

Usability and Transferability of a Visualization Methodology for Medical Data

Margit Pohl, Markus Rester, and Sylvia Wiltner

Vienna University of Technology, 1040 Vienna, Austria
Institute of Design & Assessment of Technology
{margit, markus}@igw.tuwien.ac.at

Abstract. Information Visualization (InfoVis) techniques can offer a valuable contribution for the examination of medical data. We successfully developed an InfoVis application – Gravi – for the analysis of questionnaire data stemming from the therapy of anorectic young women. During the development process, we carefully evaluated Gravi in several stages. In this paper, we describe selected results from the usability evaluation, especially results from qualitative studies. The results indicate that Gravi is easy to use and intuitive. The subjects of the two studies described here especially liked the presentation of time-oriented data and the interactivity of the system. In the second study, we also found indication that Gravi can be used in other areas than the one it was developed for.

Keywords: Information Visualization, Usability, Utility, Transferability, Focus Groups, Interviews.

1 Introduction

Computers are used increasingly in a medical context. The usability of such systems is especially important as the consequences of mistakes made when using such systems can be critical. Computers are used for a wide variety of purposes in medicine. In this paper, we describe an InfoVis tool supposed to support psychotherapists in their work with anorectic young women. The aim of information visualization is to represent large amounts of abstract data (e.g. results from questionnaires, data about the development on financial markets, etc.) in a visual form to make this kind of information more comprehensible. Medicine is a very important application area for InfoVis methods [1], especially because of their flexibility and the possibility of representing time oriented data. Chittaro also points out the significance of applying design guidelines offered by Human-Computer Interaction in such applications. For abstract data, there is usually no natural mapping of data on the one hand and visual representation on the other hand (in contrast to e.g. geographical information systems where there is a natural mapping between maps and physical space). Therefore, the design and testing of InfoVis methodologies is especially important. A more general description of the significance of usability research in information visualization can be found in [2] and [3]. A description of different

methods of usability testing can be found in [4] and [5]. For the application of such methods in a study evaluating the feasibility and usability of digital image fusion see [6] and for the importance of user centered development in the medical domain [7].

InfoVis methodologies are supposed to support humans in analyzing large volumes of data. In many cases, these processes can be seen as activities of exploration. [8] describes the process of seeing as a series of visual queries. [9] investigates user strategies for information exploration and uses the concept of information foraging to explain such behavior. [10] and [11] use the term sensemaking to describe the cognitive amplification of human cognitive performance by InfoVis methodologies. These methodologies help to acquire and reorganize information in such a way that something new can be created. [12] pointed out the importance of information visualization for „creating creativity“. [13] distinguishes between two different forms of search – simple lookup tasks and searching to learn. Searching to learn (which is similar to exploration) is more complex and iterative and includes activities like scanning/viewing, comparing and making qualitative judgments. This distinction seems to be important as different forms of evaluation methods are appropriate for these two forms of search. Testing the usability of exploratory systems is more difficult than testing systems for simple lookup (see e.g. [14]). We developed a report system (described in [15]) to investigate the usability of an exploratory InfoVis methodology in medicine.

The following study describes an investigation in how best to support psychotherapists in their work. The aim of these therapists is to analyze the development of anorectic young women taking part in a psychotherapy. During this process a large amount of data is collected. Statistical methods are not suitable to analyze these data because of the small sample size, the high number of variables and the time oriented character of the data. Only a small number of anorectic young women attend a therapy at any one time. The young women and their parents 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 of this indicates that InfoVis techniques might be a better method of analysis of these data. The aim of the therapists is to predict success or failure of the therapy depending on the results of the questionnaires, and, more generally, to analyze the factors influencing anorexia nervosa in more detail. This process is explorative in nature as there exists no predefined “correct” solution for the therapists’ problems but several plausible insights might be got during the analysis of the data.

