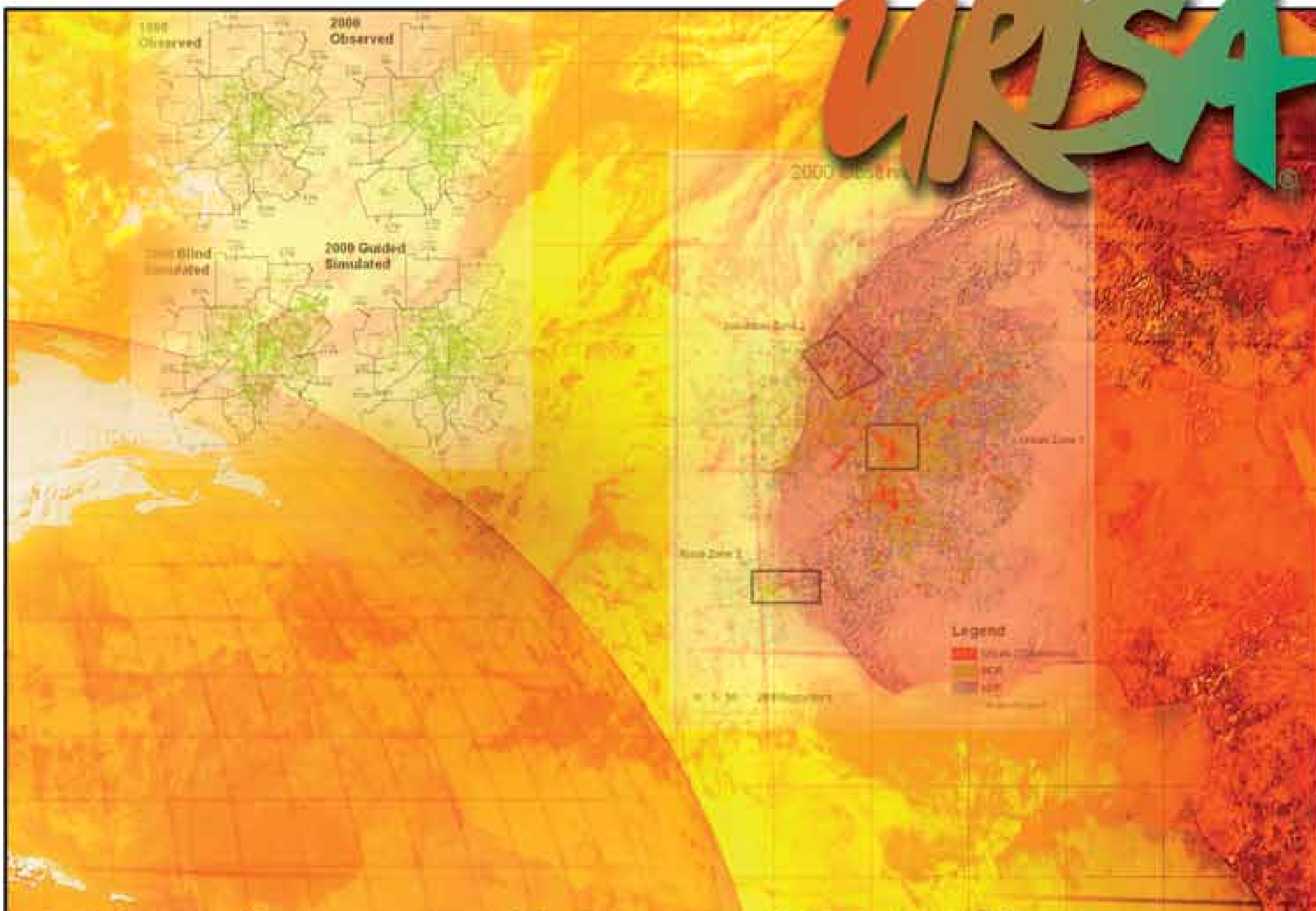
Vol. 22 No. 1
2010

Journal
of the Urban and Regional Information Systems Association

URISA®



Validation and Demonstration of the Prescott Spatial Growth Model in Metropolitan Atlanta, Georgia

