Methods for the Evaluation of an Interactive InfoVis Tool Supporting Exploratory Reasoning Processes

Markus Rester
Institute of Design and Assessment of Technology
Vienna University of Technology
Favoritenstraße 9-11/187
A-1040 Vienna, Austria
markus@igw.tuwien.ac.at

Margit Pohl
Institute of Design and Assessment of Technology
Vienna University of Technology
Favoritenstraße 9-11/187
A-1040 Vienna, Austria
margit@igw.tuwien.ac.at

ABSTRACT

Developing Information Visualization (InfoVis) techniques for complex knowledge domains makes it necessary to apply alternative methods of evaluation. In the evaluation of Gravi++ we used several methods and studied different user groups. We developed a reporting system yielding data about the insights the subjects gained during the exploration. It provides complex information about subjects' reasoning processes. Log files are valuable for time-dependent analysis of cognitive strategies. Focus groups provide a different view on the process of gaining insights. We assume that our experiences with all these methods can also be applied in similar evaluation studies on InfoVis techniques for complex data.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Evaluation/Methodology*; H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems—*Animations, Evaluation/Methodology*

1. INTRODUCTION

The application of InfoVis in ill-structured domains can be particularly valuable. By ill-structured domains we mean highly complex areas of knowledge characterized by irregular solutions to problems. There is no well-defined problem space or a limited number of clear solutions. In such knowledge domains InfoVis can be helpful as an explorative tool. Nevertheless, it is difficult to evaluate such forms of InfoVis because the traditional approach of measuring the time needed to find a solution and the number of errors can be misleading. Alternative forms of evaluation are, therefore, necessary. These should also cover activities like comparing, clustering or correlating data [10]. We, therefore, use the term 'insight' as defined by Saraiya et al. [13] to describe the users' activities in the following study. Insights are created in an interactive, hypotheses-generating process. This process is supported specifically by the visualization techniques used.

The following paper describes the evaluation of an interactive InfoVis technique. First we briefly discuss the evaluation of another InfoVis technique developed for the same application area. Both are supposed to support psychotherapists in their work with anorectic girls. During the therapy of these girls a large amount of highly complex data is collected. Statistical methods are hardly applicable for the analysis of these data because of the small sample size, the high number of variables, and the time-dependent character of the data. Only a small number of anorectic girls attend a therapy at any one time. The girls, their parents, and their therapists have to fill in numerous questionnaires before, during and after the therapy. In addition, progress in therapy is often not a linear process but a development with ups and downs. All this indicates that visualizations might be a more appropriate method of analysis than statistics. The aim of the therapists is to predict success or failure of the therapy for the individual patients depending on the answers the gave to the questionnaires, and, more generally, to analyze the factors influencing anorexia nervosa in more detail. In addition, they want to reduce the number of questionnaires the patients have to fill out. The Interactive Stardinates [5] and Gravi++ [3] were developed to support them in this work.

The evaluation of these two InfoVis techniques should not only yield overall information about their quality. The methods of evaluation should also indicate which specific features of an InfoVis technique are especially useful for given tasks. Furthermore, it would be valuable to compare different problem-solving strategies. In ill-structured domains, several ways to find a solution are usually possible. Software logs or observation techniques can be used to identify such strategies. The report system developed for the evaluation of Gravi++ allows more detailed information about how subjects got their insights.

The evaluation of the Interactive Stardinates technique has been finished recently. The evaluation of Gravi++ is part of an ongoing project. Most of the methods discussed in this paper have been already applied in this project but not all of the data has been analyzed.

BELIV 2006 Venice, Italy

2. RELATED EVALUATION APPROACH: THE INTERACTIVE STARDINATES

The Interactive Stardinates [5] are a hybrid InfoVis technique, which combines geometric- and glyph-based features. The data of one patient is represented by a bundle of lines with their vertices on the axes arranged in a circle. Each Stardinate can represent the data of one patient so the data set is decomposed into small multiples (see Figure 1).