The study presented here is part of a larger project called in2vis. The aim of the project was the development of an appropriate InfoVis methodology to support the therapists in their work with anorectic young women. This methodology is called Gravi (for a more detailed description see below). This methodology was evaluated extensively. In addition, it was compared to other methods of data analysis (Machine Learning and Exploratory Data Analysis). The evaluation process took place in two stages – the usability study and the utility study (insight study and case study). In a last stage of the utility study, we tested whether the visualization methodology developed for a certain area of application could be useful for other areas as well. We called this transferability study. For an overview and short description of the different stages see Fig. 1.

We distinguish usability and utility aspects of the system to avoid mixing up usability problems and problems with the InfoVis methodology and its mapping of data and visualization as such. In this paper, we present results from the usability study and from the so-called transferability study which also investigated usability aspects. Machine Learning and Exploratory Data Analysis were not considered in the usability and the transferability study. The results of the utility study have been presented elsewhere [16, 17]. A preliminary description of parts of the usability study were described in [15].

Stage	Method	Subjects	Aim	Outcome
usability study	usability inspection	1 usability expert	spot most obvious glitches	31 severe usability problems
	heuristic evaluation	27 semi-experts in usability (students)	in depth testing	447 reports ⇒ 576 problems 221 different problems
	focus groups		additional usability assessment	no new problems BUT different perspective
insight study	insight reports	32 domain novices (students)	patterns of insight & cognitive strategies	876 reports ⇒ 2166 insights 668 different insights
	log files		used vis. options & exploration strategies	56055 log file entries
	focus groups		relativize findings & aid correct interpretation	transcription of 3x 100min
case study	interviews	2 real users (clinicians)	feasibility & usefulness in real life	transcription of 1x 60min
	observation & thinking aloud			notes on 1x 180min
transferability study	interviews	14 experts of other domains	usefulness in other domains	transcription of 14x 60min

Fig. 1. Overall evaluation study design. At four different stages diverse evaluation methods are utilized. Quantitative and qualitative methods complement each other. Appropriate subjects are tested according to the various aims of the respective stages. The two parts which results are discussed in this paper are highlighted (red and blue).

2 Description of the Visualization Methodology

The human perceptual system shows great ability to locate and organize things spatially, perceive motion, etc. This is utilized in Gravi [18] 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 Fig. 2 left side). This leads to the formation of clusters of patients who gave similar answers (see Fig. 2 right side). The therapists are especially interested in those variables which predict the outcome of the therapy (successful or not successful). By analyzing clusters of “positive” and “negative” cases they can identify those variables.

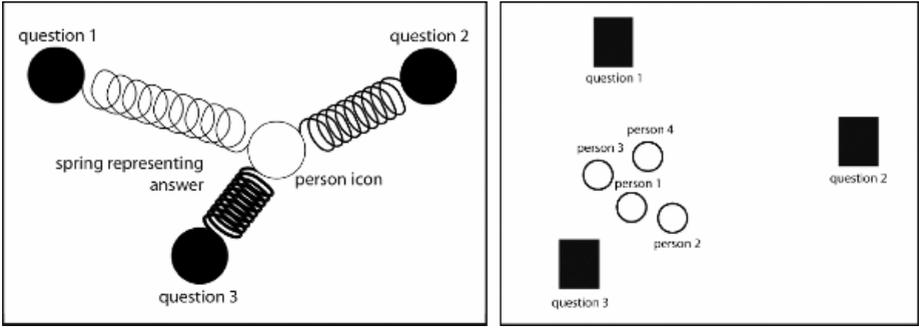


Fig. 2. Concept of spring-based positioning (left), leading to formation of clusters (right)

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 [19].

