



Using Gaze Data in Evaluating Interactive Visualizations

Harri Siirtola, Kari-Jouko Rähkä

► To cite this version:

Harri Siirtola, Kari-Jouko Rähkä. Using Gaze Data in Evaluating Interactive Visualizations. 2nd Human-Computer Interaction and Visualization (HCIV) (INTERACT), Aug 2009, Uppsala, Sweden. pp.127-141, 10.1007/978-3-642-19641-6_9 . hal-01572648

HAL Id: hal-01572648

<https://inria.hal.science/hal-01572648>

Submitted on 8 Aug 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Using Gaze Data in Evaluating Interactive Visualizations

Harri Siirtola, Kari-Jouko Rähkä

Visual Interaction Research Group (VIRG)
Tampere Unit for Computer-Human Interaction (TAUCHI)
Department of Computer Sciences
FIN-33014 University of Tampere, Finland
{harri.siirtola, kari-jouko.raihä}@cs.uta.fi

Abstract. Evaluations have long been missing or imperfect in a publication presenting a new visualization technique, but proper evaluations are now becoming a standard. There are many reasons for the reluctance of evaluating visualization techniques, including the complexity of the task and the amount of work required. We propose a simple evaluation approach that consists of a set of tasks carried out in an experimental setting coupled with eye tracking to approximate the focus of the user’s attention. In addition, we discuss three methods to visualize the gaze data to gain insight into the user’s attention distribution, and show examples from a study where a parallel coordinate browser was evaluated.

1 Introduction

Evaluation of interactive visualizations is a challenging task. This is easy to verify by observing the attention it is given in the literature, and the number of workshops, panels, sessions, and keynote speeches dedicated to it. The complexity of evaluation comes from the large number of factors that affect how a particular technique performs. These factors include interplay of the user’s internal models, perceptive processes, interpretations, and cultural elements [1]. It is often difficult to cancel out even some of these factors which leads to hard-to-analyze experimental results.

The lack of evaluation is a serious problem in visualization research. Often the published ideas are visually or technologically interesting, but the lack of evaluation leaves it open if they are actually useful. Ellis and Dix [2] did in 2006 a survey of 65 papers describing a new visualization application or technique, and found that only 12 of them had any kind of evaluation at all, and of those 12 evaluations only two were of any use! This is clearly a problem for the visualization community since development of visualization techniques should be based on experiences from prior work.

The proposed evaluation methods for interactive visualizations vary a great deal. In one end we have recommendations to perform longitudinal and extensive in-depth user studies, and in the other end there is an exasperated demand to

“try it at least by yourself.” Somewhere, between these two extremes, there must be an approach that is both practical and acceptable to the community.

In recent years the eye tracking technology has become affordable and is now considered as standard in usability testing [3]. When the early usability methods used video taping of the computer screen and test participant in analysis, it is now possible to overlay the screen recording with the test participant’s gaze path. This is immensely useful in analyzing interactive visualization techniques as well, because when a person maintains her gaze on a single location for a certain period of time, it can be used as an approximation of the person’s focus of attention [4]. Without this information it is really hard to say if the user’s attention is drawn to the relevant, task-specific parts of the visualization.

The main hindrance in using eye trackers in evaluation is the sheer volume of data they produce. Unless the analyst is willing to watch all the gaze-overlaid videos and transcribe the events, the data must somehow be abstracted, summarized and visualized to allow the analysis. Analyses of usability tests often abstract the data into heatmaps that do reveal the highly useful overview of gaze behavior, but abstract away much of the gaze data. In the evaluation of interactive visualizations a bit more detail is often desired.

In this paper we consider a number of methods to abstract the gaze data for the analysis of interactive visualizations. The following approaches to represent the gaze data are explored:

- Heatmaps of fixations
- Proportion of time spent on Areas-Of-Interest (AOIs)
- Transitions between AOIs

Each of these approaches provide a different view of how a visualization performs. In Section 4 we introduce a parallel coordinate browser and its evaluation data that is used as material in examples.

2 Related work

Saraiya et al. [5,6] discuss the limitations of using a short-term controlled experiment to study interactive visualizations and present an “insight-based longitudinal method” for evaluation. They argue that the evaluation should cover the entire analysis process from a raw data set to the insights sought from the data. In similar vein, Shneiderman and Plaisant [7] propose a method titled “multi-dimensional in-depth long-term case study (MILC)” which is based on a vast array of methods applied in a longitudinal study.

Isenberg et al. [8] propose a “grounded evaluation” method that uses qualitative data to “ground” the evaluation to the correct context of the intended use of the visualization technique. They advocate application of qualitative methods early on and through-out the entire development life cycle, and greater sensitivity to context.

