

A. M. Heinecke, H. Paul (Hrsg.): Mensch & Computer 2006: Mensch und Computer im StrukturWandel.
München, Oldenbourg Verlag, 2006, S. 175-184

Personalised Focus-Metaphor Interfaces: An Eye Tracking Study on User Confusion

Sven Laqua, Gemini Patel, M. Angela Sasse

University College London, Department of Computer Science

Abstract

Personalised web interfaces are expected to improve user interaction with web content. But since the delivery of personalised web content is currently not reliable, a key question is how much users may be confused and slowed down when personalised delivery goes wrong. The aim of the study reported in this paper was to investigate a worst-case scenario of failed personalised content presentation – a dynamic presentation of content where content was dynamically presented, but content units were selected at random. We employed eye-tracking to monitor the differences in users' attention and navigation when interacting with this "dysfunctional" dynamic interface, and a static version. We found that subjects who interacted with the dysfunctional version took 10% longer to read their material than those with static content, and displayed a different strategy in scanning the interface. The relatively small difference in navigation time in first-time viewers of dynamically presented content, and of the results from the eye-tracking patterns, suggests that users are not significantly confused and slowed down by dynamic presentation of content when using a Focus-Metaphor interface.

1 Introduction

The Internet is currently undergoing a major shift towards more service-oriented and integrative architectures with the aim to access, manage, create and communicate information more efficiently. These efforts of integration can be easily recognised in Microsoft's "live" strategy (Windows live), and similar undertakings by Yahoo (My Yahoo) and Google (Google's Personalised Homepage). On the one hand, rich internet applications (RIA) summarise developments towards more complex web based applications that span communication, information, collaboration and content creation in a wide range of ways (e.g. Flickr, Digg, Writely, Skype, Breeze or del.icio.us). On the other hand, huge efforts are made on improving the provision of personalised content (Lin 2005; Novak et al. 2003; Ting et al. 2005) with foci on agent-based systems (Buhler et al. 2003; Maamar et al. 2005), semantic web tech-

nologies (Dill et al. 2003) and quite recently and already very popular – collaborative tagging (Golder & Huberman 2006).

2 Related Work

Much research has been done on providing personalised content using adaptive websites (De Bra et al. 2004). At its core is the creation of sophisticated user models based on clickstream data (Ting et al. 2005), web usage mining (Lin 2005) or similar techniques. But the suggested reorganisation of websites based on these formally developed models has some major drawbacks. One of them is the widely recognised lack of reliable models of information-seeking behaviour (Ting et al. 2005) – information-seeking behaviour varies not just across users, but also within individuals (Boardman & Sasse 2004). Users often have unique aims and objectives and different interaction contexts.

Another problem is the visual component in the interaction. Simply recognising, recording and analysing clickstream data does not answer the key question: How long did it take the user to locate relevant information, and how long did she spend actually reading it?

Simply measuring the time between two clicks does tell us little about what actually is going on inside the user – is she confidently using the website, which would be reflected by a long **study time** (time with attention on content) and a short **scan time** (time to scan navigation and make decision). Or is she confused, with a shorter study time and a longer scan time?

In our study, we used eye-tracking to determine how much time participants spent on navigation, and how much on reading content with static and dynamic content presentation.

2.1 Considering the Interface

The biggest issue for personalised websites is the applied method of visualisation. Scanning a grid or table-based layout spanning two, three or more columns is a task of high cognitive load, and users have developed efficient techniques to facilitate interaction with these websites. It is widely accepted that websites are mainly scanned rather than read and that effects like second-visit blindness (Wirth 2003) and banner blindness (Norman 1999) affect the perception of a website. This is a big thread to personalisation, since most websites make extensive use of navigational elements. The actual content gets more or less visually hidden behind the navigational framework of a website and processes of visual search (Duchowski 2003) become crucial. Often it takes users a substantial amount of time to understand the structure of a website, which spans visual layout, navigation and structure of content.

In this study, the design of the prototype user interface applies the **Focus-Metaphor approach**, which combines aspects of design theory, cognition psychology and educational theory to create a more natural way of interaction (Laqua & Brna 2005).

Especially in personalisable environments, aspects of usability, learnability and satisfaction are crucial as they expect users to engage and immerse themselves to a degree that is far beyond the usual usage of a website.

Approaches to assist users in navigating through large information spaces – and thus making content more accessible – include Degree of Interest (DOI) Trees (Card & Nation 2002) and Fish-eye views (Gutwin 2002). Whereas DOI Trees still come with the problem of separation of navigation and content, which causes disruptions in cognitive processes, Fish-eye view interfaces provide a better focus + context solution but are mainly applicable to a very limited number of scenarios (like map visualisation) in a usable way.