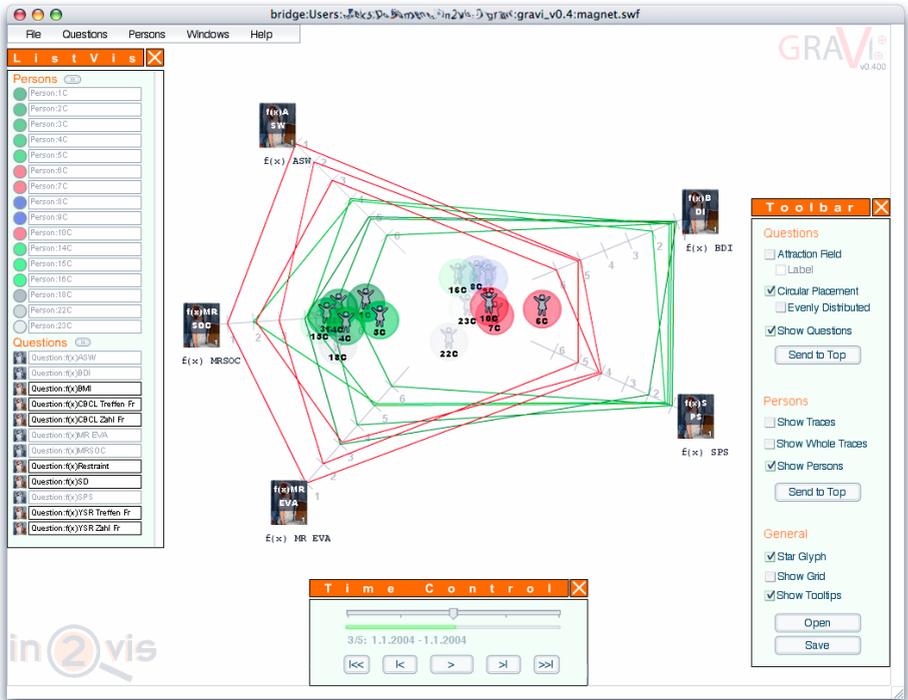


Fig. 3. Typical Screenshot of Gravi – Two clusters show similar answering behavior of patients with positive therapy outcome (green icons) and those with negative outcome (red icons). Star Glyphs – a connected line of the answer scores on each questionnaire for every patient – communicate exact values of all patients on showed questionnaires.

Many different visualization options are available, like Star Glyphs (see Fig. 3). To represent the dynamic, time dependent data Gravi uses animation. The position of the patients' icons change over time. This allows for analyzing and comparing the changing values. The therapists need this feature to visualize information recorded at different points in time. The development in time is a very important aspect of the analysis of the progress of the therapy. To visualize the change over time of patients in one view there is the option to show the paths the patients' icons will follow in the animation. These paths are called traces (see Fig. 4)

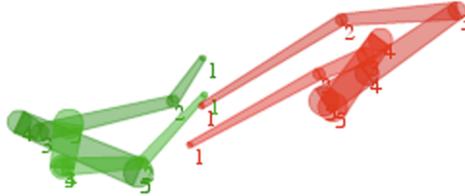


Fig. 4. Traces allow for the analysis of the changing answering behavior of multiple patients over all five time steps (i.e. the whole therapy which lasts for about one year). Shown here are the traces of four patients who start at almost identical positions according to the same five questionnaires of Fig. 3. Already at time step “2” we see divergent changes in answering behavior of those with positive and those with negative therapy outcome.

3 Usability Study (Focus Groups Following Heuristic Evaluation)

We utilized three different evaluation methods in our usability study: an informal inspection, heuristic evaluation, and focus groups (see Fig. 1). The first method revealed severe usability issues which have been corrected before any further evaluation took place. This should increase the quality of following assessments. Heuristic evaluation [20] and focus groups complemented each other with their different outcomes. For results of the heuristic evaluation see [15].

“Focus groups are structured group interviews that quickly and inexpensively reveal a target audience’s desires, experiences, and priorities” [21]. They are a method of qualitative research which involve group discussion and group interviewing. Participants discuss and comment on the topic that is under investigation. This relies heavily on their subjective view and experience [22]. Although one major strength of focus groups is that they may reveal issues not anticipated so far, due to the interactive setting, they have to be prepared well. So a guideline for the group moderator is very advisable to ensure that the discussion stays focused and covers predefined questions [cf. 23].

Whereas the outcome of heuristic evaluation was an extensive list of usability problems, focus groups allowed the subjects to give detailed arguments for their views on the usability of the assessed visualization methodology. Also a form of severity rating was compiled when the subjects discussed their views on the most important usability problems.

The subjects were computer science students who attended a course where they not only received lectures on usability but also had to work out several assignments. Therefore we can describe them as semi-experts in usability. Furthermore, they also received an introduction to the application domain in order to ensure basic understanding of real users' data and tasks.