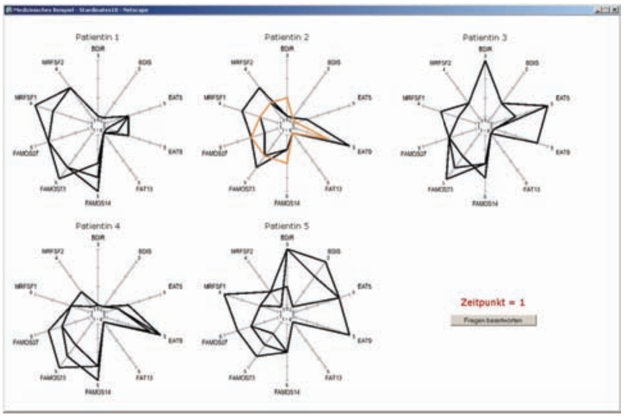


Figure 1: Five Stardinates Each Visualizing One Patient With Three Time Steps and 10 Questions (Time Step 1 of Patient 2 is Highlighted)

Lanzenberger et al. [6] investigated whether the Interactive Stardinates were a more appropriate InfoVis technique than Parallel Coordinates for the dataset described above. The subjects who tested the tools were no domain experts. The evaluation of the Stardinates used time measurements, categorization of insights and key statements as variables. The categorization of insights is slightly similar to the one proposed by Saraiya et al. [13] and includes categories as, for example, “comparing patients”, “overview”, “changes over time”, etc. The system of categories is domain specific, in contrast to other systems of categorization (see e.g., Pillat et al. [9]). Generic categorization of insights can be very valuable, but in the context of this study it did not seem to be appropriate. Key statements were formulated by experts after extensive study of the data. Especially, the latter two variables enable researchers to make more detailed conclusions about the quality of single features of the system (e.g., whether a visualization technique rather supports more general or more detailed insights or whether it supports the processing of dynamic data, etc.).

3. GRAVI++

The ability of the human perceptual system to locate and organize things spatially, perceive motion, etc. is utilized in Gravi++ [3] by positioning icons on the screen. These represent the patients and questionnaires they answered. According to the answer a patient gave to a question, the patient’s icon is attracted by the question’s icon. A spring-based system model is used to depict this so that every patient is connected to every question (see Figure 2 top left side). This leads to the formation of clusters of patients who gave similar answers (see Figure 2 bottom left side).

To deal with the time dependent data Gravi++ uses animation. The position of the patients’ icons change over time. This allows analyzing and comparing the changing values. Many visualization options are available, like Star Glyphs and attraction rings to communicate the exact values of each answer or traces to show the paths of the patients’ icons over all time steps (see Figure 2 right side).

Gravi++ provides various interaction possibilities to explore the data and generate new insights. The icons and visual elements can be moved, deleted, highlighted and emphasized by the user. Each change leads to an instant update of the visualization. For details on mental model, visualization options, user interactions, and implementation see [3].

4. GENERAL STUDY DESIGN

The usefulness of an InfoVis tool is not as predictable as the one of ‘classic’ software, because of the remarkable influence of human reasoning processes on success in application. So even after participative design and faithful development the outcome has to be evaluated to a high extent.

Usability not only matters but may become vital due to the interactive and explorative nature of many tasks users will perform. Therefore, on the one hand, one has to pay particular attention to usability questions in an iterative design process. On the other hand, a serious examination is also essential for the assessment of the InfoVis technique because of its interdependency with usability.

In many cases it is necessary to conduct ‘classic’ experiments as well as to do some kind of field observation. Often this is the one and only way to evaluate the usefulness for the ‘real world’ and ensure ecological validity. The same applies to the question of generalization. Sometimes, the usefulness of applying an InfoVis technique in another than the original context can only be assessed by asking the respective users to decide for themselves whether a technique makes sense in their setting and for their data and tasks and thus allowing for profound assessment of external validity.

Therefore, we believe that a sustainable assessment of any interactive InfoVis technique has to include the following four areas: (1) a usability evaluation to distinguish between weaknesses of the implementation and of the technique as such, (2) an experiment to collect quantitative data, (3) a case study to ensure ecological validity, and (4) an assessment of the transferability to other contexts. Each of these parts requires suitable investigation methods. Also the subjects who are tested may differ considerably. Such mixes of different methods complementing each other are also suggested in e.g., [14, 10].