2.2 The Focus-Metaphor

The Focus-Metaphor approach aims to provide a novel framework for building web-based applications and consists of two interdependent parts. Part one addresses the construction of the back-end suggesting a novel unitised structure of ‘detached’ information modules. Each of these modules is hierarchically structured, providing different levels of detail of information at different levels of the hierarchy. Part two targets the front-end counterpart promoting the unitised and hierarchically structured visual interface. This Focus-Metaphor interface displays a contextual subset of the available information modules which themselves act as means for the navigation (see Figure 1 for the prototype interface).

One core strength of the Focus-Metaphor is that it delivers intuitive interfaces (Laqua & Brna 2005) that are very easy to interact with, despite their novelty. An earlier study has shown that, compared with a standard website layout, a Focus-Metaphor interface (FMI) leads to faster navigational decisions and longer “on content” time (Laqua & Brna 2005). In addition to these objective outcomes, the majority of the participants also reported that they preferred the novel interface over the standard layout. It could be argued that ease of use, seamlessness of the interface, consistency and the provision of focus and context are the key reasons for participants’ fondness for the FMI prototypes.

3 The Prototype

In order to evaluate effects of user confusion with a Focus-Metaphor interface (FMI), the prototype interface simulating personalised delivery was constructed for this experiment (see Figure 1). It uses the same visualisation style as in an earlier experiment by Laqua & Brna (2005). The FMI prototype models a simple information space based on textual and figurative elements. It is a high-fidelity prototype in the sense of allowing users to freely interact with all the integrated content. However, the prototype has been designed low-level in terms of integrated functionality, mainly due to its determination of solely being used for the eye-tracking study. To ensure a useful and efficient analysis later on, reading text, watching pictures and navigating through the prototype should be regarded as the highest feasible degree of freedom (based on experience with earlier studies).

The celebrity Michael Jackson was chosen as informational domain to ensure familiarity across participants, and because much content is publicly available. The prototype consists of 35 modules, covering various topics of Michael Jackson’s career. For the purpose of the experiment, two versions of the prototype were developed.



Figure 1: Focus-Metaphor prototype with primary focus and secondary focus modules

The first version dynamically displays content out of the pool of available modules. This version will be referred to as **dynamic version** and is used as treatment group. In order to maximise the measured effect and in order to use the findings of this study for future reference, a randomisation function has been used to determine the dynamically displayed content. This allows to measure the greatest possible impact in terms of user confusion using a Focus-Metaphor interface and to use the results as a benchmark for future dynamic personalisation. In the future, comparisons of the results of this study with those of subsequent studies¹, shall provide insights into the effectiveness of the personalisation.

The second version of the prototype displays static content which does not change during the session of the experiment. This **static version** is used as control group for this experiment. Due to its static nature, this version can only display a subset of the available modules. Consequently, 7 modules have been chosen which reflect a good mixture of the available 35 modules.

4 Eye-Tracking Study

The overall aim of this study was to measure user attention on the two different versions of the FMI prototype. The experiment has been conducted in a usability lab using the Eyegaze system (LC Technologies).

To have comparable results, the same visual layout (factor X) has been used during the experiment and the amount of accessible information has been altered through the randomly selected display of content in the dynamic version (experimental group). By comparing an

¹ Future studies will apply dynamic personalisation based on methods of machine learning and social tagging.

interface that provides randomly personalised content (dynamic version – X[1]) with one that keeps all information static and thus makes the interface predictable (static version – X[2]) the main objective has been to measure the biggest possible **effect of user confusion** through the means of visual attention (behaviour Y).

X[1]: Dynamic experiment version (Focus-Metaphor): animated primary focus module; not animated secondary focus modules centred around the primary focus module; 35 content modules in total randomly loaded into the 7 visual placeholders

X[2]: Static control version (Focus-Metaphor): animated primary focus module; not animated secondary focus modules centred around the primary focus module; 7 content modules in total remaining static within the 7 visual placeholders

As the Focus-Metaphor interface is based on a minimalist visualisation to keep cognitive load low and match the user's short term memory, the random personalisation has been chosen solely for the purpose of this experiment: to simulate a **worst-case scenario of personalisation**. By knowing the effects of this worst-case scenario, we can measure future prototypes that apply more sophisticated methods of personalisation with the benchmarks set by this study on user confusion.