The focus groups lasted about an hour each. The same 27 subjects who also took part in the heuristic evaluation were split into two groups of similar size for the focus groups. This is necessary because otherwise the discussion would be too chaotic or, and this would be even worse for an interesting outcome, the participants would not start a lively discussion referring to each other at all due to a too big group size. Another important argument for holding two different groups is that there may be some group members who are rather dominant regarding discussed topics. By holding at least two focus groups the overall outcome of this evaluation method will probably more fruitful.

The guideline for the discussion started with an introduction including an explanation of the focus group methodology and the showing of some screenshots of Gravi, so the subjects remember all the main features of the visualization. First, the overall impression of the usability was discussed. Most important here were the following questions: *Can Gravi be operated intuitively? What was the most severe usability problem? What was the best designed feature from a usability perspective?* Thereafter various parts of the user interface were subject to discussion. Understandability was the basic question for all these parts, but also structure or positioning were discussed for main menu, windows and toolbars, context menus, and workplace.

3.1 Intuitive Operation of Gravi

In focus group A the participants made 16 statements, 11 of which were on the positive or even very positive side. Nevertheless there were some constraints mentioned. The basic interactions and the mental model were assessed very understandable. Complex functionalities were rated more confusing.

In focus group B 11 statements are documented. 8 of these were – similar to group A – on the positive side. But in contrast to group A, the mentioned restrictions in group B have in common that at the beginning Gravi cannot be used intuitively but some time is needed to learn how to operate the system.

3.2 Most Severe Usability Problem

There were 46 statements concerning the most important usability problem. These 46 statements documented 27 different problems, 10 of which were mentioned more than once. With 7 mentions the problem “undo/redo is missing” was the most frequently documented one. For more details on this result see [15].

An evident characteristic of the gathered material is that, except for a few statements, the problems mentioned by the different focus groups were quite divergent. That confirms the claim for the necessity of conducting more than one focus group.

3.3 Best Designed Feature from a Usability Perspective

Members of group A stated 7 good features, including drag & drop interactions (two times), instant updates of visualization after any interaction, time control window to navigate through the time-dependent data, and the idea of only one workbench where all interaction and visualization takes place. Two more statements were concerned with positive aspects of the visualization methodology and are not usability issues as such.

Participants of group B stated similar positive features like drag & drop interaction but also some additional good elements: the possibility to quickly switch between much and less detail of the visualized data, the neutral design of the user interface that does not distract from reading the visualization, the beneficial color coding of the data, the existing alternative ways to accomplish some desired interaction (see Fig. 5), and direct interaction which gives the feeling of touching the visualization and therefore eases understanding.



Fig. 5. Alternative ways to remove icons of patients from the workbench: drag & drop interaction with a single icon: moving from workbench to ListVis window (top left), context menu of a single icon (top right), context menu of workbench to remove all icons (bottom left), and main menu entry to remove all icons from workbench (bottom right)

4 Usability and Transferability (Interviews with Experts)

The visualization methodology Gravi has been developed for a very specific purpose – the analysis of the therapy process of anorectic young women. As the development process for such methodologies is costly and time consuming, it is interesting to assess whether this tool could be used in other contexts as well. To achieve this, we conducted an investigation about the transferability of Gravi to other content areas (as e.g. other forms of psychotherapy, history, teaching, etc.). We especially asked content experts from these other areas whether they could think of concrete examples how to use Gravi in their area of expertise and discussed these examples with them. As far as we know, no other study of this kind has been conducted. This study was especially interesting because we could also test Gravi with content experts in similar domains as the one it was originally intended for (that is, other forms of psychotherapy).

An important aspect of this transferability study was also to assess the usability of Gravi with content experts (as opposed to the original usability study which had been conducted with students). Such studies can yield very interesting information about the usability of InfoVis tools gained in more realistic contexts than studies with students as subjects. But there are problems as well. One of the main problems of usability testing with content experts is their lack of time. This was also obvious in this study. The experts did not have enough time to really get acquainted with the system, and the results are, therefore, not as detailed as the ones from the usability study with the students who got a long introduction to the system and had enough time to try out the system on their own. Nevertheless, we think that such studies can be valuable if combined with other usability studies because they increase the ecological validity of the overall evaluation of the system.

