior [17].

In the blogging context, the FMI is mapped onto a blog space to provide convenient access to large amounts of blog articles. Its contextual interface elements are arranged around the central content element which displays the currently selected blog article in detail. The contextual elements function as navigation (activated through clicking) and provide previews of related blog articles much like snippets on search engine result pages (SERP).

When selecting a contextual element, its state changes: It enlarges into a content element and moves to the centre of the screen, replacing the previous element. The display of contextual elements is dynamically adapted to the new primary

content element using full-text similarity matching for the entire information space. The applied algorithm is inspired by a ‘similar posts’ plug-in for Wordpress [12] and uses MySQL’s full-text index and MATCH capabilities. This approach enables the dynamic adaptation of contextual elements to the currently displayed article. Selecting an article adjusts the context to the most similar / related articles. In a sense, the FMI “re-ranks” relevance of alternative articles and adjusts the contextual navigation accordingly; much like has been proposed for the display of search results by Teevan et al. [18].

Providing only similar content to choose from in the navigation should facilitate orienteering and support focusing on a specific task. In order to enable efficient task switching, a search tool has been integrated that provides the same functionality as traditional blog search. Switching to a completely different topic using directed search mimics Teevan’s concept of “teleporting” [17].

3. HYPOTHESIS

Findings of previous studies [9, 10] led us to hypothesize, that (H1) using the FMI will improve participants’ task performance (in the information search condition). We further suggest that the nature of the FMI will lead to (H2) increased interaction with and exploration of the provided information using the FMI. The more active engagement with the available information should lead to a superior information experience by (H3) participants favoring the FMI over the BlogUI.



Figure 2. Standard Blog Interface (BlogUI)

4. EXPERIMENTAL DESIGN

We conducted a study to investigate information experiences in blog environments, and whether a Focus-Metaphor display can improve them. A corpus of approximately 160 blog articles on lifestyles was used to create a blog environment for this study. The content came from popular blogs and blog-like news-sites that cover health related topics such as exercise, fitness, workouts, healthy foods, diets, drinking, environmental issues and fashion. The original blogs and websites have both male and female audiences.

The standard blog interface used in this study (Figure 2) deploys a Wordpress installation with a 2-column layout. A “traditional” theme has been chosen to be representative for the majority of blogs in the blogosphere. A list of articles is displayed in the left column and a category list is provided in the right column. Each blog article can be accessed through 2+ categories. Search has also been integrated as alternative means for navigation.

4.1 Participants

60 participants completed this study (31m/29f) from 18 to 67 years (median 28). Participants came from a range of

educational backgrounds and had varying levels of computer experience (though all used computers at least occasionally). Since the tasks involved a substantial amount of reading, we recruited only native English or bilingual speakers and said experiment was for those with an interest in a healthy lifestyle. Payment was £8 for a 1hour session.

A **computer expertise (CE)** measure was calculated through a number of demographic questions on computer usage and literacy from a pre-questionnaire. Normalized value ranges from 0 (very novice) to 1 (very experienced). The average CE score for our participant sample is 0.4. For comparative analyses, participants with a score of ≤ 0.4 have been labeled “novice” and participants with a score of > 0.4 have been labeled “expert”.

4.2 Independent Variables

4.2.1 UI Type

Two user interface types have been tested. A **traditional blog interface** (BlogUI, see Figure 2) has been used as baseline and compared against a **contextual focus-metaphor interface** (FMI, see Figure 1).

4.2.2 Task Type

Information experience was tested for two different task types:

- 1) A more flexible and user-centred condition, where **information exploration** has been tested through topical scenarios.
- 2) A more guided and goal-oriented condition, where **information search** has been tested through specific search tasks.

Exploration tasks (1) and search tasks (2) represent contrasting scenarios. They have been tested in separate sessions.

4.3 Dependent Variables

Task performance measures have been taken for task errors and task completion times (for task type 1).

Eye-tracking analysis focused on measure of fixation counts, gaze time and average fixation durations.

Navigation and interaction strategies have been measured through search usage and other types of interaction possible with the respective UI type.

Subjective evaluation measures have been taken through detailed usability questionnaires and a wealth of qualitative feedback.

4.4 Tested Scenarios (Task Types)

4.4.1 Information Exploration Tasks

The exploration tasks provided participants with a problem scenario and topic (see Table 1). Participants were then given time to “*explore information that ... provides useful insights concerning the given task*”. Participants were allowed to stop the task themselves (usually if they could not find more interesting information or felt they read enough) or were stopped after 5 minutes to keep the overall experiment time in a reasonable timeframe.

After completing a task, participants were asked to briefly reflect on the information found (e.g. “what was useful, or not?”, “which article was most interesting?”, etc.). The aim was to foster a deeper (and more realistic) involvement in the given tasks (and stimulating intrinsic motivation).

Table 1. Example of Information Exploration Task

Task:

“After your holidays, you have gained a few pounds. You are not happy and want to lose weight. You have seen from your friends that diets don't work. You believe that exercise is a much better way of getting back into shape.

There is a lot of useful information on sport and fitness in these pages Please have a look and find information on what sport is suitable for you and fits into your lifestyle.”

4.4.2 Information Search Tasks

Search tasks provided participants with a specific scenario describing a particular article to be found (see Table 2). Participants were free to choose and switch between the integrated search and other means of interaction. If a participant was unable to find the target article, she was allowed to stop the current task and proceed with the next task. In addition to 6 standard search tasks, we included an additional difficult task: Title and image of the target article did not provide an obvious link to the problems statement in task 3 (difficult task).