4.1 Demographics

The eye-tracking study used a population sample of 24 participants, 16 male and 8 female, all university students (average age 20.7 years). Participants were from a variety of ethnic backgrounds. The most common first language spoken was English (14) and Gujarati (5). Participants were randomly assigned to either the experimental or the control group. The experimental group (11 participants) used the dynamic version X[1] and the control group (13 participants) used the static version X[2]. To allow fair comparisons none of the participants did have prior knowledge or expectations of the prototypes. This allowed the experiment to be more realistic. However, in preparation for the experiment, participants were given a brief warm-up session to familiarise with the way the prototypes worked and to rule out effects that are due to the novelty of Focus-Metaphor interfaces.

4.2 Procedure

Prior to each session, participants were given a scenario form which briefly described the procedure of the experiment. The studies used a between-subjects design, where each participant conducted one session with either version X[1] or X[2] of the prototype.

The used Eyegaze system does not require any head mounted parts, but in order to increase validity of the data, a chin-rest has been used to minimise head movements and to avoid interruptions in tracking the users' gaze. The experiment was set up such that that the participants were free to explore and interact with the prototype, simulating realistic browsing tasks.

During each session of the experiment, the raw eye-tracking data has been collected, the screen was video-captured and all user interaction was logged within the prototype for later analyses. Each participant interacted with the prototype for approximately 5 minutes, how-

ever if the participant viewed all the information faster, the experiment was terminated earlier. As expected this occurred more often with the static version X[2] as there was less information to read than in the shuffled version X[1]. Nevertheless with $t_{totalX[1]} \sim 274$ sec. and $t_{totalX[2]} \sim 253$ sec. the actual difference of average session duration between the two versions was less than 10%. For the analysis, the raw data has been normalised to cope with this variance.

4.3 Analysis

The main goal of the analysis has been to compare and find differences in user attention between the two different versions X[1] and X[2] over the complete length of sessions. The analysis of user attention includes a heatmap analysis to visualise the general distribution of user attention, and an attention analysis that reports on the effects of user attention on the actual content and navigational elements between static and dynamic version.

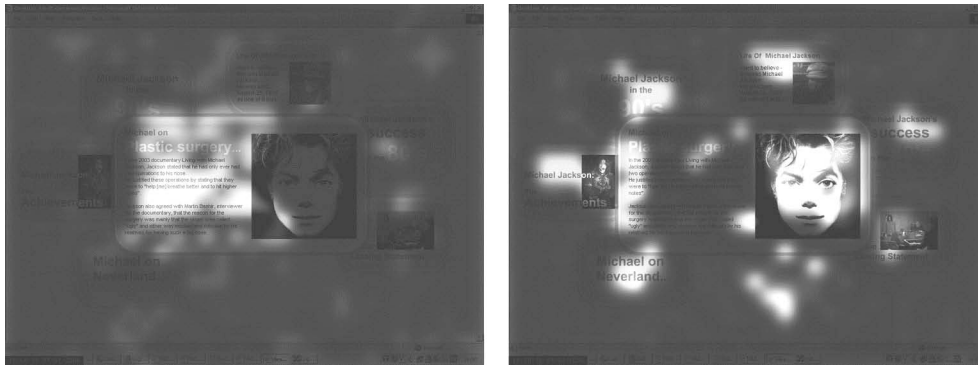


Figure 2: Heatmaps visualising the attention losses on content (left) and attention gains on navigation (right) of the dynamic version X[1] versus the static version X[2]

4.3.1 Heatmap Analysis

Screen-grids have been used to record fixations and to generate heatmaps that visualise the distribution of attention across the interface. In a first step, separate heatmaps for X[1] and X[2] have been created that visualise the mean of overall attention of the experiment group X[1] and the control group X[2]. Although differences between the two versions have already been clearly visible, additional ‘difference heatmaps’ have been generated that directly highlight the attention gains and losses between X[1] and X[2] (see Figure 2). In detail, the left heatmap shows the attention surplus of X[1] (static version) over X[2] (dynamic version) and the right heatmap shows the attention surplus of X[2] over X[1]. This analysis visually confirms that participants of the control group using the static version spent more time on the primary focus module (and thus on the content) than participants of the experiment group using the dynamic version. In contrast, the experiment group spent more time on the secondary foci modules (and thus on the navigation) than the control group.

4.3.2 Attention Analysis

In addition to the screen-grids, regions of interest have been set up to precisely record the fixations on the elements of the interface – mainly the primary focus module (the main content) and the navigation modules (the secondary content). These regions have been used in the attention analysis to compare the dynamic version X[1] with the static version X[2] over the length of the complete session, measuring the overall visual **attention on content (Y_cont)** and the overall visual **attention on navigation (Y_navi)**. This analysis provides insight into how long users' actually spent reading the content (and looking at pictures) indicated through eye fixations on the primary focus module and how much time they required for navigation, which is accumulated by the individual measures of eye fixations for each navigational module.

