

User-Centered Evaluation of Information Visualization Techniques: Issues and Perspectives

Carla M.D.S. Freitas¹, Marcelo S. Pimenta¹ and Dominique Scapin²

¹ Universidade Federal do Rio Grande do Sul (UFRGS) – Instituto de Informática
Porto Alegre, RS, Brasil, 91.501-970

² INRIA Rocquencourt

{mpimenta, carla}@inf.ufrgs.br, Dominique.Scapin@inria.fr

***Abstract.** This paper discusses some important issues related with usability evaluation of information visualization techniques, particularly using a user-centered point of view. This way, we address one of the goals established in the EDGE (“Evaluation methods, Design Guidelines and Environments for Virtual Reality and Information Visualization Techniques”) project, namely developing new methods for evaluating information visualization techniques, taking advantage of the experience of both teams in this matter.*

1. Introduction

Following the same path of desktop graphical user interfaces (GUI) and Web-based interfaces, a large use of information visualization techniques will depend on their usability. Whereas the first information visualization (IV) techniques were presented without thorough evaluation studies, researchers have become aware of the importance of such usability studies [Plaisant 2004]. However, despite an evident progress in recent years to establish some ‘good practices’ for design and usability evaluation of such techniques, the definition of several aspects related to an user-centered perspective for IV techniques remains an open issue.

Historically, a user-centered perspective has its origin related to User Centered Design (UCD), which involves basically simplifying the structure of tasks, making things visible, getting the mapping right, exploiting the powers of constraint, and designing for error. Thus, a user-centered perspective tries to optimize the user interface around how people can, want, or need to work, rather than forcing the users to change how they work to accommodate the system or function. In short, evaluation with a user-centered perspective is an evaluation based on the needs of the user and, for that, we need to know them, their goals and tasks.

In the last years, we have been particularly interested in evaluation of IV techniques. Although there is a great variety of models and techniques for information visualization, there is not yet a consensus about their usability evaluation: what is the meaning of usability for IV techniques? Which characteristics we have to evaluate and how to evaluate them? In addition, which aspects are generic to all types of interactive systems and which are specific to IV techniques? Which of these issues are particularly relevant to IV techniques’ evaluation?

Our goal in the EDGE (Evaluation methods, Design Guidelines and Environments for Virtual Reality and Information Visualization Techniques) project is to investigate if and how sound ergonomic user-centered knowledge can be transferred to IV techniques' context. This paper intends to start a discussion and presents some issues concerning IV techniques evaluation we are working on based on past experience [Scapin and Bastien 1997][Freitas et al. 2002].

2. Current Studies in Evaluation of Information Visualization Techniques

The different aspects related to the evaluation of IV techniques have become research issues among the (Human-Computer Interaction) HCI and Visualization community. We are particularly interested in finding answers to the following question: "How do we know if information visualization tools are useful and usable for real users performing real visualization tasks?", i.e., how do we know if a IV technique is able to provide insight to users? In fact, for effective and well-accepted adoption of information visualization tools, they have to be effective, efficient and satisfying for the intended users [Yi et al. 2008].

As new applications are more often related to larger and more complex datasets, the challenges of information visualization involves not only the selection of typical datasets and tasks but defining evaluation methodologies [Shneiderman and Plaisant 2006] and finding the appropriate case studies and users [Seo and Shneiderman 2006] [Valiati et al. 2008]. There has been already some research aiming at consolidating visualization tasks [Amar et al. 2004], collecting benchmark datasets [Plaisant et al. 2008], and experimenting human-computer interaction evaluation methods [Valiati et al. 2008].

The work on identifying and understanding "visualization" users tasks started in 1990, but remained set aside for a long time. Recently reports have been published devoted to understanding and representing the tasks users perform to accomplish their goals. Indeed, the identification and understanding of the nature of the users' tasks in the process of acquiring knowledge from visual representations of data is a recent branch in information visualization research, and some taxonomies have been proposed, for example, [Amar et al. 2004]. We have been working on that subject [Valiati et al. 2007], and have achieved a stable set of tasks that are useful both in the evaluation of new developments and to compare different techniques for some specific application.

Many of the evaluation reports we find in the literature are descriptions of experiments with users targeting the comparison of different IV techniques or assessing a specific technique regarding different tasks [Bertini et al. 2006] [Bertini et al. 2008]. Most of those works focused on controlled experiments in-laboratory, with specific tasks selected from the typical ones performed by the potential users of the visualization technique. Some works addressed different aspects of the evaluation process [Whiting et al. 2008]. The experimental procedure may be inadequate sometimes, mainly during a research exploratory stage when goals and tasks may not be already defined. Thus, for evaluating visualizations, longitudinal or (more broadly) qualitative research studies involving actual users' participation [Shneiderman and Plaisant 2006] [Isenberg et al. 2008] have been strongly recommended.

3. Some Important Issues for a User-Centered Viewpoint of Usability Evaluation of Information Visualization Techniques

A review of existing work on IV techniques evaluation allowed us to identify a significant number of problems.

First of all, from those methods used in traditional interaction evaluation, few are actually used for IV techniques evaluation: *the diversity of methods is quite limited*. In practice, most of the evaluation in IV is oriented to user testing methods. These methods try to check if usability goals are met and also to identify usability problems by conducting experiments, in which users try to solve realistic tasks with the technique. The dependent variables usually measured in this process are task time and task accuracy. Data collected is analyzed statistically, in order to capture the central usability measures effectiveness, efficiency and user satisfaction. However, user testing is not always the best choice because it is a very time consuming process, with high costs.