Table 2. Example of Information Search Task

Task:

“A colleague told you, he just read an interesting article on the pros and cons of drinking coffee every day. Since you worry about your consumption, you want to have a look at this article yourself...”

4.5 Method

Each participant completed 2 independent conditions in a between-subjects design *UI Type x Task Type* (see Figure 3 for details). Results of a small pilot study showed that switching between types of tasks is prone to errors. Particularly the more complex procedure of the information exploration tasks required careful explanation to the participants.

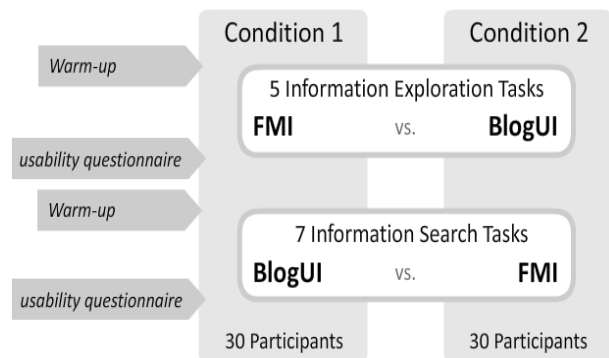


Figure 3. Schema of Experimental Design

In a pilot study, we found the best order to be running the information exploration tasks first and the information search tasks second (as the search tasks are more self explanatory).

The experiment was conducted in a usability lab using a Tobii X50 eye tracker. Warm-up sessions before running the actual experiment conditions ensured that all participants were introduced to the test environment and were given time to familiarize with the functionality of the user interface. Participants were calibrated separately for each condition to improve the quality of the collected data.

5. RESULTS

For the goal-oriented search condition, we computed task performance measures - error rates and task completion times. For the more behaviourally-oriented information exploration condition, we focused our analysis on user interaction measures such as articles read, searches made and categories chosen. Our analysis includes traditional eye tracking measures such as fixation counts and average fixation durations to support findings of performance and behavioural measures. Areas of Interest (AOIs) have been defined for navigational elements (Navigation AOI) and the content sections (Content AOI) within each respective UI. In addition we collected user feedback through detailed usability questionnaires.

We present the task performance results first, and the behavioural interaction analysis findings second. We then summarize the results of the eye-tracking analysis, and demonstrate how they relate to the performance and behavioural findings. Following the objective analysis to test H1 and H2, we present subjective measures collected through questionnaires to evaluate H3. Finally, we present some results of a more detailed analysis by demographic group and computer experience.

5.1 Task Performance

The information search condition required participants to conduct 7 tasks. As mentioned in the description of the search condition, 6 of these tasks were “standard” tasks, 1 task was more difficult, with the respective information hidden within the final few paragraphs of one particular article. The error distribution across individual tasks (see Table 3) highlights the problems participants had to successfully finish this difficult task (T3). Due to the structural differences of T3, we analyzed task performance for the remaining tasks separately. A detailed reflection on T3 and its design implications can be found in the discussion section.

Table 3. Task Errors for search tasks (across all participants)

	Task1	Task2	Task3	Task4	Task5	Task6	Task7
FMI	0	0	13	0	0	2	0
BlogUI	3	0	9	1	1	3	2

5.2 Error Rates

Overall, with 98.8% (SE_x=0.8%) for FMI versus 94.4% (SE_x=1.7%) for the BlogUI, both interfaces had high success rates (for tasks 1, 2, 4-7) in the information search condition (see Figure 4). Participants using the contextual navigation of the FMI made **~80% fewer errors**, than participants using the traditional BlogUI. A two-sample t-test underlines the significance of this difference with $t_{57} = 2.00$, $p < 0.02$.

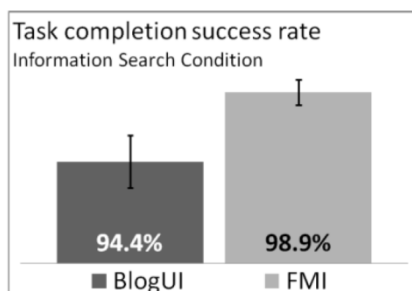


Figure 4. Task completion success rate in the Search Condition

One additional aspect to consider is the distribution of errors over the various tasks. Using the FMI, in 5 out of the 6

traditional tasks, all participants successfully completed their tasks. In contrast, using the BlogUI, only in 1 out of the 6 traditional tasks, none of the participants failed to complete the task successfully. This finding hints at a general task-independent problem with the BlogUI for information search tasks. In contrast, the error concentration for the FMI on specific tasks hints at a task-dependent problem for information search tasks. Future iterations of the FMI prototype should investigate this phenomenon to further minimize task error rates.

5.3 Task Completion Times

For successfully completed tasks, participants’ task completion times were **~19% faster** using the FMI (39 seconds, SE_x=2s) than for participants using the BlogUI (48 seconds, SE_x=3s) (see Figure 5). This difference is significant with $t_{338} = 1.97$, $p < 0.01$ (using a two-sample t-test).

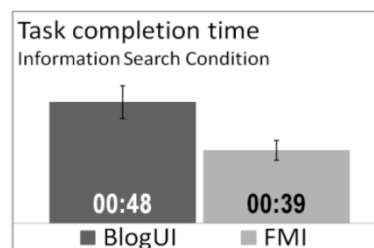


Figure 5. Task completion time (for successful tasks)

Another interesting finding is the much larger spread of successful task completion times for the BlogUI. Ranging from 6 seconds for the fastest to 212 seconds for the slowest task completion time, the spread of 206 seconds for the BlogUI is 44% larger than the spread for the FMI, with 115 seconds (9 seconds for the fastest, and 124 seconds for the slowest task completion time).