In addition to the MILC approach, the other publications and position statements from the BELIV workshops (“Beyond time and errors: novel evaluation

methods for information visualization”) [9,10] are relevant as well. Andrews [11] gives a clarification of different flavors of evaluation and discusses which method is applicable in which phase of the development cycle. O’Connell and Choong [12] stress that the evaluation should be based on large, realistic data sets and tasks, and advocate the user-centered approach to evaluation. Robertson [13] stresses the importance of ecological validity, both in tasks and in the choice of participants.

Eye tracking has a long history in evaluating eye movements for a variety of visual inspection tasks [14]. The richness of the data brings with it an added challenge: which metrics to use for evaluating the data and the interaction with the application? Jacob and Karn [15] list more than ten possibly useful metrics, each fit for and used in for various purposes in varying contexts. For a researcher mainly interested in the interactive visualization and not on the analysis tool, finding the right analysis metric can be a challenge. Systematic, standardized approaches would be desirable.

In keeping with the tradition of the visualization community, to make sense of the gaze data one should naturally try to find suitable visualizations of it. Ramloll et al. [16] address the challenges and opportunities in designing gaze data visualization tools. Various tools have since been produced by eye tracker vendors and by researchers (e.g., [17]). This paper builds on the existing tools by proposing a method suitable for analyzing interactive visualizations.

3 Evaluation method

The proposed evaluation approach can be seen as an extension of what North [18] suggested in his column about measuring “insight generation.” Our approach has two elements:

1. Artificial tasks that simulate the true tasks;
2. Approximation of visual attention by eye tracking.

North suggested that (1) should include both simple benchmarking tasks and more complex, open-ended tasks. The former allow us to check the performance of low-level mechanics of the visualization, and the latter bring at least some ecological validity to the evaluation. The open-ended tasks are highly useful as post-test interview material as well. The “artificiality” of tasks means that they can be real tasks, if possible, but often we need to resort to something that is only mechanically close to the real ones.

The second element (2) proposes that we use eye tracking data to inspect where the user is looking at while using a visualization. An eye tracking device captures participants’ gaze data, especially their fixation targets and lengths. A *fixation* occurs when the person maintains his or her gaze on a single location for a certain period of time, and it can be used as an approximation of the person’s focus of attention. The rapid transition between fixations is called a *saccade*, and no visual percepts accumulate during the saccades.

These elements are combined in a controlled experiment that provides both quantitative (task execution time, correctness, gaze distribution) and qualitative (observation, think-aloud protocol, interview) data about the visualization. While this data may often be quite noisy, it is highly useful when comparing two designs or when trying to understand why a certain visualization performs as it does.

We believe that the “light-weight” or “discount” evaluation method proposed here might be a good tradeoff which reveals enough about the evaluation target to draw informed conclusions. Depending on whether the main interest is in quantitative or qualitative information, the elements (such as think-aloud and eye tracking) may be combined in different ways [19], lending flexibility to the approach.

4 Evaluating Parallel Coordinate Explorer

Evaluation of *Parallel Coordinate Explorer (PCE)* [20,21] is used as an example to illustrate how gaze data helps in studying interactive visualizations. This particular implementation was chosen because there are several studies of it, its strengths and weaknesses are well known, and we recently redid one of the studies by adding gaze tracking to the experimental setup [22]. The implementation of PCE is available on the web [23] as well.

PCE is a parallel coordinate browser that allows users to interact with a parallel coordinate plot by creating and modifying persistent selections. These selections highlight the lines that fulfill the constraints that can be combined with logical connectives (AND, OR, XOR).

Figure 1 shows a typical interaction with PCE. Suppose the task is to find the most powerful four-cylinder car in the data set (i.e., its engine having the highest horsepower value). The four-cylinder cars have been highlighted (constraint on axis CYL), another constraint on axis HP focuses the selection on the cars with high engine horsepower, and one of the cars in the selected set has been drilled down (the values are displayed under the axis bottom labels in red). The end-points of an axis show normally the minimum and maximum values of the corresponding variable, but change to display in red the minimum and maximum of a constraint if there is exactly one range selection on the axis.

The details of the experimental setup and the results are presented in Siirtola et al. [22] and only the nine tasks are given here:

0. How many American cars are there in the data set?
1. How many cars have a four or six cylinder engine?
2. What is the average mileage for the six-cylinder cars?
3. How would you describe the cars that weigh over 4500 pounds?
4. What is the origin of the six-cylinder cars that were manufactured in 1971?
5. Which Japanese cars have the best acceleration?
6. What else is common to the most powerful, best accelerating, and heaviest cars in the data set?