ParticipatoryGIS: A Web-based Collaborative GIS and Multicriteria Decision Analysis

Traditional Genetic Algorithm and Random-weighted Genetic Algorithm with GIS to Plan Radio Network

Task-oriented Usability Evaluation of a WebGIS for a Real-World Community

Web-PPGIS Usability and Public Engagement: A Case Study in Canmore, Alberta, Canada

w w w . u r i s a . o r g

Volume 22 • No. 1 • 2010

Journal of the Urban and Regional Information Systems Association

Contents

Refereed

URISA Journal

Publisher: **Urban and Regional Information Systems Association**

Editor-in-Chief: **Jochen Albrecht**

Journal Coordinator: Wendy Nelson

Electronic Journal: http://www.urisa.org/urisajournal

EDITORIAL OFFICE: Urban and Regional Information Systems Association, 701 Lee Street, Suite 680, Des Plaines, Illinois 60016; Voice (847) 824-6300; Fax (847) 824-6363; E-mail info@urisa.org.

SUBMISSIONS: This publication accepts from authors an exclusive right of first publication to their article plus an accompanying grant of non-exclusive full rights. The publisher requires that full credit for first publication in the *URISA Journal* is provided in any subsequent electronic or print publications. For more information, the "Manuscript Submission Guidelines for Refereed Articles" is available on our website, www.urisa.org, or by calling (847) 824-6300.

SUBSCRIPTION AND ADVERTISING: All correspondence about advertising, subscriptions, and URISA memberships should be directed to: Urban and Regional Information Systems Association, 701 Lee Street, Suite 680, Des Plaines, Illinois 60016; Voice (847) 824-6300; Fax (847) 824-6363; E-mail info@urisa.org.

URISA Journal is published two times a year by the Urban and Regional Information Systems Association.

Educational programs planned and presented by URISA provide attendees with relevant and rewarding continuing education experience. However, neither the content (whether written or oral) of any course, seminar, or other presentation, nor the use of a specific product in conjunction therewith, nor the exhibition of any materials by any party coincident with the educational event, should be construed as indicating endorsement or approval of the views presented, the products used, or the materials exhibited by URISA, or by its committees, Special Interest Groups, Chapters, or other commissions.

SUBSCRIPTION RATE: One year: $295 business, libraries, government agencies, and public institutions. Individuals interested in subscriptions should contact URISA for membership information.

US ISSN 1045-8077

Editors and Review Board

Task-oriented Usability Evaluation of a WebGIS for a Real-World Community

Jens Ingensand and François Golay

Abstract: *This paper describes the usability evaluation of a WebGIS. The system under evaluation has been developed for a real-world community in the Swiss canton of Vaud, namely, for people who are involved in wine making. The system's usability is tested with users from this community. The goal of the study is to validate hypotheses and answer research questions focusing on users' performance and specific factors that influence users' performance, such as experience and training or particular interface features. To analyze these hypotheses and research questions, a methodology has been developed to capture the users' interaction during the hands-on evaluation and to analyze the captured data. We found that there are different strategies for the interaction with the WebGIS and that users perform differently depending on their strategies. We believe that these differences in performance are related to the users' experience and expertise in the use of similar systems, but also to specific interface features. The findings presented in this paper are useful for developers and designers of WebGIS and can help to improve these systems.*

INTRODUCTION

In recent years, online geospatial systems for people who are not experts in the domain of spatial information science have emerged. For many of these systems, however, the data representation and the way of interacting with this data were inspired by systems that were created for expert users such as ArcGis, MapInfo, or Manifold.

The usability of these expert systems has been evaluated and criticized by several groups of researchers (Traynor and Williams 1995), but few researchers have yet explored whether data representation and interaction manners, taken from expert systems, are applicable to spatial information systems that are not used by experts.

Despite the fact that major online Web-mapping systems such as Google Maps or Microsoft Live Search have been significantly improved regarding their usability, systems implemented for specific groups of users still lack such improvements. One reason for this problem might be that many of these systems are based on components that already have predefined interaction manners built in and that only offer a certain set of possibilities for data representation. Another reason is that, up until now, almost no specific usability guidelines exist for online geospatial systems that are built for specific groups of users such as real-world communities (Haklay and Zafiri 2008).