Table 4. Task Completion times for individual search tasks (across all participants) in seconds

	Task1	Task2	Task4	Task5	Task6	Task7
FMI	46 (SE _x =4s)	48 (SE _x =6s)	26 (SE _x =2s)	39 (SE _x =4s)	39 (SE _x =4s)	36 (SE _x =4s)
BlogUI	81 (SE _x =10s)	56 (SE _x =7s)	30 (SE _x =3s)	42 (SE _x =8s)	34 (SE _x =6s)	43 (SE _x =4s)

This finding is particularly interesting, when considering the very low error rate for the FMI. Although this user interface and its interaction technique were completely new to all participants, not only did they make much less errors, were faster overall, but participants also showed much more consistent task completion times, than in the more familiar BlogUI.

As both measures, error rates and task completion times, indicate a clear performance advantage for the FMI, we accept the first hypothesis (H1).

The limited scope of the tested information space is likely the cause for overall higher success rates and shorter task completion times than would be measured in a real-world context with large-scale and unstructured information spaces. However, considering that the significant advantage of the FMI might scale, its contextual navigation could offer huge performance benefits over traditional web-based navigation.

5.4 Information Interaction

In both UIs, participants interacted with information through the use of search and additional UI-specific interactions. Participants were free to choose at any point of each condition which means of interaction to use. In the BlogUI, a category list

has been the central navigation element. Choosing a particular category would load a list of related articles, which participants could scroll through. In the FMI, the main interaction mechanism has been its contextual navigation (elements).

5.4.1 Search Usage

For the use of search, our findings show that participants conducted significantly more searches (approx. 47%) using the FMI, with 1.52 searches per task for the FMI versus 1.03 searches per task for the BlogUI ($t_{56} = 2.00$, $p < 0.005$) in the information search condition (see Figure 6).

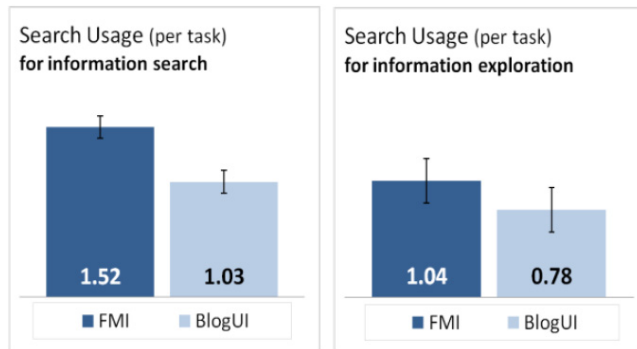


Figure 6. Search Interaction per task for both information task conditions and both UIs

Participants also conducted approx. 34% more searches using the FMI in the information exploration condition. However this result is not significant, due to strong individual differences in user behaviour: Approx. 1/3 of participants did not use search at all during information exploration, whereas other participants heavily relied on search (with a peak of 4.4 searches per task for both UIs). The lower scores for search usage during information exploration tasks are particularly interesting when considering the overall task duration of approx. 5min for this condition – compared to a task duration of 48 seconds (BlogUI) or 39 seconds (FMI) for the information search condition.

5.4.2 Other Interactions

In the exploration condition, participants preferred alternative means of interaction for both UIs (FMI: contextual elements, BlogUI: category list). In contrast, when faced with a specific search task, usage of the search tool increased (see Figure 7).

Using the BlogUI, participants clicked on average on 2.3 categories during each exploration task (approx. 5 min). Using the FMI, participants interacted on average with 7.2 articles per task (through selection of a contextual navigation element).

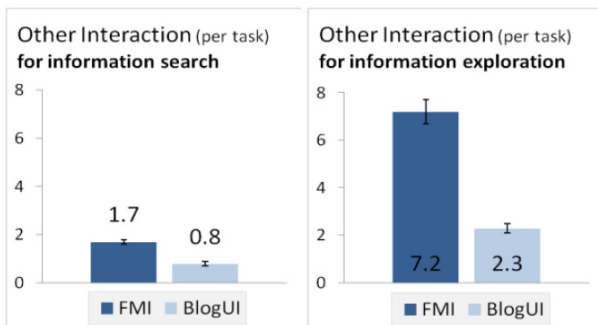


Figure 7. Other Interaction per task for both information task conditions and both UIs (category selection for BlogUI and context selection for FMI)

The distinctly different visual representations and types of interaction on the same blog space make a direct comparison of category interaction (BlogUI) and context interaction (FMI)

difficult. In the FMI, selecting a contextual element will display the one selected article plus previews of related articles by updating the contextual navigation elements. In the BlogUI, selecting a category item will load the respective category page with a number of related articles (up to 10). Here, the participant has no direct control over which specific article(s) will be loaded, and particularly, which article will appear on top of the page and will therefore be visible “above the fold”².

The interaction technique underlying the BlogUI allows participants to scroll through a list of given articles, more or less one by one. The amount of articles skimmed over varies greatly across participants with some barely scrolling down and others scrolling down to the last element. Hardly ever did a user request a 2nd result page (when more than 10 articles would belong to this category).

This distinct difference in interaction behaviour promotes a more passive consumption of information when using the BlogUI. By offering another article directly following the previous (through scrolling down), the “system makes the decision” for the participant. Considering the lack of overview, as participants cannot perceive which articles are contained within the list (without scrolling to the bottom), and the lack of screen estate, which rarely allows more than 1 or 2 articles to be visible concurrently, the BlogUI encourages participants to simply read the next article.

