

# Evaluating Information Visualization Applications with Focus Groups: the CourseVis experience

Riccardo Mazza

Faculty of Communication Sciences

University of Lugano

Via Buffi, 13 CH-6900 Lugano - CH

mazzar@lu.unisi.ch

## ABSTRACT

This paper reports our experience of evaluating an application that uses visualization approaches to support instructors in Web based distance education. The evaluation took place in three stages: a focus group, an experimental study, and a semi-structured interview. In this paper we focus our attention on the focus group, and we will show how this evaluation approach can be very effective in uncovering unexpected problems that cannot be identified with analytic evaluations or controlled experiments.

## Categories and Subject Descriptors

H.5.2 [Information interfaces and presentation] User Interfaces -Evaluation/methodology, Graphical user interfaces (GUI)

## General Terms

Human Factors, Verification.

## Keywords

Focus group, Human factors, Information visualization evaluation.

## 1. INTRODUCTION

Evaluating an information visualization system (InfoVis), as other user-centred applications, is a challenging task that is quite different from evaluating systems in other areas of Computer Science, mainly because it is sometimes hard to construct experiments or observations that give definitive quantitative answers regarding a particular visualization [5][8]. Evaluation of user-centred tools often makes use of analytic and empirical techniques with the objective to investigate the usability, the functionality, and the effectiveness of an interactive system [5]. Chaomei Chen on his inspiring article on top 10 unsolved information visualization problems [3] put as the first three problems issues related to human factors (namely: usability issues, understanding elementary perceptual-cognitive tasks, prior knowledge), highlighting that there is still a lot of work to do on defining evaluation methods that involve real users and perceptual-cognitive tasks.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

BELIV 2006 Venice, Italy.

Copyright 2006 ACM 1-59593-562-2/06/05 ...\$5.00.

Current evaluation practices are based on methodologies established in Human Computer Interaction (HCI). Plaisant [17] cites four main areas where HCI techniques have been used with InfoVis: controlled experiments comparing design elements, usability evaluations, controlled experiments comparing two or more tools, case studies. However, such methodologies have been criticised by some authors [3][17] because they don't address the specific InfoVis applications needs. For instance, a key point on InfoVis is to understand whether users can comprehend the meanings of the underlying data, recognise patterns, or identify salient structures. Although we agree with these authors, we have to acknowledge that some user studies may be still appropriate to evaluate some aspects of visualizations such as the usefulness of the graphical representations.

HCI evaluation methods fall into two types: analytic evaluations and empirical evaluations.

*Analytic evaluation methods* come from psychological models of human information processing and are based on studies of human cognition and behaviour. They are performed with expert-based methods such as *heuristics evaluations* (where an expert evaluates an interface and judge its compliance with recognized usability principles called "heuristics"), or *cognitive walkthroughs* (where an expert walks through a specific task using a prototype system, thinking carefully about potential problems that could occur at each step) [18]. They are also used to evaluate usability and accessibility issues. One of the most known models is GOMS (Goals, Operations, Methods and Selection of rules) proposed by Card et al. [2]. This is a class of evaluation techniques for modelling and describing human task performance. Analytic evaluations usually occur during the system's design and are oriented to identify problems and guide modifications during the development of a system.

*Empirical evaluation methods* (also known as *user studies*) involve real users in the study and allow designers to obtain qualitative and quantitative data [18]. Usually they are performed with system already implemented (in form of prototypes or demonstrators), as they are suitable to make formal claims. Empirical evaluations can be further distinguished between quantitative studies and qualitative studies.

*Quantitative studies* consist of an analysis of determinate hypotheses tested through direct measurements [5]. Examples of hypotheses can be the user's performance in relation to a specific task, or the number of trials required in order to accomplish a specific task. This requires the definition of one or more variables related to the hypotheses examined and a metric associate to each of them. The evaluation is carried out usually by the means of *controlled experiments* (also known as *experimental studies*) [8]. They consist of asking the user testers to run a task and

performing some measurements using observation, and completing the study with questionnaires or interviews.