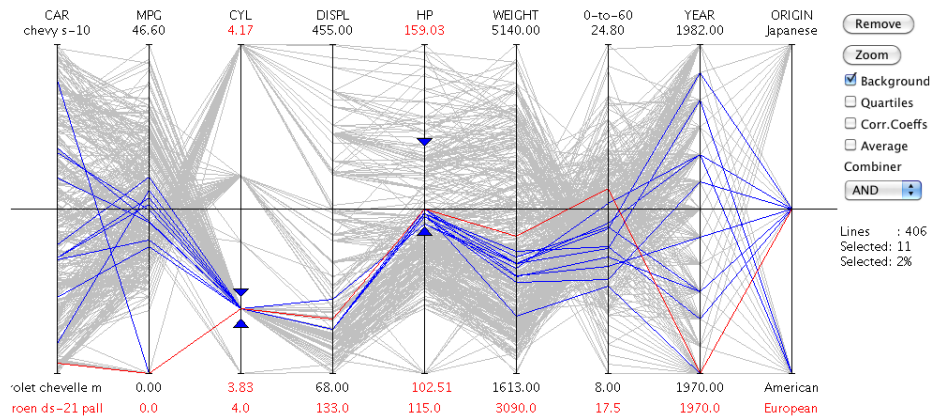


Fig. 1. Interacting with Parallel Coordinate Explorer: finding the most powerful engine among the four-cylinder cars.

7. How many non-American cars are there in the data set?
8. What is the most common number of cylinders for cars manufactured in 1973?

There are three distinct task types: simple selections, complex selections, and explorative tasks. All tasks except 3 and 6 had an unambiguous answer. Task 0 was always performed first as a practice task but the participants did not know about this.

Here we will focus on the approximation of visual attention by visualizing gaze data. The other elements of evaluation mentioned in Section 3 (task execution measurements, live observations, think-aloud, and interviews) are similar as in any usability test [24], and an example of applying them in the evaluation of an interactive visualization can be seen, e.g., in Siirtola et al. [25].

The following sections will show what kind of observations about an interactive visualization tool can be made by abstracting and visualizing the fixations and saccades of evaluation participants. In the following, the gaze data is represented as heatmaps of fixations, as balloon plots of attention proportions, and as gaze transitions between areas of interest.

5 Heatmap-based observations

The most popular approach to visualize fixation data is to ignore the order of the fixations completely and to base the analysis on how long the gaze has been fixated on different areas of the screen. Here “long” can mean either number of fixations or total duration of fixations in the area. Long fixations are usually an indication of increased cognitive processing, and a high number of fixations suggests problems in the visual search. By smoothing the data appropriately and encoding it with different levels of gray or color, the interest or attention can

be visualized in a style familiar from maps. Such visualizations have been called “attentional landscapes” [26] and “fixation maps” [27]. The most common term for this concept nowadays is “heatmaps” or even “attention heatmaps”.

Figure 2 shows the heatmap of task 4 for a single participant where the objective was to find the origin of six-cylinder cars manufactured in 1971. Solving the task requires making a selection with two constraints, one on the CYL axis that selects the six-cylinder cars, and another on the YEAR axis that further prunes the selection to cars that were manufactured in 1971. With this selection it is easy to verify that those cars were manufactured in the U.S.A. In (the color version of) the figure the number of fixations is encoded ranging from green to red via yellow, green being the lowest number of fixations observed and red being the highest.



Fig. 2. Heatmap of fixation counts for task 4 for a single participant (“What is the origin of the six-cylinder cars that were manufactured in 1971?”).

There are two red “peaks” in this landscape in the bottom of axes YEAR and ORIGIN (appearing darker in grayscale). One might expect that there should be three peaks, one for each selection and one for reading the result. It appears that this particular participant had problems making the selection for year 1971 and also spent a relatively long time verifying the result. Making the selection from the YEAR axis is more difficult than making the selection from the CYL axis because of the higher number of items which translates into lesser space to mouse the selection.

Figure 3 shows the heatmap of task 4 for all participants. The overall situation appears more like expected – there are three distinct peaks for the task-critical



Fig. 3. Heatmap of fixation counts for task 4 over all users.

areas, and the trickiest selection has received most of the attention. In addition, it seems that the participant of Figure 2 remembered the location of axes CYL and ORIGIN, but some of the participants had to scan the axes names to locate them (indicated by the light green color over the axis names).

Heatmaps are good for summarizing large quantities of data that would be next to impossible to gain insight into if presented numerically. Sometimes, as in the example above, heatmaps may give useful insight into the distribution of users’ attention, especially if the task at hand is focused and simple enough.

Heatmaps are typically used in single-condition situations, or in multi-condition experiments without tests for statistical significance. Therefore heatmaps should primarily be used for data visualization instead of data analysis [28]. In addition, the heatmaps often do not come with all the relevant metadata, and being unaware of it can lead into false interpretations. Another issue is the attractiveness and apparent intuitiveness of heatmaps which can lead into over-interpretation.