In contrast, the interaction technique underlying the FMI requires participants to actively select each article they want to read in full. Selecting a contextual element will load the full article in the centre of the UI and load a set of related elements (which replace the previously displayed elements) in the contextual area of the UI. This approach not only provides a better overview over a given set of articles (thus improving awareness and understanding), but its interaction technique might also encourage a more active decision making process.

5.4.3 Visual Interaction through Heatmaps

The physical interaction with a user interface (via a mouse or keyboard) can be regarded as the decision part of sensemaking processes (see a detailed discussion of sensemaking in section 2). Awareness and understanding are crucial stepping stones in leading the user to an interaction decision. This means that the visualization of participants’ perception (using gaze fixations) of the tested UI’s via heatmaps provides a starting point for the interpretation of participants’ interaction decisions.

Our analysis shows that users scan a substantial amount of contextual elements in the FMI before making their navigational decision. The heatmaps in Figure 8 highlight the even distribution of visual attention for both information exploration (left) and information search tasks (right) across all participants.

Heatmaps visualize the spatial distribution of overall fixation counts, through a black to red colour spectrum. The longer the timeframe represented in the heatmap, the more fixations are naturally collected across all areas. The heatmap for information exploration (left) has more intense colouring than the heatmap for information search (right) because exploration tasks lasted approx. 5min, and thus much longer than the tasks for information search (with less than 1min duration in

² The part of the web page “above the fold” is defined by screen resolution and relates to the area visible without scrolling down.

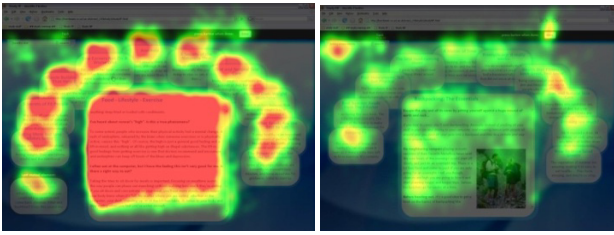


Figure 8. Heatmap across all participants for information exploration (left) and information search (right)

average). The comparison of both heatmaps also shows the increased (visual) relevance of the search functionality (top left corner of the UI) in the information search condition.

Heatmaps for the BlogUI (see Figure 9) show a much more uneven distribution of visual attention for the home page (heatmap no.1), for category pages (no.2, 4, 5, 6) and for search result pages (no. 3) across all participants for information exploration tasks.

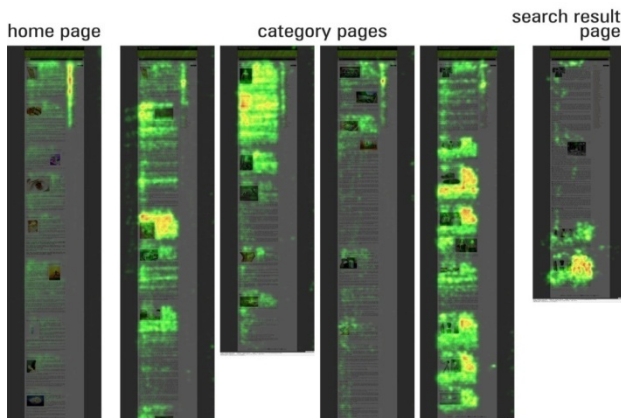


Figure 9. Heatmaps for various pages in the BlogUI (across all participants)

The heatmap for the home page of the BlogUI (left) visualizes the strong focus on the category list. Participants might skim over the first article displayed, but largely ignore the rest of the article list when starting their information exploration.

5.5 Gaze Fixation Analysis

A summative analysis of the eye-tracking data reveals significant differences in how participants use both UIs. The BlogUI collects significantly more gaze fixations in the content section than the FMI and significantly less fixations in the navigation. This finding is valid for exploration tasks (see Figure 11) and for search tasks (see Figure 10).

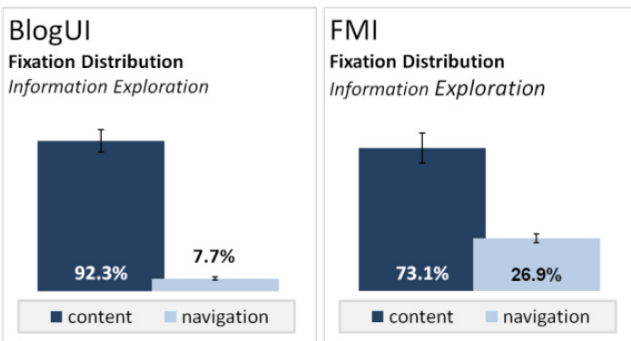


Figure 10. Fixation Analysis for Information Exploration Tasks

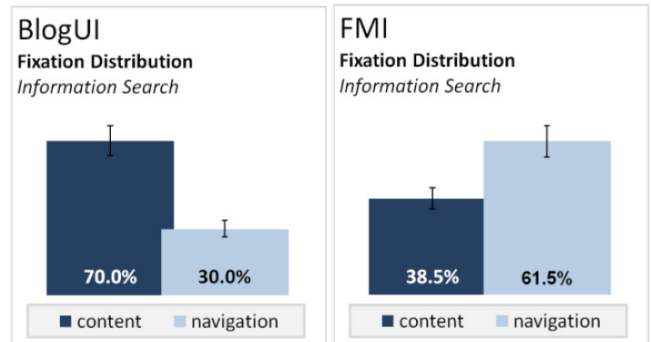


Figure 11. Fixation Analysis for Information Search Tasks

For exploration tasks, both UIs show significantly more fixations in the content than the navigation section. For search tasks, the FMI attracts significantly more fixations in the navigation section. The search tasks did not involve any reading of articles beyond the amount required to judge whether or not a potential article is the “target article”.