*Qualitative studies* can also be quite useful for empirical evaluation of user-centred systems. Qualitative research involves the analysis of qualitative data, which may be obtained through questionnaires, interviews and observing users using the system, to understand and explain social phenomena. They are opposite to quantitative methods used in experimental studies for their ability to analyse phenomenon from the point of view of the participants that it is largely lost when textual or analytical data are quantified [9]. Focus group [10][7][15] is a form of qualitative research which involve group interviewing of individuals selected and assembled by researchers to discuss and comment on, from personal experience, the topic that is the subject of the research [6]. This is less formal than controlled experimental study, but has the advantage of getting the user's viewpoint directly and "may reveal issues which have not been considered by the designer" [6](p.431). Additional qualitative studies include semi-structured interviews that can be conducted with users, asking specific questions to elicit information about users' impression and general comments. In InfoVis both focus groups and interviews can be appropriate to detect the users' opinion about the usefulness of the proposed graphical representations. The combination of both qualitative (e.g. focus group and individual interviews) and quantitative methodologies (e.g. experimental study) is known as *cross examination* or *triangulation* [10] and it is appropriate to InfoVis applications because it allows to examine data gathered with different and complementary ways, establish commonalities or differences and to provide rigour to the study [4].

## 2. COURSEVIS

Course Management Systems (CMS) have been widely used nowadays for their great potential to enable interactive web-based teaching and to support the administration of distance courses. CMS are used to distribute information and content material to students, prepare assignments and tests, engage in discussions, and manage distance classes without any time and space restrictions [1]. Although these systems support many tasks related to teaching at distance, the instructors still face a number of problems with managing distance courses effectively, most of which are brought by the difficulty to gain sufficient understanding of social, cognitive, and behavioral aspects of distance students. For these reasons we designed CourseVis, a visual student tracking tool that obtains tracking data from a CMS (in particular, our prototype is based on the WebCT [19] platform), transforms the data into a form convenient for processing, and employ InfoVis techniques to produce various graphical representations that can be explored and manipulated by course instructors to examine social, cognitive, and behavioral aspects of distance students. Describing the CourseVis research is out of the scope of this paper. A detailed description is available in [11][13][14]. In order to provide the flavour of the graphical representations obtained with CourseVis, and to contextualize the evaluations, three visualizations that represent data from courses collected in real settings are showed in Figures 1,2, and 3.

Figure 1 shows the students' accesses to the course. The student access plot in CourseVis is a composition of a scatterplot and a histogram. The scatterplot represents bivariate data of students and dates of the course, where a corresponding bullet represents at least one access to the course made by the student on the corresponding date. The histogram represents the number of pages

of the course accessed by all students on a particular date. The instructor has an overview of the global student access to the course with a clear identification of patterns and trends.

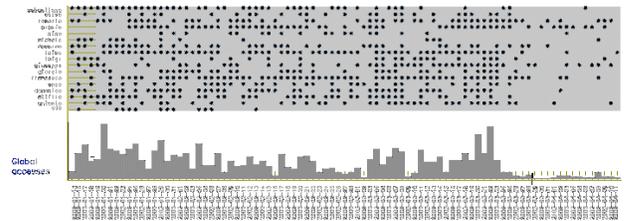


Figure 1: Accesses plot (© 2005 IEEE).

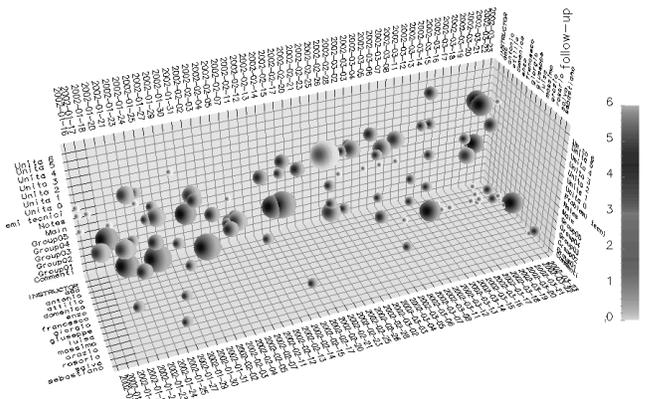


Figure 2: Discussion plot

Summary of student's behaviours  
from 2002-01-15 to 2002-04-11  
Student: "francesco"

