

# Incorporating Eyetracking into User Studies at Google

Laura Granka, Kerry Rodden

Google

1600 Amphitheatre Parkway

Mountain View, CA 94043, USA

granka@google.com, kerryr@google.com

## ABSTRACT

In this position paper we describe our respective backgrounds, and our initial experiences with incorporating eyetracking into user studies at Google. We also present our views on the principal topic of the workshop: measurement of user satisfaction using eyetracking.

## INTRODUCTION

The authors both work as usability analysts at Google. We have been actively involved in designing and running eyetracking studies, and figuring out how we can make best use of the technology to provide us with reliable insights into our products. Some more details about our backgrounds follow.

**Laura:** I have been using eyetracking for three years in a Web based context. For my Masters research at Cornell, I studied eyetracking and information retrieval, specifically addressing how people read and scan search engine results [2]. I looked at both descriptive measures of viewing behavior (e.g., how many abstracts do users look at before clicking on one, how long do users spend reading/scanning each abstract) as well as understanding the influence of task type and task difficulty on searching behaviors. I also used eyetracking in a more general Web context, looking at how users browse and view the content on commercial home pages. In these latter studies, I looked more at the bottom-up influence of page elements (namely, visual salience) and overall patterns of eye movements across the page.

**Kerry:** My background is in computer science research, and as part of my PhD work at Cambridge University I designed and conducted a series of experiments comparing different ways of arranging thumbnail images on a screen. I was not lucky enough to have access to an eyetracker then, but was able to do some analysis of users' mouse paths in an experiment where these should have corresponded roughly with users' visual scan paths. I have

worked at Google since late 2003, and have been involved in usability testing many different Google products, mostly in the area of search. With regard to eyetracking, I am particularly interested in how users scan search results, and in ways of aggregating data across users, both for statistical analysis and for visualization.

## USE OF EYETRACKING AT GOOGLE

We acquired the Tobii 1750 [5] in mid-2005, and have been using it with the bundled ClearView software for calibration, session recording, data export, and analysis.

Our team focuses primarily on running “discount” usability studies [3] with think-aloud. So far, our main use of eyetracking has been as a qualitative supplement to these studies. We have the equipment set up so that viewers in the observation room can see the user's screen, with a live overlay of their eye movements. This makes a session more compelling to watch, and has become very popular with product teams, making them more likely to come and observe. In addition, the calibration and set-up process is simple enough that it normally adds only about 5 minutes to the study time.

In general, our team members do not feel comfortable combining think-aloud with quantitative measures (such as time-to-task-completion) because the process of thinking out loud introduces bias into the measurements [1]. We believe that this caveat also applies to eyetracking data; for example, a user might pause in the middle of a task in order to explain why they were having a particular problem, and look around the screen far more than they would if simply getting on with their task in silence. So our policy at present is that when combined with think-aloud, eyetracking is used purely qualitatively.

Sometimes, we organize more formal lab experiments, with larger numbers of users, and we ask them to work silently. For these studies, we do record quantitative measures, e.g. comparing time-to-task-completion of two variants of a system, and we conduct statistical analysis of the results. We have used eyetracking data to supplement these measures, e.g. number of fixations during a task or sub-task.

Our biggest challenge so far has been certain limitations of the ClearView software, particularly with regard to data analysis. While it can generate gaze plots from individual

user sessions, and heat maps of aggregate data from several sessions, the software is severely limited with regard to the aggregation options offered. We have also found that, in general, existing eyetracking software lacks specialized features for analysis of studies where web pages are used as the stimuli, e.g., dealing with repeat visits to the same page, or page content that changes dynamically.

### MEASUREMENT OF SATISFACTION

As described in the previous section, we have been using eyetracking data as a supplement to more traditional measures, but only in formal lab experiments, not formative, think-aloud usability studies. So far in these experiments we have largely used eyetracking as a measure of efficiency, nor effectiveness or satisfaction.

In general, our usability team focuses primarily on observing the user's actions during a study, as this helps us to objectively analyze efficiency and effectiveness. We are less interested in measuring satisfaction through subjective user opinion (e.g. of task success, or system acceptability), especially because of the small sample sizes that are typically used in discount usability – we would not be comfortable with the reliability of any resulting statistics.

Because most of Google's products are web-based, we can instead use log data to gather coarse measures of satisfaction on a large scale, e.g. how many users a product has, and how frequently they are using it. It could be argued that, on the web, if users are not satisfied with a site, they will simply go elsewhere. Eyetracking can then be used (either qualitatively or quantitatively) to help us explain phenomena observed in log data.

However, a system must be built and deployed to users before log data becomes available. If we could use eyetracking to infer user satisfaction (e.g. with a task outcome, or preference for a particular system) when using prototypes in lab experiments, this would give us an opportunity to help ensure a satisfying user experience at an earlier stage in the product development process.

### INTERPRETATION OF EYE MOVEMENTS

The majority of existing eyetracking work is primarily cognitive – meaning that we try to use eyetracking to better understand what a user is thinking and what information they are processing. There is hardly any use of eyetracking in “non-cognitive” contexts – using eyetracking to understand what a user is feeling, such as feelings of satisfaction, frustration, or preference. While this type of interpretation would certainly be useful, some preliminary work is necessary to develop a shared understanding of how eye movements could be interpreted in this way.

There are a few initial studies of how eyetracking might relate to emotion, most of them conducted in the field of psychology. The most commonly used measure is pupil dilation, which taps into the degree of a user's interest or arousal in the subject matter. However, accurately interpreting pupil dilation is difficult in Web analyses, as the large variance in page color and brightness confound the data.

While interpretation standards for the meaning of ocular indices are desired, another important issue to consider is that task significantly influences viewing behavior. Task differences most notably affect a user's scan path, but they also lead to different interpretations of certain ocular indices. For instance, fixation duration can take on a different meaning based on whether a user is performing a reading task or visual search task [4]. Furthermore, scan path and the number of fixations in specific areas of interest also differ based on the task a user was given [6].

To expand on this body of research, it would be useful to supplement eyetracking data with other measures, such as clickthrough data and mouse movements, and look for relationships between them. To better understand how eye movements might correlate with user emotion, we could do some studies analyzing eye movements in conjunction with physiological data (such as skin conductance or heart rate change). However, we are doubtful that usability problems would evoke emotions strong enough to be reliably measured, compared to those generated by the more extreme stimuli typically used in psychology experiments.

### ACKNOWLEDGEMENTS

We are grateful to Maria Stone for her comments on an earlier draft of this paper.

### REFERENCES

1. Ericsson, K.A. and Simon, H.A. *Protocol analysis: Verbal reports as data*. MIT Press, 1993.
2. Granka, L., Joachims, T., and Gay, G., Eye-tracking analysis of user behavior in WWW-Search, poster abstract, *Proceedings of ACM SIGIR 2004*.
3. Nielsen, J. *Usability Engineering*. Academic Press, Boston, MA, 1993.
4. Rayner, K. Eye movements in reading and information processing: Twenty years of research. *Psychological Bulletin*, 124: 372-422, 1998.
5. Tobii Technology, <http://www.tobii.se/>
6. Yarbus, A.L. *Eye Movements and Vision*. Plenum Press: New York, 1967.