The contrast between BlogUI and FMI is rooted in the conceptual differences of how navigation is used. The preview snippets in the navigational elements of the FMI allow people to skim through navigational choices before selecting an article. In contrast, the category links in the BlogUI do not provide any details about the underlying articles. Users are required to skim through a list of the actual articles (after choosing a category) to identify relevant information. Scanning or skim-reading is an important part of “navigational decision making” – particularly when browsing the Web. As a result, a substantial part of collected fixations within the content section of the BlogUI is conceptually used for navigating. The distribution of fixation counts suggests a link to the differences found in interaction behavior between FMI and BlogUI.

5.6 Gaze Time Analysis

The analysis of participants’ attention distribution across individual contextual navigation elements in the FMI shows a quite homogenous distribution for both task conditions, in particular for the exploration tasks (see Figure 12).

For the search tasks, the analysis shows a preference for the first contextual element (which is also the most related one). The trendline for search tasks (with a quite good fit of $R^2=0.74$) indicates a steady but slow decay of attention (negative slope) from the first to the last contextual element.

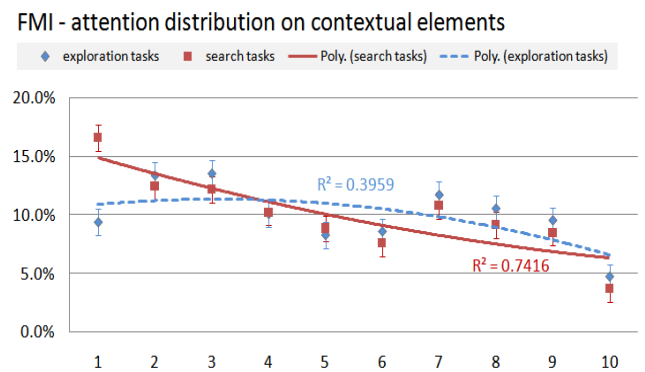


Figure 12: Overall Gaze Time Distributions for Contextual Elements of the FMI for Information Search and Information Exploration

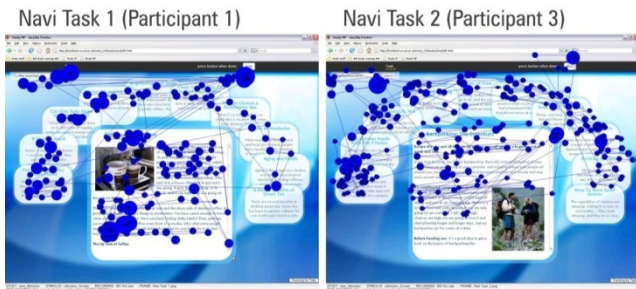


Figure 13. Individual Gaze Plots for Information Search (Task 1, Task2)



Figure 14. Gaze Plots for various pages in the BlogUI

Exemplary gaze plots for specific search tasks using the FMI visualize the attention distributions (see Figure 13) for individual participants. They illustrate the calculated trends across all participants and search tasks (see Figure 12).

Gaze plots for the BlogUI illustrate the much less homogeneous distribution of attention (see Figure 14). Participants' attention moves down the page to process articles in a more linear manner. This interaction requires a lot of scrolling, and can easily lead to "attention gaps", where individual articles are not fixated at all – intentionally or unintentionally. These findings further support the hypothesis that participants explored more of the available information in the FMI than using the BlogUI (H2).

5.7 Average Fixation Duration Analysis

The overall analysis of average fixation durations reveals significant differences between FMI and BlogUI for the exploration condition ($t_{56} = 2.00, p < 0.005$) and for the search condition ($t_{52} = 2.01, p < 0.02$) (see Figure 15).

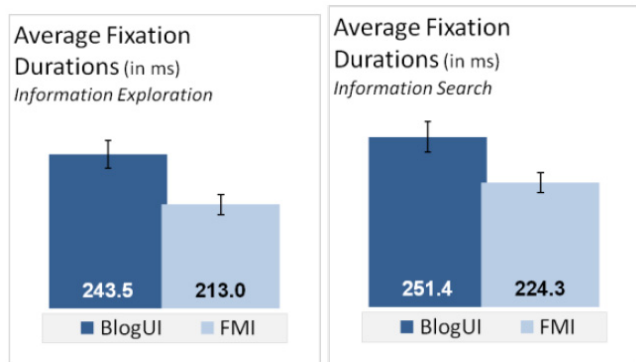


Figure 15: Average fixation durations - BlogUI and FMI for Information Exploration and Information Search condition

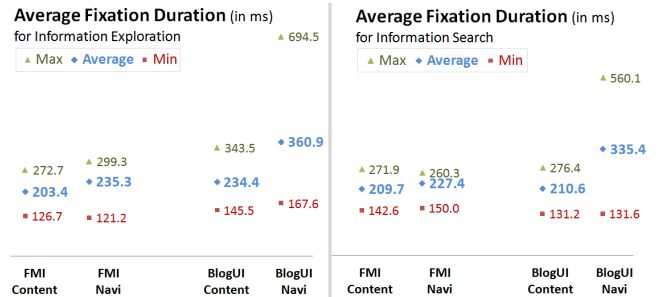


Figure 16. Average fixation duration breakdown for Content and Navigation AOI (for both UI types and both task types)

A further distinction for content and navigation areas of interest (AOIs) in the respective UIs reveals that the navigation AOI is responsible for the increased average fixation durations in the BlogUI (see Figure 16).