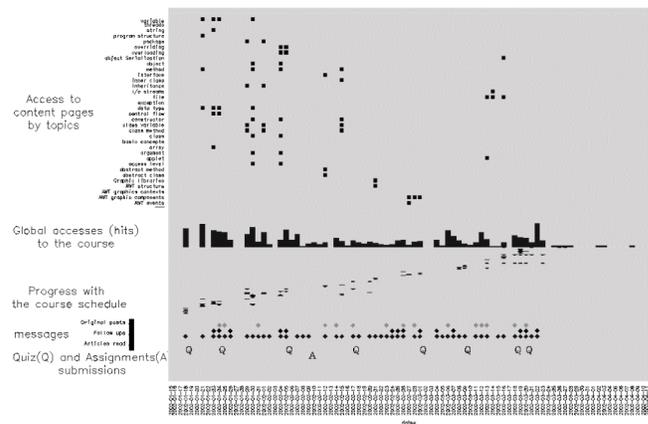


Figure 3: Behavior graph (© 2005 IEEE).

In Figure 2 a 3D scatterplot represents data about discussions throughout the course. The dimensions of the plot are time, topics and students. A sphere represents a new thread of discussion, initiated by the student on a specified date and topic. The follow-up in a thread (the number of messages that compose the whole thread) is encoded with the color and the size of the sphere.

Figure 3 is a comprehensive image that represents in a single view the behaviors of a specific student in an on-line course. It takes

advantage of single-axis composition method for presenting large number of variables in a 2D metric space. With a common x-axis mapping the dates of the course, a number of variables are represented. The information represented here are namely the student's access to the content pages (ordered by topics of the course), the global access to the course (accesses to content pages, quiz, discussion etc.), a progress with the schedule of the course (if the content of the course is organised in a linear sequence, we put in the y-axis the page ordering according to this sequence, a mark then represents a page accessed by the student on a day, where the y-location of the mark is proportional to the sequence of the pages in the schedule, first page on bottom, last page on top), messages (posted, read, follow-up), and the submission of quizzes and assignments.

### 3. THE EVALUATION

CourseVis is an instructor-centered tool, hence it is important to evaluate how this system impacts its intended users. An empirical evaluation of CourseVis was conducted [11] to evaluate whether those representations comply with the instructors' requirements, and whether instructors may take advantage of the proposed representations. The CourseVis evaluation focused on its *effectiveness* (can it help instructors gain an understanding of what is happening in distance classes), *efficiency* (can instructors infer required information quickly), and *usefulness* (to what extent the information provided is useful to the instructors). The evaluation involved instructors with experience in using CMS in distance learning, and took place in three stages.

- 1) A *focus group*, conducted with five instructors who were provided with CourseVis.
- 2) A *controlled experiment*, conducted with six instructors with experience in distance learning.
- 3) Finally, *semi-structured interviews* were conducted with participants involved in the controlled experiment.

The focus group and the semi-structured interview aimed to assess the usefulness of the graphical representations generated in CourseVis, while the experimental study aimed at verifying the effectiveness and the efficiency of the representations.

In this paper we focus our attention to the focus group, showing how this evaluation approach can be very effective in uncovering unexpected problems that cannot be identified with controlled experiments and structured interviews.

#### 3.1 The focus group

Focus group is a technique that can help to investigate user attitudes, feelings, and beliefs on proposed representations with a group interview [6][7][16]. It is conducted by bringing together some representative users to discuss issues and concerns about the features of the system being evaluated. The discussion is lead by a moderator whose role is to facilitate the interaction between group participants and to keep them focused on the topic of investigation. This investigation technique is useful because "*attitudes, feelings and beliefs ... are more likely to be revealed via the social gathering and the interaction which being in a focus group entails*" [6]. Compared to other investigation methods, such as observation, interviews or surveys, focus groups are particularly suitable for the challenges specific to InfoVis, such as "*formulate and answer questions [users] didn't anticipate having before looking at the visualization*" [17] (p.111).