Two other problems arise from that one. *Evaluation happens too late when employing only user testing*, because such testing should be mainly applied in later stages of development: to testing, a running system is mandatory, and in general, *evaluation process does not follow a general usability evaluation methodology*. Sometimes, instead of empirical methods like user testing (which can only be used after some form of interaction design is available, and for which direct access to end-users is mandatory), it may be interesting to adopt some analytical methods like expert evaluation (inspection based solely on the evaluator's knowledge and experience), document-based evaluation (inspection based on some guiding documents, at various degrees of precision) or even model-based methods (inspection based on some theoretical models, usually cognitive models). These methods are particularly useful when it is not possible to collect data directly from users; but also they can be simply a useful first step to uncover some major usability flaws before investigating more complex problems further.

Our intention is to put emphasis on the proposal of guidelines (not strong recommendations or a strict methodology). From several case studies conducted by Valiati et al. (2008) following the MILCs approach [Shneiderman and Plaisant 2006], we devised that: (1) The context of usage for evaluation must be defined before the beginning of evaluation; (2) Evaluation needs to know who the users are, of what their goals are and to decide which users to support; (3) Evaluation needs to understand which tasks users need to perform and their characteristics (steps, constraints, and other tasks attributes like frequency, priority, etc) and to decide which tasks to support; (4) Evaluation should be performed earlier in the design-development cycle.

We are also focusing on the identification of appropriate usability methods for IV techniques evaluation. Methods for evaluating usability are described in terms of the characteristics, which should be present in the way the evaluation is conducted. The detailed prescription of a method to be used is in relation to a narrowly defined set of user goals, in a specified task domain, with limited metrics. However, much work is needed to extend the scope of current evaluation methods to cope with the many possible IV techniques' usability dimensions. For example, how do we apply usability inspection methods (like heuristics evaluation, for example) to IV techniques? Or how do we guide IV techniques development with these usability considerations?

Consequently, some interesting questions are open for discussion: (1) Can the sound ergonomic knowledge (style guides, architectures, and design and evaluation methods that have been proved adequate for GUIs and Web-based contexts) be transferred to this novel context of IV techniques? If so, how do we deal with the idiosyncrasies of IV techniques: in a generic or in a customized to IVT-specificities way? How do we ensure user involvement in usability evaluation? In fact, user involvement is a direct way to accelerate the process of improving usability evaluation of IV techniques. We think that doing user testing earlier allows usability knowledge be gained rapidly, rather than having simply the technology perfected without user concern.

References

- Amar, R., Eagan, J., Stasko, J. (2005) "Low-Level Components of Analytic Activity in Information Visualization". *Proc. Information Visualization 2005*, pp. 111-147.
- Bertini, E., Plaisant, C., Santucci, G. (2006) "BELIV: Beyond Time and Error: Novel Evaluation Methods for Information Visualization". *Proceedings (Workshop of the Advanced Visual Interfaces Conference)*. Venice, Italy, May 23, ACM Digital Library.
- Bertini, E. et al. (2008) "BELIV: Beyond Time and Error: Novel Evaluation Methods for Information Visualization". *Proceedings (Workshop of the Conference on Human Factors in Computing Systems)*. Florence, Italy, April 5, ACM Digital Library.
- Freitas, C.M.D.S. et al. (2002) "Evaluating Usability of Information Visualization Techniques". In *Proc. Symposium on Human Factors in Computer Systems (IHC 2002)*, Fortaleza, Brazil.
- Isenberg, P. et al. (2008) "Grounded evaluation of information visualizations". In *BELIV'08*, April 5, Florence, Italy. ACM Digital Library.
- Plaisant, C. (2004) "The Challenge of Information Visualization Evaluation". *Proc. AVI'04 - ACM Conference on Advanced Visual Interfaces*, ACM Press, pages 109-116.
- Plaisant, C., Fekete, J.-D., Grinstein, G. (2008) "Promoting Insight Based Evaluation of Visualizations: From Contest to Benchmark Repository". *IEEE Transactions on Visualization and Computer Graphics*, 14(1): 120-134.
- Scapin, D. L. and Bastien, J. M. C. (1997) "Ergonomic Criteria for Evaluating the Ergonomic Quality of Interactive Systems". *Behaviour and Information Technology*, 16 (4) 220-231.
- Seo, J. and Shneiderman, B. (2006) "Knowledge discovery in high dimensional data: Case studies and a user survey for the rank-by-feature framework". *IEEE Trans. on Visualization and Computer Graphics* 12(3): 311-322.
- Shneiderman, B., Plaisant, C. (2006) "Strategies for Evaluating Information Visualization Tools: Multi-dimensional In-depth Long-term Case Studies". In *BELIV'06*, May 23, Venice, Italy, 38-43.
- Valiati, E. et al. (2007) "Experimental Evaluation of Tasks Classification in Multidimensional Information Visualization". In *Proc. CLIHC'07 – Workshop of INTERACT'07*.
- Valiati, E., Pimenta, M., Freitas, C.M.D.S. (2008) "Using Multi-dimensional In-depth Long-term Case Studies for Information Visualization Evaluation". In *BELIV'08*, April 5, Florence, Italy. ACM Digital Library.
- Whiting, M.A., Haack, J., Varley, C. (2008) "Creating realistic, scenario-based synthetic data for test and evaluation of information analytics software". In *BELIV'08*, April 5, Florence, Italy. ACM Digital Library.
- Yi, J.S. et al. (2008) "Understanding and Characterizing Insights: How do People Gain Insights Using Information Visualization?" In *BELIV'08*, April 5, Florence, Italy. ACM Digital Library.