The transferability study was not only restricted to experts in medicine but also included experts from other areas with similar problems (historians, teachers, etc), that is, complex, sometimes time oriented data and ill-structured knowledge domains. All in all, there were 14 content experts participating in this study. In this paper, we concentrate on the experts from the medical area. This group consists of five subjects, all of whom have a background in psychotherapy and have at least passing knowledge of the domain for which Gravi was developed. These subjects had varying degree of previous experience with computers.

We decided to conduct interviews with the subjects. Because of the time constraints of the subjects, a longer interview and a more rigorous procedure of investigation (e.g. thinking aloud) was not possible. We developed a framework of questions for the interview (for a description of qualitative data analysis see e.g. [24]). The aim of this framework was, on the one hand, to get information about the usability of the finished system and, on the other hand, to investigate whether Gravi could be transferred to other domains. The questions asked are described in the next paragraph. The questions asked in the interviews were also used as categories for the interpretation of the interviews. In addition, we also looked for indications for major usability problems. An indication for a usability problem might, e.g., be if users could not understand a feature of the system after two or three explanations and made the same mistake again and again.

Table 1. Guideline for transferability interviews

Topic	Time Target
1. Introducing interviewers, research project, research question, and interview procedure. Get permission to record interviews.	5 minutes
2. Brief presentation of application domain: <ul style="list-style-type: none"> • Anorexia nervosa: definition, causes, consequences, and treatment of disease • Abbreviations and meanings of used questionnaires (e.g., BDI – Beck’s Depression Index) Initial training (i.e. presentation of most basic features and interactions with Gravi): <ul style="list-style-type: none"> • Drag & Drop • Highlight • Add and remove of patient icons and questionnaire icons • Color-coding of icons (e.g., green – positive therapy outcome) • Transparency of icons (all data is missing) • Tooltips for details on patient and questionnaires • Spring metaphor of visualization for positioning the patients’ icons • Time-dependent data and navigation through time steps 	10 minutes
3. Scenario 1: What is the influence of 5 given questionnaires (i.e. the answering behavior of the patients on these) on the therapy outcome? Visualization is already operated by subjects, although assistance is provided whenever needed.	15 minutes
4. Scenario 2: What statements can be made about patient 18? What is the predicted therapy outcome of this so far unclassified patient? (See Fig. 6) Visualization is operated by subject without any support.	10 minutes
5. Presentation of further features and visualization options of Gravi (e.g., traces for visualizing paths of patients’ icons over all time steps)	5 minutes
6. Questioning of subjects on usability and utility of Gravi: <ul style="list-style-type: none"> • Was the general impression of Gravi positive or negative? (Give reasons) • Was the system understandable? (Give reasons) • Are there any concerns about shortcomings of the visualization? • What was the best feature of Gravi? • How was the learnability of Gravi? Questioning of subjects on transferability of Gravi: <ul style="list-style-type: none"> • What are possible fields of application in the subjects’ domains? Sketch concrete examples if possible. • Is the visualization technique suitable for the data of the subjects’ domains or are improvements and/or modifications necessary? 	15 minutes

The experiment for the transferability study lasted approx. one hour. In the beginning, the subjects were introduced to Gravi. This usually took about 30 minutes. It should be mentioned that this gave subjects only a superficial impression of the system. Then, the subjects had to solve a very easy example on their own to get a better understanding of how the system worked. Then the subjects were asked how

confident they were about the solution they found, which positive and negative features of the system they could identify and whether the system was easy to use. The last question was whether they could think of a concrete application area for Gravi in their own domain. In the guideline for the interviews (see Table 1), usability aspects were explicitly addressed. In addition, usability issues could also be derived from the interaction of the subjects with the system.

Subjects were tested in their own offices (which meant that they were sometimes disturbed by phone calls etc). The same laptop was used for all tests. The interviews were recorded and then transcribed.