In our case a visualization technique was developed for very specific users, data, and tasks. Although there are other appropriate concepts for this special domain (see section 2 Stardinates), these visualizations are not available in an advanced stage of implementation. Therefore, a comparative study of different visualizations like the one of Saraiya et al. [13] is not possible. But it seems a viable approach - in trying to understand how reasoning processes take place and what qualities of InfoVis support them - to compare the InfoVis technique with other methods used so far. We decided for a comparative study of Explorative Data Analysis (EDA) as introduced by Tukey in 1977, various methods of Machine Learning (ML), and Gravi++ concentrating on insights during the exploration process.

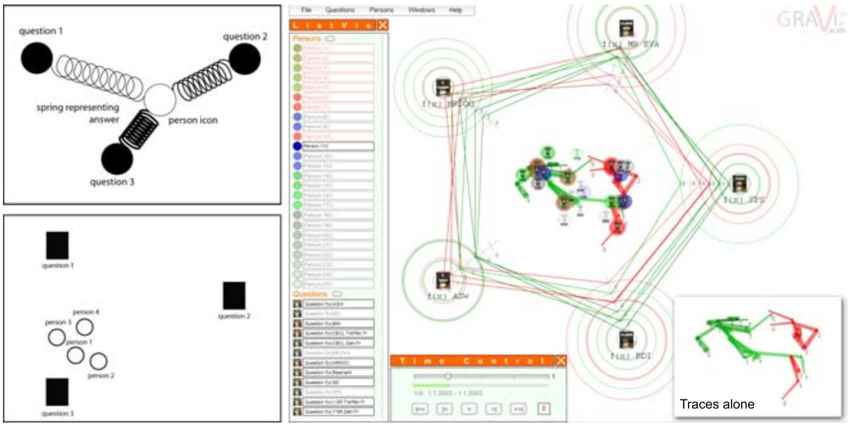


Figure 2: Gravi++ Concept of Spring-Based Positioning (Top Left), Leading to Formation of Clusters (Bottom Left), and a Typical Screenshot (Right)

Method	Usab.	Vis. Techn.	Case Study	Transfer
Usability Insp.	X			
Heuristic Eval.	X			
Insight Reports		X		
Focus Groups	X	X		
Log Files		X		
Thinking-Aloud			X	
Interviews			X	X

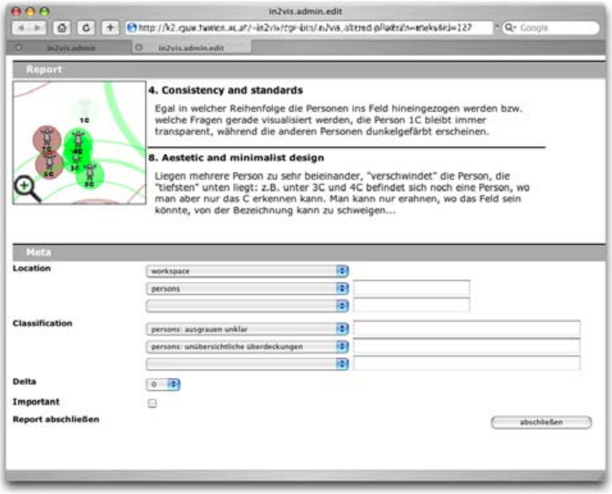
Table 1: Used Methods in Various Stages of Evaluation (Usability, Visualization Technique, Case Study, and Transferability)

5. EVALUATION OF GRAVI++

In the four different stages of evaluation various methods are used: usability inspection, heuristic evaluation [8], insight reports, focus groups [4], log files, thinking-aloud [1], and interviews (see Table 1). Of course, also different populations are studied according to the respective focus in the different stages.