6 Proportion of time spent on Areas-of-Interest

The heatmap-based analysis of fixations can be made more detailed if we divide the underlying visualization into “Areas of Interest (AOI)”, basically by naming the parts of the visualization that are of interest. In many visualizations we have at least some static elements in the interface that allow to track when the user’s gaze is fixated into them. If the visualization does not have any static elements at all, it is still possible to compute the fixations for dynamic elements, but it will be a technical challenge. In PCE, the only dynamic parts of the visualization

are the selection constraints, making it easy define the AOIs. Figure 4 shows the AOI definitions that were used in the experiment [22].

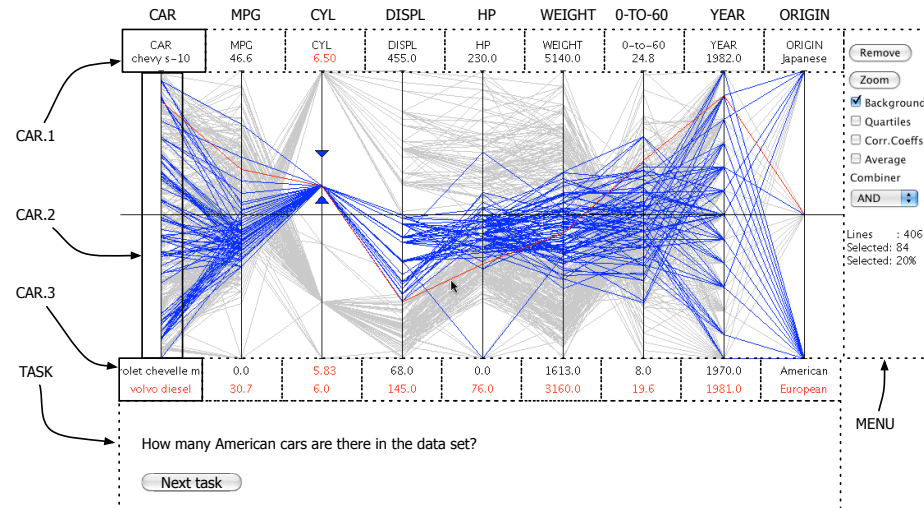


Fig. 4. Areas Of Interest (AOI) used in the analysis and their naming conventions.

Each area is named in Figure 4 after the axis and divided into three parts: top of the axis (1), axis itself (2), and bottom of the axis (3). For instance, for the first vertical axis from the left (CAR), CAR.1 refers to the area enclosing axis label and the maximum value, CAR.2 to the axis itself, and CAR.3 to the minimum value and the selected value, if one polyline is chosen. Figure 4 shows in addition a visual query where all the six-cylinder cars have been highlighted (triangles at the CYL axis) and the one with the smallest engine displacement has been selected (a Volvo diesel, values shown in red below the labels of the axes).

Dividing the user interface of PCE into AOIs allows to list the areas that must be attended to solve a task. For example, solving task 0 (“How many American cars are there in the data set?”) requires looking into the areas TASK (to read the task), ORIGIN.2 and ORIGIN.3 (to find and select “American” cars), and MENU (to read the number of selected cars). This is the optimal course of events which is rare in a real first-use situation. However, the comparison makes it easier to evaluate the design and implementation of a visualization by revealing which proportion of fixations is spent on the areas that are relevant in carrying out the task.

Figure 5 shows two balloon plots of the fixation data (produced with the package *gplots* of statistical system R [29]). In a balloon plot each cell contains a dot whose size reflects the relative magnitude of the corresponding component. The left-hand side plot shows the medians of fixations before entering an AOI,

and the right-hand side plot shows the fixation duration means in an AOI. On the left, the grey background in a cell indicates an AOI that must be first attended to solve a task, and on the right the grey background shows all the cells that must be attended in a course of the task. Therefore, to approach the optimal solution, on the left a smaller value on grey background is “better” and on the right a bigger value on grey background is “better”.