#### 3.1.1 Procedure

We performed a focus group with the aim to identify problems with the representations generated in CourseVis and assess their usefulness. It involved five people with experience on on-line learning or computer-based educational systems. The selected participants had a complementary background, to avoid that diverse opinion and experiences may not be revealed with a group too homogeneous with regard to specific characteristics [6]. The experimenter acted as a moderator for the discussion. The participants were given an introductory explanation about the aim of CourseVis, about the purpose of the study and their role. Then, each graphical representation was displayed to the group on a wide plasma screen and the participants were asked to discuss that representation. The moderator made the interactions with the CourseVis system, i.e. the participants did not use the interface, they just looked at the graphical representations showed on the wide screen and requested at times appropriate manipulations (zooming, rotating).

To promote a debate, the moderator asked some open questions, e.g.

- What kind of information can you gather from these images, what can you understand about the students and the class as a whole?
- Can you identify general patterns or tendencies?
- Is the representation useful, and why? Is it interesting to you? Would you be interested to have this information displayed?
- Do you think it must be presented differently and how?
- Is there anything else which can be useful that is not captured in this representation?
- What kind of activities may you use this representation for?

The session lasted about one hour and was recorded on videotape. The videotape helped the data analysis, as it provided a complete record of the session, capturing all verbal and nonverbal behavior. The recording was analyzed by two people with complementary backgrounds, and the results will be explained in the next section. Participants were informed that the discussion was recorded and were assured that it would be used confidentially for the sole purpose of evaluating and improving CourseVis. The focus group revealed some problems concerning details of the representations that were judged as misleading from all members of the group. These problems were fixed in the subsequent version of the tool.

We have to stress some limitations of the focus group. In literature there is no universal indication about the recommended number of people per group. Even though some authors give precise suggestions about the group size (Krueger [10] recommended 7 to 10, Morgan [15] suggested as a range of acceptable size from 4 to 12) we have to acknowledge that the few number of participants to the group led to a partial analysis. The number of groups is questionable as well. We used only one group, while some experts suggest that the number of groups should be at least two [15] or three [10]. The difficulties in finding instructors available for a group study led us restrict our analysis to only one group. Moreover, involved individuals belong to specific context, within

a specific culture (although there were representatives from two countries, UK and USA). In particular, all of them teach courses at a university level. Their point of view may differ from others teaching in different level, such as in secondary school.

### 3.1.2 Results

Each graphical representation illustrated in Section 2 was proposed to the group and was subject to discussion. A summary of predominant statements and relevant participant thoughts is provided here. Results are divided for representations to help the reader to contextualize the discussion on a particular image.

#### Accesses plot (Figure 1)

All participants expressed interest in having this graphical representation in their courses. They would use this graphical representation to identify students' accesses and find early problems, as one of the participant commented:

*"If I was running this course I would want to use it in order to see if a particular student hadn't accessed in to very much, to find this early on"* .

It was their opinion that this representation could give instructors the opportunity to have a clear picture of students who are actively participating in the course and those who are not. Instructors could encourage students not very persistent to be more active, with the intent to do appropriate activities before it is too late:

*"you want to provoke them by e-mail at some point and say: look, you've got a week free and there is no access"*,

and

*"I think that the image helps me to realize that something may be wrong. I should be doing something before it is too late"* .

One participant argued that some instructors may be more interested in having not only the information about the access on a date, but also how long students accessed the course, how often, and even the period of the day (e.g. morning, afternoon, night). It was pointed out that one dot may represent a number of accesses and this may be misleading. However, other participants thought that, even without this information, this graphical representation is useful to assess some aspects of students' participation.

Participants identified some patterns looking at the dots, which referred to the group exercises given at 3/4 of the course and the termination of the course, which occurred 15 days before last date reported in the graph.

Participants made also some suggestions to improve the graphical representation: (a) to put a scale on the global access histogram to better clarify what it represents; (b) to indicate week-ends and holidays with different colors on dates and mark topics of study on dates to see if some topics required long time to work on them. Future development of CourseVis could investigate how the representations could incorporate these recommendations.