5.1 Usability Evaluation

In the first place we conducted extensive usability studies (see [12]) with three different methods: usability inspection by an expert to spot the most obvious glitches and fix them so that the subjects of the following heuristic evaluation were able to concentrate on more sophisticated problems. Subjects in this part do not have to be domain experts. Therefore, 27 computer science students who can be described as semi-experts in the field of usability participated. Because of the importance of usability we decided for rather a large number of evaluators to ensure the acquisition of relevant data for improving the inspected software. Due to the number of evaluators this test took place in a laboratory setting. A report system (see Figure 3), similar to the one used later for the documentation of insights, was provided to assist the subjects in documenting their findings. Although we did not extend Nielsen's heuristic used by the evaluators with categories like visual representation usability (e.g., completeness, spatial organization) or data usability suggested by Freitas et al. [2], some reports covered problems of this areas. In a final step we held focus groups with the same subjects. They did not reveal any new usability problems but gave an important and quite different perspective on the experiences the subjects made (see [12] for details).


in2vis.admin.edit
http://k2.cguw.tuwien.ac.at/~in2vis/cgi-bin/in2vis_altered.pl?adts2n=meks&rid=127
Google
in2vis.admin
in2vis.admin.edit
1C
7C
3C
4C
2C

Report

4. Consistency and standards

Egal in welcher Reihenfolge die Personen ins Feld hineingezogen werden bzw. welche Fragen gerade visualisiert werden, die Person 1C bleibt immer transparent, während die anderen Personen dunkelgefärbt erscheinen.

8. Aestetic and minimalist design

Liegen mehrere Person zu sehr beieinander, "verschwindet" die Person, die "tiefsten" unten liegt: z.B. unter 3C und 4C befindet sich noch eine Person, wo man aber nur das C erkennen kann. Man kann nur erahnen, wo das Feld sein könnte, von der Bezeichnung kann zu schweigen...

Meta

Location workspace
persons

Classification persons: ausgrauen unklar
persons: unübersichtliche überdeckungen

Delta 0

Important □

Report abschließen abschließen

Figure 3: Administrative Access to a Typical Report Generated by One of the Subjects Showing a Screenshot of a Usability Problem and Two Descriptions of Violated Usability Principles with Multiple Classification Options for the Investigator

5.2 Evaluation of Visualization Technique

In an experiment 33 subjects, once again students, who were domain novices received both an introduction (about an hour) to the domain, data, and tasks and an explanation of the three methods to use (about half an hour each). Although the real users are clinicians, a comparative study is also interesting with domain novices if one keeps this fact in mind and does not jump to conclusions in assessing the investigated technique. Often there are not that many subjects available that belong to the group of real users so that it is much more reasonable to use qualitative methods, like interviews with them. Moreover, it is an important question how useful an investigated technique is for experts compared to domain novices. Divided into three groups the novices used EDA (histograms, scatterplots, boxplots, and descriptive statistics), ML (a pruned C4.5 decision tree and the sequential minimal optimization algorithm for training support vector machines), and Gravi++ in different order. The subjects received a handout explaining the test procedure, any used abbreviations, and a short documentation of the three methods. EDA and ML was available as printout material. Due to the complexity of domain and data, scenarios were provided: two specifying meaningful subsets of data (questions, patients, time steps) to explore and two more stating concrete questions. The time constraints of the four scenarios were 25, 20, 10, and 10 minutes. During the use of Gravi++ software logs were recorded. In the end three focus groups were held.

5.2.1 Insight Reports

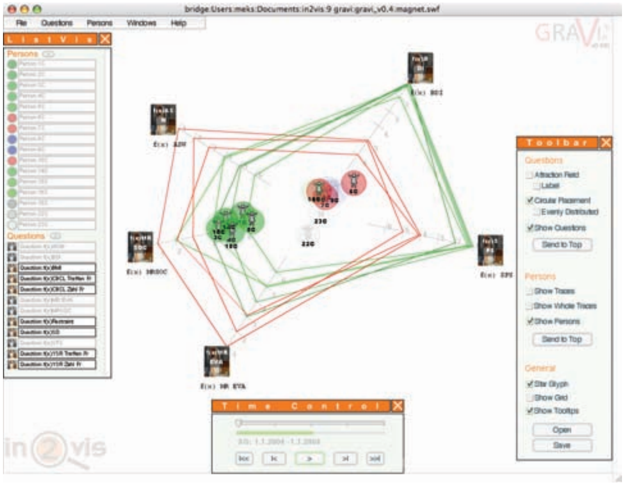
The report system used by the subjects to document their insights during the exploration process was similar to the one used in the heuristic evaluation (compare Figure 3). It is implemented in Perl and MySQL and has following features: screenshot upload in the case of Gravi++ and specification of used materials (e.g., histograms, boxplots, C4.5) in the case of EDA and ML, insight description, and confidence rating (three-step: low, medium, and high). Once again an administrative access to the documented insights provides multiple classification options (e.g., insight, plausibility, complexity, argument) for the recorded 909 reports.