Participants' fixations on the category list in the BlogUI last ~48% longer (335.4ms, $SE_x=18.6ms$) than their fixations on contextual elements in the FMI (227.4ms, $SE_x=5.8ms$) for information search tasks and even ~53% longer for information exploration tasks (BlogUI: 360.9ms, $SE_x=21.3ms$ vs. FMI: 235.3ms, $SE_x=7.0ms$). Both differences are significant.

Increased average fixation durations for the search condition, independent of the UI used, are caused by participants' increased attention on navigational elements and the fact that average fixation durations are generally higher for the Navi AOI, than for the Content AOI. The more goal-driven nature of the search tasks required participants to find the right article (making more use of navigational elements), by skimming potentially relevant articles rather than reading them thoroughly. Thus, the BlogUI shows a ~250% increase in attention on navigational elements for the search tasks (when compared to the exploration tasks). The FMI also shows a ~100% increase in attention on navigational elements (see earlier Fixation Count Analysis).

5.8 Questionnaire Findings

Differences measured for task performance and eye movements indicate a clear advantage of FMI over BlogUI. However, to successfully shift users' information experiences, they need to be comfortable with using a novel UI. Other approaches of focus+context visualizations have proven in the past to be superior to traditional UIs but not popular with users.

To build a rich picture of participants' subjective preferences, this study applies a three-part usability questionnaire. We captured participants' first impression of the UIs, and asked them to rate the UI's on usability criteria such as "ease of use", "learnability", and "productivity" (through 24 standard and tailored usability questions; 6-point Likert scale).

The overall rating for the BlogUI is 4.8 ($SE_x=0.05$), and 4.7 ($SE_x=0.05$) for the FMI and thus quite positive for both UIs, considering the baseline being at 3.5 for a 6-point Likert scale. Results in Table 5 show no clear advantage for either of the two UIs. The only significant difference between BlogUI and FMI is for learnability in the search condition, favouring the BlogUI ($t_{57} = 2.00, p < 0.05$).

In light of the very one-sided results of performance, interaction and eye-tracking analysis, the conformity of overall subjective evaluations came at a surprise. Particularly considering that most participants had a rather strong view on which UI they preferred.

Table 5. Usability Questionnaire Results (normalized)

	Information Search		Information Exploration	
	BlogUI	FMI	BlogUI	FMI
Ease of Use	4.7 (SE _x =0.11)	4.7 (SE _x =0.16)	4.8 (SE _x =0.12)	4.8 (SE _x =0.09)
Learnability	5.0 (SE _x =0.10)	4.7 (SE _x =0.09)	5.1 (SE _x =0.12)	5.0 (SE _x =0.08)
UI & Content	4.6 (SE _x =0.09)	4.6 (SE _x =0.13)	4.8 (SE _x =0.12)	4.8 (SE _x =0.09)
Productivity	4.7 (SE _x =0.16)	4.7 (SE _x =0.14)	4.8 (SE _x =0.15)	4.6 (SE _x =0.10)
1st Impression	4.5 (SE _x =0.10)	4.6 (SE _x =0.11)	4.6 (SE _x =0.11)	4.6 (SE _x =0.07)
average	4.7 (SE _x =0.11)	4.7 (SE _x =0.13)	4.8 (SE _x =0.12)	4.7 (SE _x =0.09)

5.8.1 Direct Comparison between BlogUI and FMI

After completing both conditions, we gave participants a final questionnaire to compare both UI types directly. The overall results of the direct comparison show no clear preference for either of the 2 UI Types. However, there seems to be a slight preference for information search tasks independent of the UI used. This could be due to shorter task times, search tasks being more straightforward, or simply the fact that these tasks were tested 2nd to the information exploration tasks. Nevertheless, the experimental design counterbalanced such effects for a comparison between UIs (see Figure 3).

The comparison between FMI and BlogUI across task types reveals a slightly higher rating for the FMI for “ease of use” (53% vs. 47% for BlogUI). Although this difference seems marginal, it should be mentioned that ¼ of participants favouring the BlogUI mentioned “familiarity” as the key reason why they found it “easier to use”.

Familiarity might also explain why 58% of participants rated the BlogUI “easier to navigate” despite clear advantages of the FMI for task completion times and error rates.

Considering the fact, that the BlogUI is a much more traditional type of web-based UI, and none of the participants used the FMI prototype in advance to this study nor were familiar with its concept, familiarity could also play a crucial role in explaining the difference for the learnability measure in the results of the main usability questionnaire.

The balance of evaluations for FMI and BlogUI leads us to reject the hypothesis of a general preference for the FMI (H3). Nevertheless, when considering participants’ contrasting statements, such as “easy to get lost” or “did not feel in control” versus “more intuitive”, “more natural to use” or “very precise” - all on the FMI by different participants, it becomes apparent that strong individual differences exist.

5.9 Impact of Computer Expertise

Other studies in HCI that focus on eye-tracking metrics [5, 7] commonly rely on sample sizes of 15 to 20 participants (or even less). Their participants are often recruited from easily accessible and homogeneous participant groups, such as university students or employees of technology companies (e.g. Microsoft). In addition, these studies often focus entirely on eye-tracking and potentially performance metrics, excluding any subjective evaluations.

With a sample size of 60 participants it seems both feasible and necessary to conduct some post-hoc analysis, considering the inconsistencies between task performance measures and user feedback as discussed in the previous sections.

In a pre-experiment questionnaire, we collected a range of demographics, which allow us to calculate a computer expertise

(CE) score, as explained earlier. To validate the applicability of the CE score, we calculated some related statistics. There are 53% novice and 47% expert participants. The average CE score for participants using the FMI for information search tasks is 0.38 vs. 0.41 for participants using the BlogUI. The average CE score for male is 0.38 vs. 0.41 for female participants.