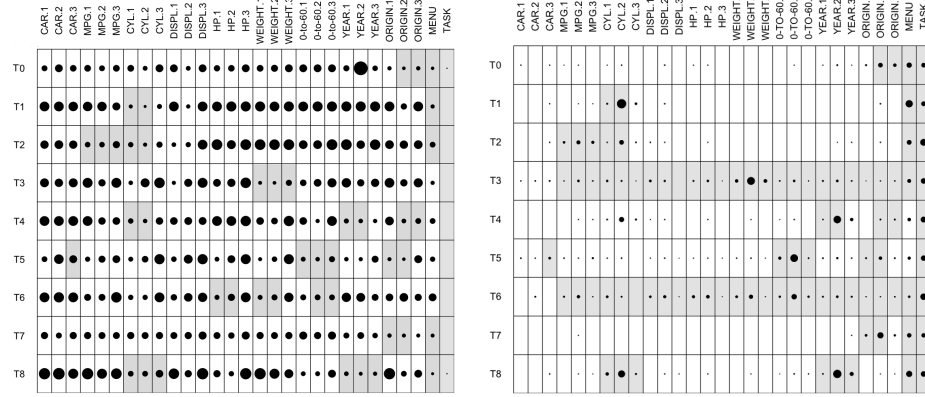


Fig. 5. On the left: balloon plot for AOI by Task: Medians of fixations before entering an AOI. On the right: Balloon plot for AOI by Task: Fixation duration means in an AOI. The mandatory AOIs for a task have a grey background.

In the experiment [22], the percentage of time spent on task-specific AOIs varied between 53% and 92%. The lowest percentages were observed on two open-ended tasks (3 and 6, 53% and 58%, respectively) and on the task with a complicated selection (task 4, 55%). The left balloon plot in Figure 5 shows the medians of fixations before entering the respective AOI and indicate that the important areas were found quickly (the balloons with grey background are small). The right side of Figure 5 plots the duration means in an AOI, and shows that the middle part of the axes drew longer fixations than either the top or bottom area. This is caused by the fact that the interactive part in the middle required more intense attention than the (mostly) static values at the ends.

7 Transitions between Areas-of-Interest

Both the heatmap and the AOI-based visualizations of fixation data ignore the order of fixations completely, which may leave important aspects of interaction unnoticed or even skew the observations. In this third method to visualize fixation data the focus is on the *transitions* between the Areas-of-Interest. This is a piecewise representation of gaze paths and does not show them as continuous objects as in a video overlay. Instead, the idea is that the relevant parts of

the interaction should “pop out” as having more inter-area transitions, providing better insight to the course of events.

The following visualizations were produced by extracting the AOI transitions from eye tracker log files with an AWK [30] script that transformed the AOI data into a transition graph defined in the language DOT (part of the Graphviz project [31]). The DOT compiler was then used to lay out and produce an image of the transition graph.

The initial, plain versions of the graphs were hard to read because of the automatic layout. It completely broke the resemblance to the user interface, trashing the mental model for the UI, and made it hard to interpret the transitions. To force the axis AOIs to stay together, they were placed in the same subgraph. Now the Graphviz layout engine may move the subgraphs around, but not the nodes inside them. Figure 6 shows a transition graph for the AOIs shown in Figure 4 with its static elements, without any transitions.

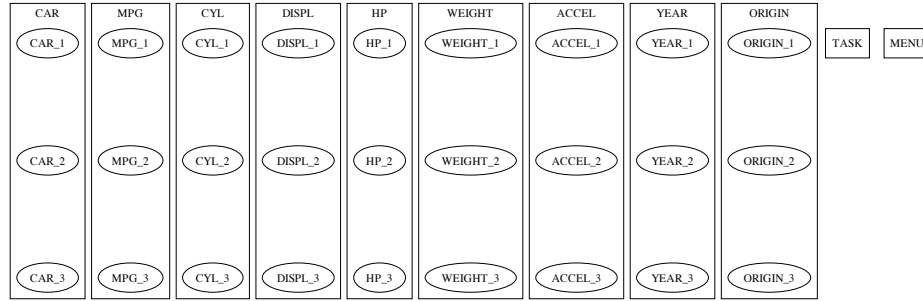


Fig. 6. Area-of-Interest elements in a transition graph. The subgraph structure constraints the nodes inside an axis area to stay together and in the correct order.

As seen in Figure 6, the period in axis names is replaced by underscore to allow indexing of associative arrays with the AOI names. Another change related to the AOIs is that the gaps between the axis AOI’s (as defined in Figure 4) were removed to prevent frequent transitions to and from the AOIs to areas that do not belong to any AOI.

To reduce visual clutter in the transition graphs two additional encodings were introduced. Repeated transitions between two AOIs are represented as a single connection line and the number of transitions is encoded into the line width. In addition, the overall fixation time of an AOI is encoded into the line width of a node symbol. A dashed line is used to indicate a single transition and an AOI without any fixations.

Figure 7 shows a single user’s transition graph for task 4. The focal point of attention is in the middle part of the YEAR axis, YEAR_2, which has the highest duration of fixations as indicated by the bold border of the node. As noted previously, it is a challenge to select one of the twelve items from a crowded

YEAR axis. A similar selection from the five-value CYL axis is easier, and this participant does not even bother to check the lower limit of selection at all (no fixations in CYL_3).