Due to the explorative nature, there is a considerable solution space for possible insights. Furthermore, it is sometimes a difficult task to rate the correctness in a range of true and false. It is often more a question of plausibility. Of course, we designed the scenarios which defined meaningful subsets of the data in close cooperation with our domain experts and also asked them for an extensive list of possible insights. But it is impossible to anticipate all valid argumentation.

There exist some complex task classifications and taxonomies and Morse et al. [7] developed sophisticated methods for exhaustively testing the capabilities of visualizations based on them. This results in very specific exercises the subjects have to solve with clear and unambiguous answers. Because our main interest lies in reasoning processes and exploration strategies we did not ask such detailed questions.

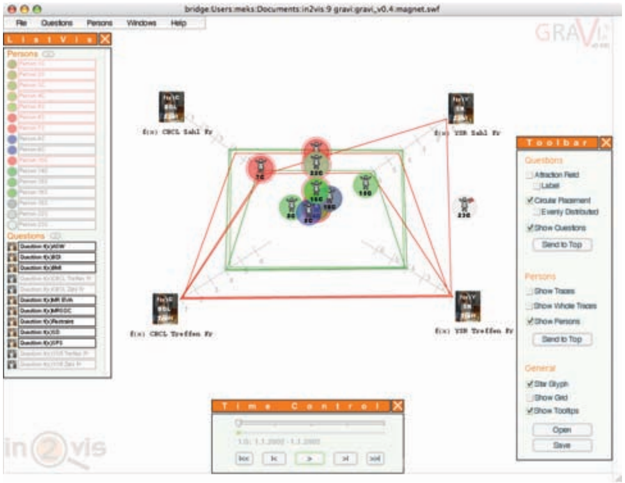
The four scenarios we used tried to cover different types of investigation on the data. Two scenarios were undirected specifications of which data to explore, just asking for documentation of whatever the subjects find out. The other two were precise questions but still requiring argumentation rather than allowing yes/no answers. The types of investigations can be characterized as follows: (1) realize the change over time of 16 patients in 5 dimensions and identify positive and negative predictors (see Figure 4), (2) recognize the consistent/inconsistent answers of parents and patients in the first time step and their role as predictor (see Figure 5), (3) analyze the effect of the therapy on one specific dimension over time (see Figure 6), and (4) predict a positive or negative therapy outcome of a so far unclassified patient with the available data of the first two time steps (see Figure 7).

We use a rather simple schema for classification of the documented insights in terms of complexity (trivial, regular, complex) and time dependency (static or dynamic observation). In a bottom-up procedure based on the collected data



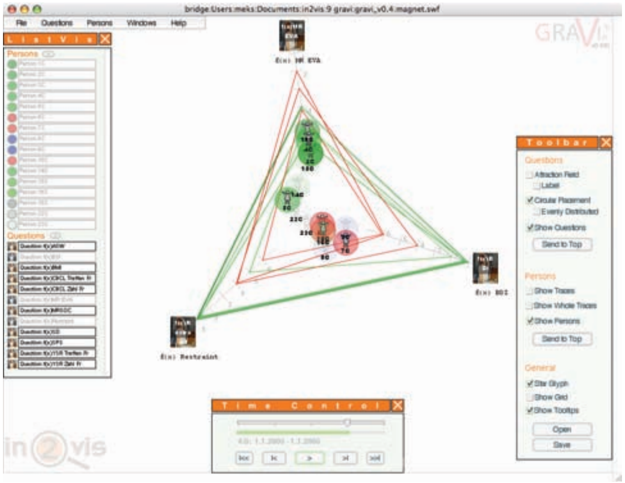
Questions
□ Attraction Field
□ Label
☑ Circular Placement
□ Evenly Distributed
☑ Show Questions
Send to Top
Persons
□ Show Traces
□ Show Whole Traces
☑ Show Persons
Send to Top
General
☑ Star Glyph
□ Show Grid
☑ Show Tooltips
Open
Save