The simple task the subjects had to solve during the experiment was to predict success or failure of a patient from the data of two time-steps (instead of five time-steps for the patients who had finished the therapy). See Fig. 6 for a prototypical visualization configuration to answer this question.

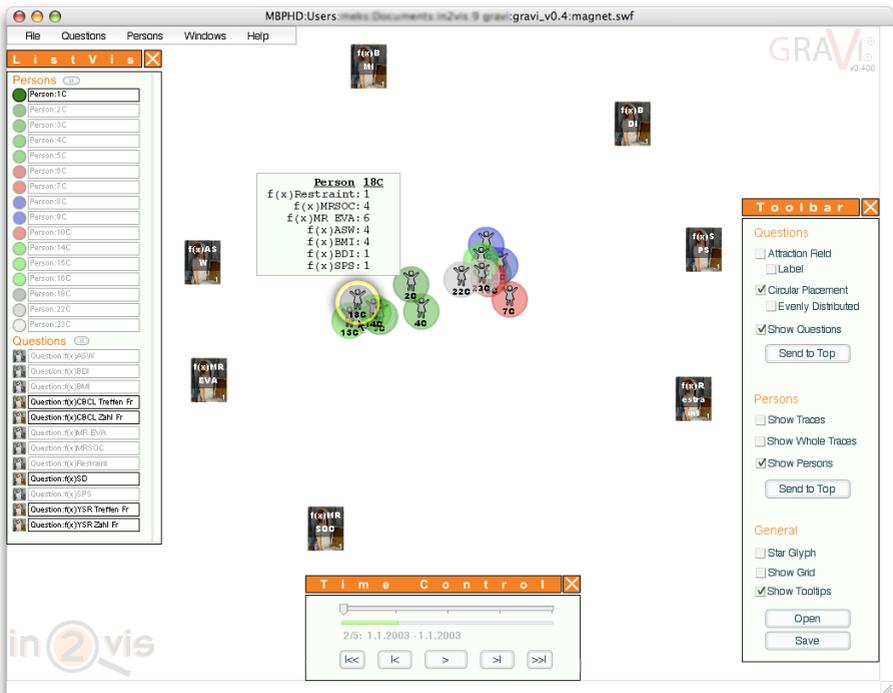


Fig. 6. Prototypical visualization configuration to predict the therapy outcome of patient 18 (highlighted). The position of the icon of patient 18 amongst patients with positive outcome allows for the hypothesis that patient 18 will also have an positive therapy outcome.

Four of the five subjects could accomplish this task. One subject felt that s/he had not enough data. This subject also felt insecure about the use of computer technology in general. This fact influenced her/his understanding of the system negatively. This subject also had tight time constraints which made the situation of the interview fairly

stressful. It should be mentioned that, despite this, the subject formulated a very interesting example from her own domain. It is probably also characteristic that this subject preferred more static features of the system to the more dynamic ones (that is, representation of dynamic data).

Of the four subjects who solved the example, only three gave a confidence rating (two of them high confidence and one medium confidence). The fourth subject mentioned that a reliable confidence rating could only be given when extensive information about the subject domains and especially the behavior of a large enough sample of other patients was available.

The answers relating to the most positive feature of the system were rather heterogeneous. Two subjects thought that the dynamic presentation of time oriented data were very advantageous. Two other subjects thought that the circular positioning of the icons for the questions was very positive. Apparently, this gives a compact impression of the data. One subject mentioned that s/he found that the visual presentation of data in general was very positive.

Several negative features were mentioned. The system quickly becomes confusing when many questions and persons are added. A good introduction to the system is necessary to avoid technical problems. One subject found it difficult to remember information from one time-step to the next. Problems of a more epistemological nature were also mentioned. One subject found that positioning the questions in an arbitrary manner on a circle is problematic. S/he thought that the system should suggest an appropriate position. Another subject thought that the level of aggregation of data might be too high so that important information might get lost. Yet another subject indicated that the system only works well if one can be sure to have identified all necessary variables for predicting success or failure of the therapy.