#### Discussion plot (Figure 2)

A participant's first impression when looked at this image was:

*"is like having an awful lot of information, initially overwhelming what you really need to"* ,

due to the many dimensions represented. This proved that, at a first stage, the 3D scatterplot created some confusion. One participant commented:

*"It looks to me like a swimming pool: it is too complex to use. It's difficult to locate the exact positions of the balls"* .

Predominant criticisms were: (a) the difficulty in locating the exact position of the balls with respect to the axis, (b) the lack of information on who is contributing in reading the discussions, (c) the rotation of the scatterplot was considered too complicated to decide to what different angle to position it. In fact, the positioning of the scatterplot that hides one of the dimensions (and which transforms the 3D image in a 2D scatterplot) was considered the only usable. But, at the same time, the same participant who criticized the rotation of the scatterplot appreciated the flexibility given by the graphical representation to exploring the dots

*"to find something interesting [...] because you never know what you need to see"* .

This graphical representation was appreciated because it may help instructors identify discussions that need attention, to see further what the discussion was:

*"If I was the instructor and I had the discussion ball in front of me, I would want to look at some of these and I would focus on the red ball for instance. The image may help me focusing on which piece of discussion I want to look at and analyze"* ,

but also:

*"I'd look at the little one, the once having no follow-ups, to see what's wrong with those time periods what's going on in that discussion, why it is not exciting, so I can change that, or take part as appropriate"* .

The connection with the topics of discussion could be useful to infer particular information:

*"I noticed that the technical problems discussion topic are at the beginning, they would be very helpful to me. When you are looking at the students to see if they find problems I can see that they know what to do now, as the technical problems seem to have been solved"* .

Participants made some suggestions to improve the interaction with the graphical representation. In particular, they would found useful to click on a particular sphere and open the discussion threads, rather than going to the whole discussions and try to search the relative thread.

#### Behavior graph (Figure 3).

Participants found it very useful to have the possibility to monitor individual students, and most of them appreciated it:

*"I think that the green stuff [the access to content pages by topics] and the blue stuff [the global access to the course] are very useful."* ,

but also, so much information may not be needed:

*"It is interesting to have an overview with this kind of picture on the student, but it also has so much information that I would not need" ,*

and:

*"I think it is best to just provide as little information as clearly as possible. If you provide 2 or 3 variables at time, you are probably more successful than trying to compete and provide as much information as possible" .*

Some participants found very difficult to look at the access to content pages to discover if a student accessed the material on every concept:

*"If there is one thing I will be looking for here is how the student access material on every concept, as it is very hard for me to tell looking at this. I see the access to a lot of things, I've got to study the lines carefully to working on it." .*

A suggested solution to this problem was to put a summary on the right part of student's access to the content pages to immediately represent what pages are accessed, which is recommended in future development of CourseVis.

Another suggestion made by participants was to modify the quizzes representation to show also the score achieved by each student, so that a better picture of the students' progress is extracted. As a future implementation the participants suggested to substitute "Q"s with the score, or to put the score next to "Q"s

#### 4. CONCLUDING REMARKS

We ran a three stages evaluation of CourseVis with the objective of understanding whether the target users may take advantage of the proposed graphical representation, and to discover any problems and possible improvements that may be taken into consideration in future developments of the tool.

A controlled experiment validated a set of hypotheses in controlled conditions (in particular, we measured the time required to accomplish a specific task and the accuracy). A semi-structured interview on topics regarding the performed tasks was conducted with subjects involved in the controlled experiment. The focus group was the only method that brought to us unexpected problems that haven't been revealed with the controlled experiment and semi-structured interview. Among the criteria addressed (effectiveness, efficiency, and usefulness), usefulness was the most interesting to our purposes. Even if particular information is provided in an effective and efficient way, instructors will have a *real* advantage only if the information provided is useful to them. The focus group and, to some extent, the semi-structured interview have proved to be suitable on investigating this issue. Focus group revealed some situations where participants regarded the information provided by the graphical representations as useful, but also uncovered some potential problems. The problems were related with missing values, confusion with rotation, and difficulty to examine too many variables at the same time. Improvements were also suggested, such as providing flexible links between the graphical representations and the corresponding data from the CMS, and enabling instructors to filter and manipulate data to be included or excluded from graphics. These problems and suggestions have been addressed in a next release of the tool, in the successor of CourseVis - GISMO [12] a visualisation tool for student

monitoring developed within the European Union funded EdukaLibre project.