The left part of Figure 3 shows the amount of time participants spent on content in percent of the overall session time by comparing the geometric means of Y1_cont (for the dynamic version) and Y2_cont (for the static version). As expected, the attention on content analysis shows a significant statistical difference between **Y1_cont = 72.1%** and **Y2_cont = 83.9%** with standard errors for the means of $\sigma_{Y1_cont} = 3.3$ and $\sigma_{Y2_cont} = 1.1$.

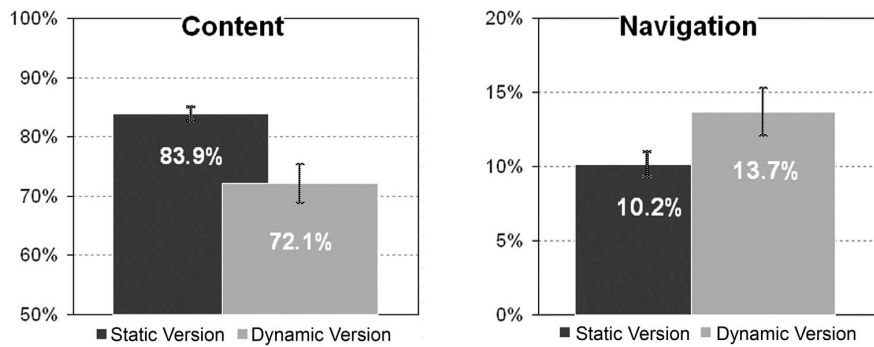


Figure 3: Comparison of user attention on content (left) and on navigation (right)

This result denotes a decrease of time spent on content by 11.8 percentage points when dynamically shuffled content is displayed instead of static content.

The right part of Figure 3 shows the amount of time participants focused their attention onto the navigational modules in percent of the overall session time by comparing the geometric means of Y1_navi (for the dynamic version) and Y2_navi (for the static version). The attention on navigation analysis also shows a significant statistical difference between the two versions with **Y1_navi = 13.7%** and **Y2_navi = 10.2%**. The standard errors for the means are $\sigma_{Y1_navi} = 1.6$ and $\sigma_{Y2_navi} = 0.8$.

The results confirm that participants spent more time navigating (Y1_navi) in the dynamic version X[1] which is an indicator for confusion, higher cognitive load and a longer decision making process to find the required content. This then consequently caused less attention on content (Y1_cont). Table 1 provides a summary of the data.

Table 1: Comparison of user attention in X[1] and X[2] on content and on navigation

	X[2] (static version)			X[1] (dynamic version)		
	in %	σ_M	Conf. Interval	in %	σ_M	Conf. Interval
Content	83.9	1.1	$81.7 \leq \mu \leq 86.1$	72.1	3.3	$65.7 \leq \mu \leq 78.5$
Navigation	10.2	0.8	$8.5 \leq \mu \leq 11.8$	13.7	1.6	$10.5 \leq \mu \leq 16.8$

5 Conclusions & Future Work

The findings of the **attention on content analysis** show a significant difference between the static version and the dynamic version. However, this has not been very surprising as we compared a static version with a dynamic version where content was shuffled randomly. As the aim of this experiment has been to measure the greatest possible effect of user confusion within a Focus-Metaphor interface it is rather surprising that the difference has not been bigger. When taking the results of this study as a benchmark for further work, a worst-case personalisation scenario losing only around 10% of user attention is very promising. Providing means to personalise the displayed information using tags and methods of machine learning is thought to improve the interaction and user experience, potentially raising user attention again to levels close to the static version. The aim of a personalised version is to provide the most relevant contextual information to the user's current task. This obviously implies that task complexity will have an effect on task performance. An investigation of these aspects will be subject of future studies.

The study presented in this paper perpetuates the research and evaluation relating to the Focus-Metaphor approach. When comparing the results of this study with the findings of an earlier study (Laqua and Brna 2005) which investigated differences of attention between a Focus-Metaphor interface (that uses the same visualisation style as in this study) and a standard grid-layout interface, the findings for the static FMI version in the earlier experiment correlate with the findings for X[2] described in this paper. With 82.5% in the earlier study and 83.9% in the current, these results highlight the validity of the results.

Even more interesting in this context is the fact that the randomly dynamic FMI version X[1] in this study with 72.1% is still on the same level as the static grid-layout version of the former study with 70.0%. Although the significance of these cross-comparisons cannot be rendered precisely, it still provides some further insides.