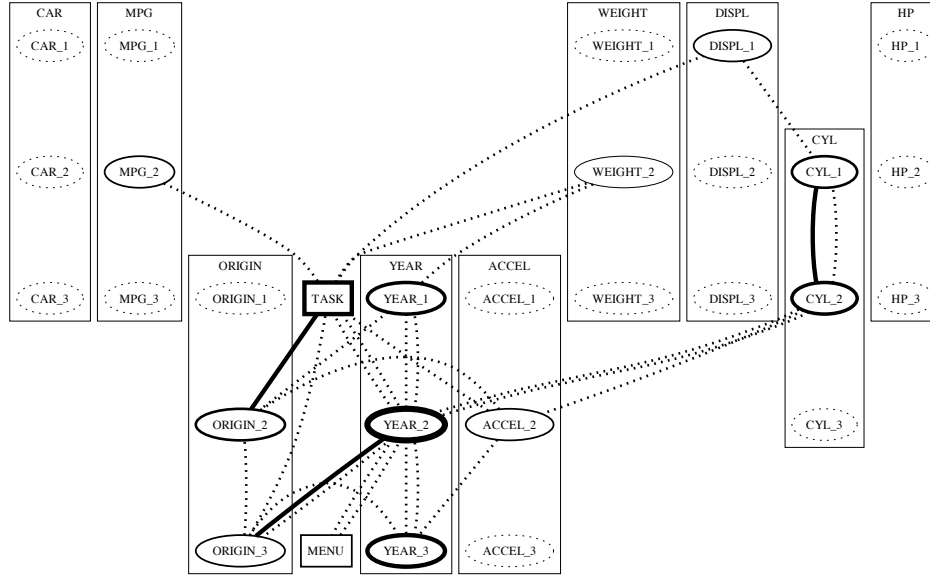


Fig. 7. Transition graph for a single user's task 4, "What is the origin of the six-cylinder cars that were manufactured in 1971?"

The inter-AOI transitions are mainly between ORIGIN, YEAR, and TASK. This participant chose first the six-cylinder cars, then the year 1971, and finally read the selection from the ORIGIN axis. There are several confirmatory YEAR-ORIGIN and ORIGIN-TASK transitions. Overall, the attention is where it is supposed to be when solving this task.

Figure 8 shows a transition graph of task 4 for all participants. Overall, the selection problem on the YEAR axis receives the most of the attention, but beyond that the situation is quite even. However, it is clear that the axes CYL, YEAR, and ORIGIN are in the focus as the layout sets them next to each other.

Figure 9 shows a typical single participant's transition graph for an exploratory task, which is here task 3: "How would you describe the cars that weigh over 4500 pounds?" The strategy is to select the cars having WEIGHT over 4500 and then scan the other axes in search of patterns. AOIs WEIGHT and TASK receive most of the attention and transitions. The middle parts of the axes are then attended and occasional drill-down hits the lower part of the axes. No attention is devoted to the axis labels, indicating that this participant knows the variable positions by now.

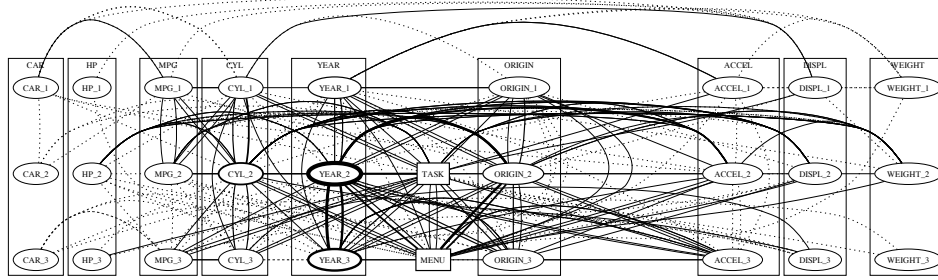


Fig. 8. Transition graph for task 4 over all participants.

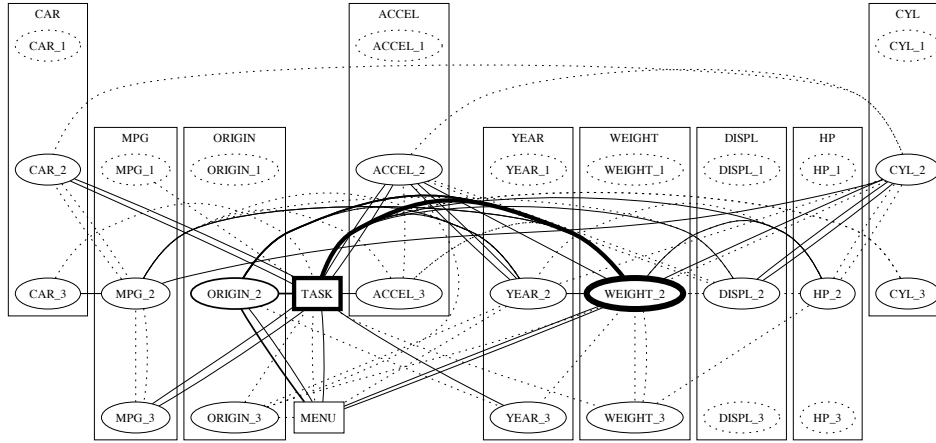


Fig. 9. Transition graph for task 3: “How would you describe the cars that weigh over 4500 pounds?”