Figure 4: Possible Visualization of Scenario 1



Questions
□ Attraction Field
□ Label
☑ Circular Placement
□ Evenly Distributed
☑ Show Questions
Send to Top
Persons
□ Show Traces
□ Show Whole Traces
☑ Show Persons
Send to Top
General
☑ Star Glyph
□ Show Grid
☑ Show Tooltips
Open
Save

Figure 5: Possible Visualization of Scenario 2



Questions
□ Attraction Field
□ Label
☑ Circular Placement
□ Evenly Distributed
☑ Show Questions
Send to Top
Persons
□ Show Traces
□ Show Whole Traces
☑ Show Persons
Send to Top
General
☑ Star Glyph
□ Show Grid
☑ Show Tooltips
Open
Save

Figure 6: Possible Visualization of Scenario 3

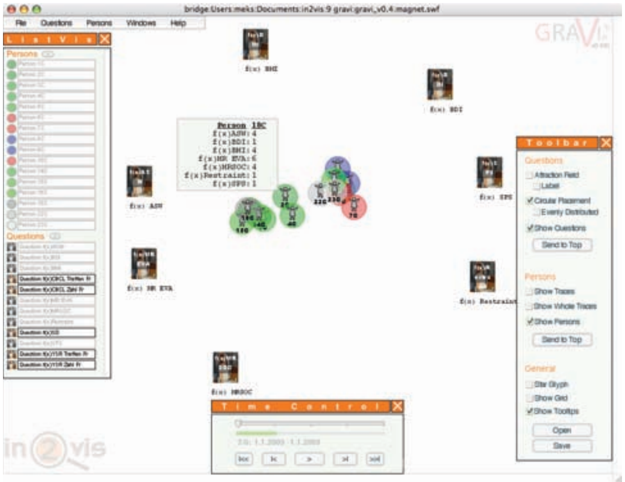


Figure 7: Possible Visualization of Scenario 4

basic categories are being developed to describe the documented insights as accurately as possible (cf. [6, 13]).

We want to concentrate on the course of insights and will try to look for different strategies of exploration. Therefore the classification of insights and the exact identification of a particular insight in different reports of different subjects will be the crux of the matter. Significant differences in the uploaded screenshots may indicate various strategies. The fact that reports vary from a few words to several paragraphs complicates the procedure. Unsolved problems include: Should we split long reports in basic insights or are they a unique occurrence of a complex insight? Are they simply a cumulative documentation from a subject who did not adhere to the test procedure of reporting insights immediately after having them?

5.2.2 *Focus Groups*

With the same subjects we conducted focus groups of about 100 minutes each. A moderator was present to structure the discussion according to a detailed guideline which was compiled to ensure comparability of collected data. The guideline also stated some questions which had to be answered by every subject. The value of this method is that it reveals subjective impressions on questions not asked before and gives a different perspective on the data collected in the experiment. There were 2-3 weeks between the lab setting and the focus groups for the subjects to have enough distance from particular incidences.

The same four topics were discussed for every of the three used methods (EDA, ML, Gravi++) in the lab: (1) appropriateness of the allowed time, (2) ease of use and usefulness of the method for gaining insights, (3) overall confidence in insights gained with the method, and (4) major strength and weakness of the method. At last the subjects debated four more questions: (5) similarity and difference of gained insights using different methods, (6) assumed comprehension rates of the complex matter with each method, (7) appropriateness of combined use of the three methods, and (8) order for best possible comprehension of the data.