Modularisation of content is one of the key advantages of the Focus-Metaphor approach. By its very nature, this solution is highly scalable, content is very easy to update, extended or also excluded without compromising the integrity of the whole environment. This can be done – like in the current existing prototypes – by content authors, but also in a collaborative way by peers – like intended for future prototypes.

The next step is the development of large scale prototype which will be available online for formative evaluation. This prototype will provide means to communicate, contribute and of course access information. The current plan envisions two separate prototype instances, ap-

plying two different informational domains, each with the aim to address a broad but unique audience and stimulate collaboration. Future studies will target both qualitative and quantitative evaluations in a mainly scenario-based fashion that is close to simulating real-world usage.

Conceptually, the integration of tagging as method for structuring the information space will be the next aspect that we will integrate and evaluate. This will enable us to investigate a growing socially personalised information space. The integration of methods of machine learning to combine human-centred and system-assisted automated adaptation will be the main body of research over the next years.

References

- 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
- Brusilovsky, P. et al. (2005): Workshop on New Technologies for Personalized Information Access. (Eds) In: 10th International Conference on User Modeling (UM'05), Edinburgh.
- Buhler, P. A. et al.(2003): Adaptive Workflow = Web Services + Agents. College of Charleston, South Carolina.
- Card, S. K.; Nation, D. (2002): Degree-of-Interest Trees: A Component of an Attention-Reactive User Interface. Palo Alto Research Center.
- Del.icio.us, <http://del.icio.us/>. Last accessed 26 March 2006.
- De Bra, P. et al. (2004): The Next Big Thing: Adaptive Web-based Systems. In: Journal of Digital Information, Volume 5 Issue 1. Article No. 247, 2004-05-27.
- Dill, S. et al. (2003): SemTag and Seeker: Bootstrapping the Semantic Web via Automated Semantic Annotation. IBM Almaden Research Center. In: WWW 2003, Budapest, Hungary.
- Digg, <http://www.digg.com/>. Last accessed 26 March 2006.
- Duchowski, A. T. (2003): Eye Tracking Methodology: Theory and Practice. Springer-Verlag, London.
- Flickr, <http://www.flickr.com>. Last accessed 26 March 2006.
- Golder, S. A.; Huberman, B. A. (2006): Usage patterns of collaborative tagging systems. In: Journal of Information Science, Vol. 32, No. 2, pp. 198-208, 2006.
- Google Labs (incl. Google's Personalized Homepage), <http://labs.google.com>.
- Gutwin, C. (2002): Improving Focus Targeting in Interactive Fisheye Views. In: Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI'02), Minneapolis, pp. 267-274.
- Laqua, S.; Brna, P. (2005): The Focus-Metaphor Approach: A Novel Concept for the Design of Adaptive and User-Centric Interfaces. In: Interact 2005, LNCS 3585. Rome: Springer-Verlag, pp. 295-308.
- LC Technologies, <http://www.eyegaze.com/>. Last accessed 20 February 2006.
- Lin, C. C. (2005): Optimal Web site reorganization considering information overload and search depth. In: European Journal of Operational Research. Elsevier, Article in Press.

- Maamar, Z. et al. (2005): Toward an Agent-Based and Context-Oriented Approach for Web Services Composition. In: IEEE Transactions on Knowledge and Data Engineering, Vol. 17, No. 5, May 2005.
- Macromedia Breeze, <http://www.macromedia.com/software/breeze/>. Last accessed 12 February 2006.
- Microsoft Windows Live, www.live.com. Last accessed 18 March 2006.
- Mika, P. (2005): Flink: Semantic Web technology for the extraction and analysis of social networks. In: Journal of Web Semantics, Vol. 3, No. 2, 2005.
- Norman, D. (1999): Banner Blindness, Human Cognition, and Web Design. In: Internetworking, Fourth Issue, 2.1, March 1999.
- Novak, J. et al. (2003): Verbindung heterogener Experten-Communities durch die Entdeckung, Visualisierung und Nutzbarmachung von stillem Wissen – das AWAKE Projekt. In: Mensch & Computer 2003: Interaktion in Bewegung. Stuttgart: B.G. Teubner, 2003, pp. 99-112.
- Ting et al. (2005): UBB Mining: Finding Unexpected Browsing Behaviour in Clickstream Data to Improve a Web Site's Design. In: 2005 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2005), France.
- Wirth, T. (2002): Missing Links – Über gutes Webdesign. München: Carl Hanser Verlag.
- Writely, www.writely.com/. Last accessed 26 March 2006.
- Yahoo – My Yahoo, <http://my.yahoo.com/>. Last accessed 16 March 2006.