8 Discussion

The proposed light-weight evaluation method is proposed as a trade-off between complex methods involving an array of qualitative and quantitative methods applied in a longitudinal study, and simple controlled experiments. The goal is to separate data collection and data interpretation as much as possible, and make the evaluation easier to repeat. When the AOIs are documented and the percentages are printed, the experiment is fairly easy to repeat and verify.

Instead of studying the participants’ focus of attention by observation and interviewing, we advocate to use the eye tracking technology for it. What is really needed is “brain tracking”, but eye tracking is a substitute while this technology matures. The problem with eye tracking is the amount of data collected even for a modest experimental setting. We discussed three approaches that avoid viewing all the video footage and provide useful visualizations of gaze data instead: heatmaps, AOI-based balloon plots, and AOI-based transition diagrams. These correspond roughly to three more general problem solving approaches: acquir-

ing general overview, working top-down, and working bottom-up. Heatmaps are good for giving an overview of how an interface is really used, but on a high level only. The AOI-based balloon plots give a bit more fine-grained insight into how an interactive visualization is used, and the AOI-based transition graphs reveal the most common paths taken while using the visualization.

The proposed approach has the advantage that it is repeatable and, with some practice, can become a technique that is fast for a researcher to use. Eye tracking is, however, no panacea. There are caveats that a researcher should be aware of, such as the inherent inaccuracy of gaze data (the combined effect of tracker inaccuracy and the perceptual span of the human eye can yield offsets of up to 2 degrees of visual angle) and the sensitivity of the results on the algorithm used for estimating fixations [32]. The inaccuracy makes it difficult to apply eye tracking for visualizations that have densely packed information, such as those in Figure 5, though even there focusing on the areas of interest (the grey areas) can make the approach usable. And in many other cases the areas of interest are more sparsely located, such as in our example, or in many timeline visualization.

The proposed evaluation approach is simple and not a substitute to a longitudinal study with real users and tasks. However, we believe that the approximation of users' focus of attention with gaze data analysis is highly useful in revealing the issues that should certainly be addressed before a proposed visualization method is acceptable. The gaze data cannot prove that a certain method is a success, but it is impeccable in pointing out the major design flaws.

In this paper we have applied the proposed evaluation approach into one visualization method and repeated an earlier experiment. In the future, we plan to refine the evaluation method further, do additional experiments, and develop a software tool to streamline the gaze data visualization. In such a tool we could link the two gaze data visualizations which would allow rapid movement between the two views, i.e., spotting an interesting cell in a balloon plot would be a link into the corresponding transition graph and vice versa. One should be able to select and unselect included participants and tasks, and adjust thresholds for the diagrams. This clearly calls for a conventional multiple-view interface with dynamic controls and brushing.

Acknowledgements

We thank Tuuli Laivo for collecting the gaze data that was used in our example visualizations. This work was supported by the Academy of Finland (project 1129300). The work of the second author was supported by the Academy of Finland grant 1130044. The paper was written while he was visiting the University of Canterbury in Christchurch, New Zealand.

References

1. Spence, R.: Information Visualization – Design for Interaction. Prentice-Hall Europe, Pearson Education Ltd., Harlow, England (2007)