Usability engineering as a field within human-computer interaction was founded in the 1980s and has become a well-established discipline since then, yet we argue that usability testing for geospatial systems is different from usability testing of other software applications. The difference is rooted in the fact that both the data displayed in a geospatial system and the interface that allows for spatial interaction are particular: The spatial information that a user is interacting with is an abstract depiction of the real world; the user needs to match this abstract depiction with his or her internal, cognitive map to understand it. Moreover, the interface that allows for interacting with spatial information (e.g., spatial navigation or adding spatial information) should take into account the relationship with the interaction with the real world. These facts stress the importance of both a geospatial system's interface and the depiction of the data itself. Therefore, the usability testing of geospatial systems includes not only the testing of an interface according to common usability measures (such as task completion time or errors), but implicates a deeper understanding of the interaction and the user.

Systems for real-world communities especially are an important research subject and challenge for geospatial science. Furthermore, this field of research needs to be addressed with evaluations involving members of the real-world community rather than with theoretical concepts (Jankowski and Nyerges 2003). Actual evaluations are necessary to find a methodology that proves to work for real-world communities.

Systems implemented for specific groups of users do only as much or as little as they are designed for. Yet each user perceives and uses them differently. These differences depend on several factors, such as specific interface features and previous knowledge and experience. Previous knowledge and experience particularly are reasons why it is important to take user context into account before, during, and after the implementation of a system.

In this paper, we describe the first phase of a research program aimed at responding to these immediate and larger challenges for GIScience and their relationship with nonexpert users as new modes of interaction become possible through new software tools.

To start making progress in a concrete and meaningful way, this research phase focused on a specific real-world application concerning a community of winegrowers in the Swiss canton of Vaud. Responding to spatial knowledge acquisition and spatial planning needs, a Web-delivered geospatial information tool was deployed and evaluated.

We provide the context of this case study and explain the main challenges and needs of the users and how the project can help this real-world community. We identify the main hypotheses and research questions, explain the methods used, what the users' activities were during evaluation, and identify the relevance of these activities to the immediate case study and larger challenges for interactive nonexpert geospatial information tools. Finally, we present the results and conclude with a discussion about the results and their validity and in which direction further research should be driven.

Case Study Context

The context of the case study described in this paper is a system called RIV (réseau interactif en viticulture—interactive network for wine cultivation). RIV is a system targeted for different aspects involved in winegrowing and wine making in Switzerland. It focuses on the spatial aspects of winegrowing (e.g., the location of the parcels (the smallest spatial winegrowing entities), the existing microclimate, the type of soil, etc.) and it is entirely accessible through the Internet.

Wine making in Switzerland usually is on a very small scale (compared to countries such as France or Australia); many winegrowers have only one or two small parcels and they deliver their harvest to small local wine cellars where the wine is made.

Winegrowers are a rather heterogeneous group of users; many winegrowers grow wine as a part-time business. To survive, various winegrowers do have second businesses such as farming or even nonagriculture-related occupations. As a result, numerous users only have a few parcels.

Approximately 7,800 winegrowers are in the Swiss canton of Vaud where RIV was developed and released, according to the winegrowers' association's Web site, and most of them are potential users.

The main idea in creating RIV was developed from a previous project where maps of the microclimate and the soils in the region had been gathered (the "terroir-project," Pythoud and Caloz 2001). At the end of the project, CDs containing static maps in PDF format were sent to the stakeholders. These maps were useful for the participants in the project. However, they did not reach all the winegrowers in the region, and there was no possibility to update those maps or to correlate the information in an interactive way.

A second reason to create RIV was an identified need to develop a tool to help winegrowers manage their parcels by "virtually" assembling parcels and correlating maps from the terroir project with their parcels. To use the system for parcel management, however, the users first have to digitize their parcels on top of aerial photos (with the support of the official cadastre).

RIV was conceived and developed in close collaboration with end users of the system in different cycles, following a user-centered system development approach (Preece et al. 2002). Several prototypes and mock-ups were developed (Ingensand 2006) and showed to end users to make sure that the system met their needs and requirements.