In general, the subjects found the system quite easy to use. Two subjects said that a good introduction was necessary. As mentioned before, one subject had difficulties to use the system (probably because of the lack of previous experience with computers and the lack of time). The protocols of the interviews which also contain the process of the introduction show no other serious difficulties in the usage of the system which indicates that the subjects' subjective impressions and their actual behavior correspond, which is not always the case in usability research.

The last question was that subjects should find an example from their own work experience which could be supported by Gravi. These examples were very heterogeneous in level of detail. To a certain extent, most of the examples were very similar to the original application domain as all subjects had a background in psychotherapy. In general, all subjects suggested the application of Gravi for the support and analysis of psychotherapy but some of them indicated that their data might not be appropriate (e.g. too many subjects, too few variables; too heterogeneous samples). One subject thought that Gravi could only be used for the analysis of group psychotherapy but another subject thought it possible to compare all her/his patients using Gravi. One subject suggested that the source of the data do not have to be questionnaires but could also be ratings made by the therapists themselves. Two subjects also suggested an application of Gravi to support university teaching. One of these examples was quite elaborate. Students should be rated according to their ability in conducting the first interview with prospective patients. The data resulting from

this rating process should be analyzed by Gravi. This analysis was supposed to improve the teaching process and give teachers more insights into what could go wrong.

5 Conclusion

InfoVis techniques can offer a valuable contribution for the examination of medical data. We successfully developed an InfoVis application – Gravi – for the analysis of questionnaire data stemming from the therapy of anorectic young women. During the development process, we carefully evaluated Gravi in several stages. In this paper, we describe selected results from the usability evaluation, especially results from qualitative studies. In a first stage, usability was evaluated using students as subjects. These students were semi-experts in the area of usability. Therefore, they could give more precise feedback concerning usability issues. Based on the results of this study, Gravi was improved considerably. The results of the transferability study, using content experts as subjects, is an indication of this. The content experts all found the system easy to use, and of a total of 14 all but one could solve a small example with Gravi on their own after a short introduction.

It is slightly difficult to compare the results of the focus groups and the interviews with the content experts because the students concentrated on usability issues whereas the answers of the content experts also cover other issues apart from usability aspects. There is some indication that the presentation of time-oriented data and the interactivity of the system (drag & drop, instant updates of the screen etc.) is seen as positive by members of both groups. The answers concerning the negative features of the system are very heterogeneous. The content experts often mentioned problems of a more epistemological nature which cannot be called usability problems.

It seems that Gravi can easily be used for other application areas in medicine. The subjects especially mentioned other uses for the analysis of therapy processes. One problem which has to be solved in this context is the question whether Gravi can only be used for group therapy or also for individual therapy.

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]. Thanks to Bernhard Meyer for the collaboration in conducting the transferability interviews.

References

1. Chittaro, L.: Information Visualization and its Application to Medicine. *Artificial Intelligence in Medicine* 22(2), 81–88 (2001)
2. Plaisant, C.: The Challenge of Information Visualization Evaluation. In: *Proc. AVI 2004*, pp. 109–116. ACM Press, New York (2004)
3. Tory, M., Möller, T.: Human Factors in Visualization Research. *IEEE Transactions on Visualization and Computer Graphics* 10(1), 72–84 (2004)