2. Ellis, G., Dix, A.: An explorative analysis of user evaluation studies in information visualisation. In: BELIV'06 - Beyond time and errors: novel evaluation methods for information visualisation (AVI'06 Workshop), ACM Press (2006)
3. Nielsen, J., Pernice, K.: Eyetracking Web Usability. New Riders Press (2009)
4. Posner, M.I.: Orienting of attention. *Quarterly Journal of Experimental Psychology* **32** (1980) 3–25
5. Saraiya, P., North, C., Duca, K.: An insight-based methodology for evaluating bioinformatics visualizations. *IEEE Transactions on Visualization and Computer Graphics* **11**(4) (2005) 443–456
6. Saraiya, P., North, C., Lam, V., Duca, K.A.: An insight-based longitudinal study of visual analytics. *IEEE Transactions on Visualization and Computer Graphics* **12**(6) (2006) 1511–1522
7. Shneiderman, B., Plaisant, C.: Strategies for evaluating information visualization tools: multi-dimensional in-depth long-term case studies. In: BELIV '06: Proceedings of the 2006 AVI workshop on BEyond time and errors, New York, NY, USA, ACM (2006) 1–7
8. Isenberg, P., Zuk, T., Collins, C., Carpendale, S.: Grounded evaluation of information visualizations. In: BELIV '08: Proceedings of the 2008 conference on BEyond time and errors, New York, NY, USA, ACM (2008) 1–8
9. Bertini, E., Plaisant, C., Santucci, G.: Beliv'06: beyond time and errors; novel evaluation methods for information visualization. *interactions* **14**(3) (2007) 59–60
10. Bertini, E., Perer, A., Plaisant, C., Santucci, G.: Beliv'08: Beyond time and errors: novel evaluation methods for information visualization. In: CHI '08: CHI '08 extended abstracts on Human factors in computing systems, New York, NY, USA, ACM (2008) 3913–3916
11. Andrews, K.: Evaluation comes in many guises. BELIV'08: BEyond time and errors: novel evaluation methods for Information Visualization (revised position paper; CHI'08 workshop) (April 2008)
12. O'Connell, T.A., Choong, Y.Y.: User-centered evaluation methodology for interactive visualizations. BELIV'08: BEyond time and errors: novel evaluation methods for Information Visualization (position paper; CHI'08 workshop) (April 2008)
13. Robertson, G.G.: Beyond time and errors – position statement. BELIV'08: BEyond time and errors: novel evaluation methods for Information Visualization (position paper; CHI'08 workshop) (April 2008)
14. Goldberg, J.H., Kotval, X.P.: Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics* **24** (1999) 631–645
15. Jacob, R.J.K., Karn, K.S.: Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In Hyona, J., Radach, R., Deubel, H., eds.: *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*. Elsevier Science (2003) 573–605
16. Ramloll, R., Trepagnier, C., Sebrechts, M., Beedasy, J.: Gaze data visualization tools: Opportunities and challenges. *Information Visualisation, International Conference on* (2004) 173–180
17. Špakov, O.: iComponent - Device-Independent Platform for Analyzing Eye Movement Data and Developing Eye-based Applications. *Dissertations in Interactive Technology*, Number 9, University of Tampere (2008)
18. North, C.: Visualization Viewpoints: Toward measuring visualization insight. *IEEE Computer Graphics and Applications* **26**(3) (2006) 6–9

19. Hyrskykari, A., Ovaska, S., Majaranta, P., Rähkä, K.J., Lehtinen, M.: Gaze path stimulation in retrospective think-aloud. *Journal of Eye-Movement Research* **2**(4) (2008) 1–18
20. Siirtola, H.: Direct manipulation of parallel coordinates. In: CHI '00: CHI '00 extended abstracts on Human factors in computing systems, New York, NY, USA, ACM (2000) 119–120
21. Siirtola, H.: Direct manipulation of parallel coordinates. In: IV'2000: Proceedings of the International Conference on Information Visualization, IEEE Computer Society (2000) 373–378
22. Siirtola, H., Laivo, T., Heimonen, T., Rähkä, K.J.: Visual perception of parallel coordinate visualizations. In: IV09: Proceedings of the Thirteenth International Conference on Information Visualization, IEEE Computer Society Press (2009) 3–9
23. Siirtola, H.: Parallel Coordinate Explorer. <http://www.cs.uta.fi/~hs/pce/> (2000)
24. Nielsen, J.: Usability Engineering. Academic Press, Inc. (1993)
25. Siirtola, H., Rähkä, K.J.: Interacting with parallel coordinates. *Interacting with Computers* **18**(6) (2006) 1278–1309
26. Pomplun, M., Ritter, H., Velichkovsky, B.: Disambiguating complex visual information: Towards communication of personal views of a scene. *Perception* (25) (1996) 931–948
27. Wooding, D.S.: Fixation maps: quantifying eye-movement traces. In: ETRA '02: Proceedings of the 2002 symposium on Eye tracking research & applications, ACM (2002) 31–36
28. Bojko, A.: Informative or misleading? Heatmaps deconstructed. In: Human-Computer Interaction. New Trends. Volume 5610 of Lecture Notes in Computer Science. Springer (2009) 30–39
29. R Development Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. (2010) ISBN 3-900051-07-0.
30. Aho, A.V., Kernighan, B.W., Weinberger, P.J.: The AWK programming language. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA (1987)
31. Graphviz: Graphviz – graph visualization software. <http://www.graphviz.org/> (2008)
32. Salvucci, D.D., Goldberg, J.H.: Identifying fixations and saccades in eye-tracking protocols. In: ETRA '00: Proceedings of the 2000 symposium on Eye tracking research & applications, ACM (2000) 71–78