5.9.1 User Satisfaction by Computer Expertise

Figure 17 visualizes accumulated ratings for the various usability criteria by UI type and computer expertise. Results show, that the homogeneous evaluation observed in the initial analysis of the questionnaire data (see Table 5) is rooted in contrasting ratings between experienced and novice participants. Whereas experts demonstrate a general preference for the BlogUI, novices generally prefer the FMI.

Usability Evaluation

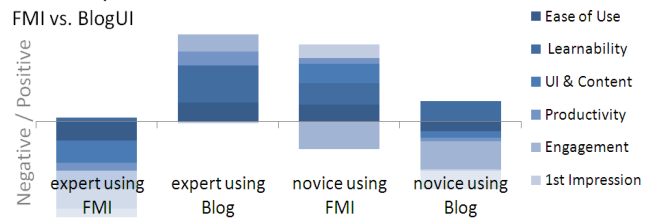


Figure 17. Usability Evaluation by UI type and IT score

In the direct comparison, ~57% of novices found the FMI easiest to use, whereas ~66% of experts found the BlogUI better to navigate.

5.9.2 Other Results

We have conducted a detailed analysis of the impact of demographics on task performance, user interactions and eye-tracking results. Details of those results will be published separately due to the limited space available in this format.

6. DISCUSSION

In this study a traditional blog interface (BlogUI) has been compared to a Focus-Metaphor interface (FMI) with dynamic contextual navigation. Measures for task performance and interaction behaviour were taken for information search and information exploration tasks. Subjective feedback was collected through usability questionnaires to evaluate users’ preferences.

Participants using the FMI completed search tasks faster and made fewer errors (H1 accepted). Moreover, participants interacted more actively with the provided information in the exploration tasks using the FMI (H2 accepted). Overall, no significant differences in user preference could be found between BlogUI and FMI (H3 rejected).

Participants were generally quite engaged in the given tasks, particularly in the information exploration condition, where they were free to read articles of their choosing. General comments on both conditions either favored the exploration condition for its more in-depth information, or favored the search condition for being easier and faster to complete. Participants contrasting views could result from different levels of engagement and interest in the topic. Overall, feedback suggests a slight preference for the UI used during information search tasks.

Higher average fixation durations are commonly linked to increased cognitive load. Our findings suggest that the FMI is generally easier to use for both information exploration and information search. This finding might partly explain the strong advantage found for the FMI in the task performance analysis.

In the search condition, participants particularly used search interaction. Traditional search result pages do not allow the user to conveniently look at all returned results [5]. In contrast, the layout of the FMI enabled participants to effectively scan and compare all returned search results and to efficiently identify the target article or refine their search strategy. Selecting one of the search results automatically adjusts the contextual navigation to provide choices that are more relevant to the selected article. To some extent, this approach realizes the promoted personalized re-ranking of search results, which Teevan et al. (2004) identify as beneficial to cope with users' individual information goals.

Simple search queries and traditional search engines are not sufficient for the challenges information workers face today. They need to make sense of relevant information on a higher level, by 1) acquiring detailed information from a variety of sources, 2) comparing information, 3) evaluating relevance and 4) making decisions based upon the gained insights and previous knowledge. Simple performance measures tested in traditional search experiments do not withstand these complex scenarios. Similarly, traditional search engines do not facilitate complex information work as described above. The concept of how people interact with the contextual navigation of the FMI aims to address these issues by providing means for seamless exploration of information spaces.

However, mixed results for subjective evaluations suggest that the level of dynamic adaption (with each interaction) in the FMI might have gone one step too far. Users unfamiliar with this novel interaction strategy were partly overwhelmed by the constant adaptation of the contextual navigation. Particularly experienced participants who are very familiar with traditional web-usage might have experienced a lack of control over the navigation.

Huge individual variances across measures such as task performance, interaction strategy, etc. have been found in our experiment. Some participants relied entirely on search, others entirely on using categories (or contextual navigation in FMI). Some participants interacted very frequently; others were much more passive and particularly in the BlogUI, would thoroughly go through article lists. The realism and less controlled setup of this study may have facilitated these individual factors to surface.

The particularly strong variances in the task performance for search task 3 indicate the impact that task design has on the actual results of such studies. Task 3 highlights participants' focus on headlines and image information. Many participants would not be able to find the right information, if they cannot deduct the relevance of an article quite easily from the heading and image information only. Information that is not related to headline, image or the first paragraph at least will only be discovered by experienced users or by chance. As blogging involves a lot of scrolling, the eye might coincidentally fixate on the correct paragraph in a longer article. However, this strategy should not be regarded as reliable and satisfying as numerous participants were scrolling over the target paragraphs with fixations in close proximity to the "answer" without noticing it. Scrolling has often been equated with bad usability [15]. Error rates and task performance times for Task 3 show the positive side of scrolling and how particularly experienced users mastered this interaction. Scrolling through long lists of text seems to allow experienced users to process large amounts of information in a very short time. Experienced users were thus able to find less prominently placed bits of relevant information more often and commonly faster than novice users. The post-hoc analysis by gender hints at gender-related influences on

task performance and user satisfaction. It is not within the scope of this thesis to investigate this issue in detail, but the findings suggest that the potential impact of gender differences has to be carefully considered for the design of future studies.

7. CONCLUSIONS

Previous studies investigating the usability and effects of novel user interfaces or user interface features hardly reported gender differences to be a significant factor in the analysis. The problem most eye tracking experiments face in that respect is the limited number of participants involved. Our experiment involved 60 participants – approx. 3 times the sample size of traditional eye-tracking studies (around 15 - 25 participants).