Some of the results of the focus groups will help to relativize or better understand findings of the analysis of the lab experiment (e.g., confidence ratings, similarity of insights). Other results will be important for a correct interpretation like the appropriateness of allowed time with regard to the amount of documented insights.

5.2.3 *Log Files*

During the experiment which took place in a computer science laboratory each and every interaction of the subjects with Gravi++ was recorded in log files. After having troubles with log file analysis elsewhere [11], we knew the importance of carefully designing the structure of the various log file entries. For a straightforward interpretation one has for instance to ensure that every entry is self-contained. Otherwise simple processing is impossible because of the need to carry along context information of many preceding entries or possibly even succeeding lines to derive the original user interaction.

After precise formulation of logging requirements, the source code of Gravi++ was extended to produce the needed output, such as activation and deactivation of visualization options (Star Glyph, attraction rings, traces, etc.), data selection (add/remove person or add/remove question via drag and drop or context menus), and exploration activities (hypotheses generation/verification/falsification via drag and drop, navigation through time steps).

The recorded data will, once aligned with the documented insights, allow for correlation between used visualization options or exploration strategies and types of insight.

5.3 Case Study

The importance of longitudinal studies has been outlined above [10]. With the real users - in our special case there are only two - qualitative evaluation methods are often more appropriate. Their in-depth knowledge is invaluable and should in any case be reached by the chosen method. We utilize interviews and thinking-aloud with the experts in their real work environment to collect data on feasibility and usefulness. Although they were tightly involved in an iterative design and development process, success is not guaranteed and has to be proven. For instance, the ability of Gravi++ in supporting the experts in the generation of new hypotheses through exploration is one interesting question.

5.4 Transferability

One of the methods used in the case study, namely interviews, is used with a population of 20 experts of other domains (e.g., social sciences, science of history). This will help to assess the usefulness of Gravi++ in knowledge domains other than the narrow area of medical science it was developed for.

6. CONCLUSIONS

Developing InfoVis techniques for complex knowledge domains makes it necessary to apply alternative methods of evaluation. Different areas should be covered in a sustainable evaluation of an interactive InfoVis tool: usability study, controlled experiment, case study, and transferability assessment.

In the evaluation of Gravi++ we used several methods and studied different user groups. We developed a reporting system for the evaluation process yielding data about the insights the subjects gained during the exploration of the data. The results of this system can be compared to the results from a thinking-aloud investigation. On the one hand, the reporting system is more efficient than the thinking-aloud method. On the other hand, it yields less detailed data. Nevertheless, it is very valuable in getting complex information about subjects' reasoning processes. Log file analysis can also make reasoning processes more transparent. They are especially useful for the time-dependent analysis of cognitive strategies. The data resulting from the reporting system and the log file analysis can be correlated to the number and complexity of the insights gained by the subjects. In this way, useful cognitive strategies for working with Gravi++ can be identified. We also developed categorization systems for the insights. Analyzing data resulting from the categorization might help to identify specific strengths and weaknesses of InfoVis techniques. We combined the reporting system with focus groups. These two methods complement each other and provide two different views on the process of gaining insights. We assume that the experiences we made with all these methods can also be applied in similar evaluation studies on InfoVis techniques for complex data.

Most of these methods have been implemented by now. The analysis of most of the data is currently work in progress. We have already studied the data from the usability study. The results show that the methods used there (i.e. usability inspection, heuristic evaluation in a laboratory setting, and focus groups) can be combined in a meaningful way. After having analyzed the data of the visualization technique evaluation (obtained with insight reports, focus groups, thinking-aloud, and interviews) we will know whether our assumptions about the specific uses of the various methods are valid there too.

7. ACKNOWLEDGMENTS

The project "Interactive Information Visualization: Exploring and Supporting Human Reasoning Processes" is financed by the Vienna Science and Technology Fund (WWTF) [Grant WWTF CI038].