4. Holzinger, A.: Usability Engineering Methods for Software Developers. *Communications of the ACM* 48(1), 71–74 (2005)
5. Holzinger, A.: Application of Rapid Prototyping to the User Interface Development for a Virtual Medical Campus. *IEEE Software* 21(1), 92–99 (2004)
6. Wiltgen, M., Holzinger, A., Groell, R., Wolf, G., Habermann, W.: Usability of image fusion: optimal opacification of vessels and squamous cell carcinoma in CT scans. *Springer Elektrotechnik & Informationstechnik, e&i* 123(4), 156–162 (2006)
7. Holzinger, A., Geierhofer, R., Ackerl, S., Searle, G.: CARDIAC at VIEW: The User Centered Development of a new Medical Image Viewer. In: Zara, J., Sloup, J. (eds.) *Central European Multimedia and Virtual Reality Conference*, pp. 63–68, Czech Technical University, Prague (2005)
8. Ware, C.: The Foundation of Visual Thinking. In: Tergan, S.-O., Keller, T. (eds.) *Knowledge and Information Visualization. Searching for Synergies*, pp. 27–35. Springer, Heidelberg (2005)
9. Pirolli, P.: Exploring and Finding Information. In: Carroll, J.M. (ed.) *HCI Models, Theories, and Frameworks. Toward a Multidisciplinary Science*, pp. 157–191. Morgan Kaufmann, San Francisco (2003)
10. Russell, D.M., Stefik, M.J., Pirolli, P., Card, S.: The cost structure of sensemaking. In: *Proceedings of the INTERCHI 93, ACM Conference on Human Factors in Computing Systems*, ACM Press, New York (1993)
11. Card, S.: Information Visualization. In: Jacko, J.A., Sears, A. (eds.) *The Human-Computer Interaction Handbook. Fundamentals, Evolving Technologies and Emerging Applications*, pp. 544–582. Lawrence Erlbaum, Mahwah (2003)
12. Shneiderman, B.: Creating Creativity: User Interfaces for Supporting Innovation. In: Carroll, J.M. (ed.) *Human-Computer Interaction in the New Millenium*, pp. 235–258. Addison-Wesley, San Francisco (2002)
13. Marchionini, G.: Exploratory Search: From Finding to Understanding. *Communications of the ACM* 48(4), 41–46 (2006)
14. Bertini, E., Plaisant, C., Santucci, G. (eds.) *BELIV 2006. Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaLuation methods for Information Visualization*, p. 3. ACM Press, New York (2006)
15. Rester, M., Pohl, M., Hinum, K., Miksch, S., Popow, C., Ohmann, S., Banovic, S.: Assessing the Usability of an Interactive Information Visualization Method as the First Step of a Sustainable Evaluation. In: Holzinger, A., Weidmann, K.-H. (eds.) *Empowering Software Quality: How can Usability Engineering reach this goal?* Wien: Oesterreichische Computergesellschaft, pp. 31–43 (2005)
16. Rester, M., Pohl, M., Wiltner, S., Hinum, K., Miksch, S., Popow, C., Ohmann, S.: Evaluating an InfoVis Technique Using Insight Reports. In: *Proceedings of the IV 2007 Conference*, pp. 693–700 (2007)
17. Rester, M., Pohl, M., Wiltner, S., Hinum, K., Miksch, S., Popow, C., Ohmann, S.: Mixing Evaluation Methods for Assessing the Utility of an Interactive InfoVis Technique. In: Jacko, J. (ed.) *HCI 2007. Human-Computer Interaction. Interaction Design and Usability*. LNCS, vol. 4550, pp. 604–613. Springer, Heidelberg (2007)
18. Hinum, K., Miksch, S., Aigner, W., Ohmann, S., Popow, C., Pohl, M., Rester, M.: Gravi++: Interactive information visualization to explore highly structured temporal data. *Journal of Universal Comp. Science* 11(11), 1792–1805 (2005)
19. Hinum, K.: Gravi++ – An Interactive Information Visualization for High Dimensional, Temporal Data. PhD thesis, Institute of Software Technology & Interactive Systems, Vienna University of Technology (2006)

20. Nielsen. Heuristic Evaluation. In: Usability Inspection Methods, vol. ch.2, pp. 25–62. John Wiley & Sons, Chichester (1994)
21. Kuniavsky, M.: User Experience: A Practitioner's Guide for User Research, p. 201. Morgan Kaufmann, San Francisco (2003)
22. Mazza, R.: Evaluating information visualization applications with focus groups: the coursevis experience. In: BELIV 2006. Proc. of the 2006 AVI workshop on BEyond time and errors: novel evaLuation methods for InfoVis, ACM Press, New York (2006)
23. Mazza, R., Berre, A.: Focus group methodology for evaluating information visualization techniques and tools. In: IV 2007. Proc. of the 11th International Conference Information Visualization, pp. 74–80. IEEE Computer Society, Los Alamitos (2007)
24. Bortz, J., Döring, N.: Forschungsmethoden und Evaluation, 4th edn. Springer, Heidelberg (2006)