Larger participant samples can have a direct positive impact onto the quality and detail of analyzing eye-tracking data. Summarized reports for measures of average fixation distributions and fixation times blur out potentially crucial individual differences. Our results highlight the need for more sophisticated eye-tracking metrics to evaluate individual differences. Contrasting interaction behaviors may be linked to specific scanning strategies and evaluating user interfaces through gaze plots of a particular participant is not a reliable alternative. Recent results by White & Drucker [19] point at dramatic differences in how individuals use Web-search. Their efforts in developing Web-usage patterns could help to shape related scanning patterns which would greatly benefit higher-level eye tracking analyses.

Evaluating 2 contrasting task types enabled us to gain valuable insights into information exploration and information search strategies. In addition, combining both conditions helped us to understand how users adjust their interaction behavior and scanning strategies from one task type to the other.

8. FUTURE WORK

One of the weaknesses of our study still is simplification of the blogosphere-condition by using a single blog. Having just one blog provides a better-than life case of interaction with information on the blogosphere (A number of participants were very positively surprised by the richness and quality of information provided within our test blog). We are currently working on new studies which will investigate task performance and interaction behavior across realistic blog communities.

Future studies will also need to investigate alternative adaptation techniques for the context navigation of the FMI, e.g. by comparing pull vs. push adaption techniques.

As previous work has shown that people's strategies to manage and access information "may take a long time to evolve" [1], we are currently piloting a large-scale longitudinal online field trial to understand the long-term impact of our tool. Such a field trial will utilize our most recent prototype iteration which provides personalised information retrieval and exploration capabilities to a much larger information space of 100,000+ articles.

9. REFERENCES

- [1] Boardman, R. & Sasse, M.A. (2004): Stuff goes in the Computer but it doesn't come out": A Cross-tool Study of Personal Information Management. In: *Proceedings of CHI 2004*, Vienna, Austria, April 20-24, pp. 583-590
- [2] Brooks, C. H. and Nancy, M. Improved annotation of the blogosphere via autotagging and hierarchical clustering. In *Proceedings of the 15th international conference on World Wide Web (WWW '06)* (Edinburgh, Scotland, 2006). ACM Press, New York, NY, 2006, 625-632.

- [3] Byström, K. and Järvelin, K. Task Complexity affects Information Seeking and Use. In *Information Processing and Management*, Vol. 31-2, 191-213, Elsevier (1995).
- [4] Card, S. K. and Nation, D. Degree-of-Interest Trees: A Component of an Attention-Reactive User Interface. In *Proc. Advanced Visual Interfaces 2002*, ACM Press (2002), Trento, Italy.
- [5] Cutrell, E. and Guan, Z. An eye-tracking study of information usage in Web search: Variations in target position and contextual snippet length. In *Proc. CHI 2007*. ACM Press (2007).
- [6] Furnas, G. W. Generalized Fisheye Views. In *Proc. CHI 1986*, ACM Press (1986), 16-23.
- [7] Hornof, A. J. & Halverson, T. Cognitive Strategies and Eye Movements for Searching Hierarchical Computer Displays. In *Proc. of CHI 2003*, ACM Press (2003).
- [8] Kraft, R. et al. (2006). Searching with Context. In *Proc. of International World Wide Web Conference (WWW '06)*, (Edinburgh, Scotland, 2006). ACM Press.
- [9] Laqua, S. and Brna, P. The Focus-Metaphor Approach: A Novel Concept for the Design of Adaptive and User-Centric Interfaces. In *Proc. Interact 2005*, Springer (2005), 295-308.
- [10] Laqua, S. et al. Contextualizing the Blogosphere: A Comparison of Traditional and Novel User Interfaces for the Web. In *Proc. HCI 2007*, Vol.2, 21st BCS HCI Group Conference (2007).
- [11] Lin, C.C. (2005): Optimal Web site reorganization considering information overload and search depth. In *European Journal of Operational Research*. Elsevier, Article in Press.
- [12] Marsh, R. Similar Posts plugin for Wordpress. <http://rmarsh.com/plugins/similar-posts/>.
- [13] Mishne, G. & de Rijke, M. A Study of Blog Search, In: *Proc. of 28th European Conference on Information Retrieval (ECIR'06)*, 2006.
- [14] Moreville, P. Ambient Findability. *O'Reilly Media*, Sebastopol, USA, 2005
- [15] Nielsen, J. Alertbox: Scrolling and Scrollbars. <http://www.useit.com/alertbox/20050711.html>.
- [16] Norman, D. (1999): Banner Blindness, Human Cognition, and Web Design. In: *Internetworking*, Fourth Issue, 2.1, March 1999.
- [17] Teevan, J. et al. The perfect search engine is not enough: a study of orienteering behavior in directed search. In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '04)* (Vienna, Austria, 2004)
- [18] Teevan, J. et al. Beyond the Commons: Investigating the Value of Personalizing Web Search. In *Proc. of Workshop on New Technologies for Personalized Information Access (PIA)*. (Edinburgh, UK, 2005).
- [19] White, R. W. & Drucker, S.M. Investigating Behavioral Variability in Web Search. In *Proc. of WWW 2007*. ACM Press (2007).
- [20] Zeng, H. J. et al. Learning to Cluster Web Search Results. In *Proceedings of the 27th annual international ACM conference on Research and development in information retrieval (SIGIR '04)* (Sheffield, United Kingdom, 2004). ACM Press, New York, NY, 2004, 210-217.