8. ADDITIONAL AUTHORS

Klaus Hinum (Institute of Software Technology & Interactive Systems, Vienna University of Technology, Austria, email: hinum@ifs.tuwien.ac.at), Silvia Miksch (Institute of Software Technology & Interactive Systems, Vienna University of Technology, Austria, email: silvia@ifs.tuwien.ac.at), Christian Popow (Department of Child and Adolescent Psychiatry, Medical University of Vienna, Austria, email: christian.popow@meduniwien.ac.at), Susanne Ohmann (Department of Child and Adolescent Neuropsychiatry, Medical University of Vienna, Austria, email: susanne.ohmann@meduniwien.ac.at), and Slaven Banovic (Institute of Design and Assessment of Technology, Vienna University of Technology, Austria, email: banovic@xover.htu.tuwien.ac.at).

9. REFERENCES

[1] T. M. Boren and J. Ramey. Thinking aloud: Reconciling theory and practice. *Professional Communication, IEEE Transactions on*, 43(3):261–278, September 2000.

[2] C. Freitas, P. Luzzardi, R. Cava, M. Winckler, M. Pimenta, and L. Nedel. On evaluating information visualization techniques. In *Proceedings of the working conference on Advanced Visual Interfaces*. ACM Press, 2002.

[3] K. Hinum, S. Miksch, W. Aigner, S. Ohmann, C. Popow, M. Pohl, and M. Rester. Gravi++: Interactive information visualization to explore highly structured temporal data. *Journal of Universal Computer Science (J.UCS) – Special Issue on Visual Data Mining*, 11(11):1792–1805, 2005.

[4] M. Kuniavsky. *User Experience: A Practitioner's Guide for User Research*. Morgan Kaufmann, 2003.

[5] M. Lanzenberger, S. Miksch, and M. Pohl. The stardinates—visualizing highly structured data. In *Proceedings of the Int. Conference on Information Visualization (IV03), July 16–18, 2003, London, UK*, pages 47–52. IEEE Computer Science Society, 2003.

[6] M. Lanzenberger, S. Miksch, and M. Pohl. Exploring highly structured data - a comparative study of stardinates and parallel coordinates. In *Proceedings of the 9th Int. Conference on Information Visualisation (IV05), July 6–8, 2005, London, UK*, pages 312–320. IEEE Computer Science Society, 2005.

[7] E. Morse, M. Lewis, and K. A. Olsen. Evaluating visualizations: using a taxonomic guide. *Int. J. Human-Computer Studies*, 53(5):637–662, 2000.

[8] J. Nielsen. *Heuristic Evaluation*, chapter 2, pages 25–62. John Wiley & Sons, Inc., New York, 1994.

[9] R. Pillat, E. Valiati, and C. Freitas. Experimental study on evaluation of multidimensional information visualization techniques. In *CLIHC '05: Proceedings of the 2005 Latin American conference on Human-computer interaction*, pages 20–30, ACM Press, New York, NY, USA, 2005.

[10] C. Plaisant. The challenge of information visualization evaluation. In M. F. Costabile, editor, *Proceedings of the working conference on Advanced visual interfaces*, pages 109–116. ACM Press, 2004.

[11] M. Rester and M. Pohl. Ecodesign - an online university course for sustainable product design. In P. Kommers and G. Richards, editors, *Proceedings of of ED-MEDIA 2005. World Conference on Educational Multimedia, Hypermedia and Telecommunications. Montreal, Canada*, pages 316–323, Association for the Advancement of Computing in Education (AACE), Norfolk, VA, 2005.

[12] M. Rester, M. Pohl, K. Hinum, S. Miksch, S. Ohmann, C. Popow, and S. Banovic. Assessing the usability of an interactive information visualization method as the first step of a sustainable evaluation. In A. Holzinger and K.-H. Weidmann, editors, *Empowering Software Quality: How can Usability Engineering reach these goals?*, volume 198 of *books@ocg.at*, pages 31–44. Austrian Computer Society, 2005.

[13] P. Saraiya, C. North, and K. Duca. An insight-based methodology for evaluating bioinformatics visualizations. *Visualization and Computer Graphics, IEEE Transactions on*, 11(4):443–456, 2005.

[14] M. Tory and T. Möller. Evaluating visualizations: do expert reviews work? *Computer Graphics and Applications, IEEE*, 25(5):8–11, 2005.