#### 5. REFERENCES

- [1] Avgeriou, P., Papasalouros, A., Retalis, S., and Skordalakis, M. Towards a pattern language for learning management systems. *Educational Technology & Society*, 6(2):11-24. <http://ifets.ieee.org/periodical/6-2/2.html>. 2003.
- [2] Card, S. K., Moran, T. P., and Newell, A. *The Psychology of Human-Computer Interaction*. Lawrence Erlbaum Associates. 1983
- [3] Chen C. Top 10 Unsolved Information Visualization Problems. *IEEE Computer Graphics and Applications*. 25 (4). pp.12-16. 2005.
- [4] Denzin, N. K. and Lincoln, Y. S. *Collecting and interpreting qualitative materials*. Sage, Thousand Oaks, CA. 1998.
- [5] Dix, A., Finlay, J., Abowd, G., and Beale, R. *Human-Computer Interaction*. Prentice Hall, second edition. 1998.
- [6] Gibbs, A. Focus groups. *Social Research Update*, University of Surrey, UK. Num. 19. <http://www.soc.surrey.ac.uk/sru/SRU19.html>. 1997.
- [7] Greenbaum, T. L. *The Handbook for Focus Group Research*. SAGE publications, CA, USA. 1998
- [8] Johnson, P. *Human-computer interaction: psychology, task analysis and software engineering*. McGraw-Hill, London. 1992
- [9] Kaplan, B. and Maxwell, J. A. Qualitative research methods for evaluating computer information systems. In Anderson, J. G., Aydin, C., and Jay, S. J., editors, *Valuating Health Care Information Systems: Methods and Applications*, pages 45 - 68. Sage, Thousand Oaks, CA, 1994.
- [10] Krueger, R. A. *Focus Groups: A Practical Guide for Applied Research*. Third Edition. Sage Publishing, Newbury Park, CA. 2000.
- [11] Mazza R. *Using Information Visualisation to Facilitate Instructors in Web-based Distance Learning*. PhD thesis, University of Lugano, 2004.
- [12] Mazza R. and Milani C. Exploring Usage Analysis in Learning Systems: Gaining Insights From Visualisations. In: *Workshop on Usage analysis in learning systems. 12th International Conference on Artificial Intelligence in Education (AIED 2005)*. Amsterdam, The Netherlands. 18 July 2005. pp 65-72. 2005.
- [13] Mazza R. and Dimitrova, V. Generation of Graphical Representations of Student Tracking Data in Course Management Systems. In: *9th IEEE International Conference on Information Visualisation*. London 6-8 July 2005.
- [14] Mazza R. and Dimitrova, V. Visualising Student Tracking Data to Support Instructors in Web-Based Distance Education. In: *13th International World Wide Web Conference (WWW 2004) - Alternate Educational Track. 2004 May 17-22*. pp.154-161. 2004
- [15] Morgan, D. L. *Focus group as qualitative research*. Sage, London. 1988.

- [16] Nielsen, J. The use and misuse of focus groups. *IEEE Software*, 14(1). p.94 – 95. 1997.
- [17] Plaisant, C. The challenge of information visualization evaluation. *Proceedings of the working conference on Advanced visual interfaces (AVI04)*. pp. 109 – 116. 2004.
- [18] Tory M. and Möller T. Human Factors in Visualization Research. *IEEE Transaction of Visualization and Computer Graphics*. Vol. 10 N. 1. 2004.
- [19] WebCT. World Wide Web course tool. WebCT educational technologies, Vancouver, Canada. <http://www.webct.com>.