How Can Interactive Tools Help Real-world Communities?

Interactive spatial tools that are conceived and implemented for the public are a rather new phenomenon. Both the availability of spatial data and the increasing possibilities of Internet mapping systems have accelerated this process.

All these systems, however, were developed for unspecified groups of users and thus address as many users as possible. In most cases, the user can choose between different systems and select the one that corresponds best to his or her needs. With specific groups of users (such as the winegrowers in the canton of Vaud), the users' needs and competencies differ from those of the public. Therefore, such systems have to be adapted to their context.

Interactive spatial tools that are tailored to specific needs and competencies can help increase knowledge and also productivity. Groups that are working within a field with a strong connection to land, soil, and climate especially can benefit from such tools through an increased knowledge of their spatial context.

The challenge of developing tools for such groups of users lies not only in the development of the tool itself but also in the acquisition and the preparation of data. Yet the act of collecting spatial data that is adapted to a specific group is a task that can take much time and can be rather expensive. A solution to this problem, used within the context of this case study, is to let the users of this system collect their own spatial data.

To validate any development that has been created for specific groups of users, we need to develop a methodology that helps detect exactly how the users are interacting with such systems.

HYPOTHESES AND RESEARCH QUESTIONS

In this evaluation, we wanted to emphasize questions and hypotheses that will be discussed and validated through the analysis of the tasks the users had to carry out with the system. However, the analysis of the system's overall efficiency and effectiveness also played an important role. These hypotheses are essential issues when evaluating the usability of any interface that represents and interacts with spatial data. Moreover, these hypotheses are also a fundamental part of our question if interactive Web-mapping systems can improve spatial consensus and awareness in real-world community applications.

In our hypotheses, we first analyze user performance. As user performance, we consider measures that are counts of actions and behavior that one can see (Dumas and Redish 1999).

User Behavior/Interface Use Influences

User behavior/interface use significantly influences performance.

- Is it possible to identify different user strategies to solve specific tasks?

- Which strategies result in better performance?
- Is there an identifiable connection between the user's performance and the user's satisfaction?

Specific Interface Features Influence

Specific interface features significantly influence performance.

- Is there any evidence that some interface features cause a higher cognitive load?
- Were there differences in performance that were associated with features of RIV that are common with conventional GIS versus less conventional/interactive tools?

User Experience and Training Influences

User experience and training significantly influence performance.

- Are there identifiable differences between users?
- Is there any evidence to suggest that the user's geospatial technology expertise has an influence on the user's performance and the way the user interacted with the interface?

EVALUATION SETTING

To engage our hypotheses, we planned the evaluation in three parts:

- A questionnaire about the user's education, background, and computer habits;
- The hands-on evaluation; and
- A second questionnaire concerning the usability of the system.

Besides the questions about the user's background, the first questionnaire also contained some questions about the user's experience with other geospatial systems (such as address-finding systems and three-dimensional visualization software).

We conducted the evaluation according to the think-aloud method (also known as verbal protocols) (*Ericsson and Simon 1993)*. Users were encouraged to say what they think, feel, and do. The advantage of this method is that the evaluator receives much qualitative feedback that otherwise would have been uncovered in addition to the quantifiable feedback (Beer et al. 1997). The disadvantage of this method is that it takes more time; furthermore, Jacobsen et al. (1998) have described that the evaluation expert also exerts an influence on the outcome of the evaluation.

USER SELECTION

Users were selected using a database that was offered by the winegrowers' association. All winegrowers who were listed with e-mail addresses in the database were selected as potential test persons (175 people). One hundred of these 175 winegrowers were selected, together with a representative of the winegrowers' association. The invitations were sent in two series (50 and 50) with two months in between. For each of the series, ten winegrowers were ready to evaluate the system. The average age of these users was around 45 years (three winegrowers did not want to reveal their ages).

USER ACTIVITIES

The evaluations took place in an office at EPFL. Two persons—the evaluation expert and the evaluator (the user)—participated at each evaluation.

As mentioned previously, the second part of the evaluation was the actual interaction with the system. The evaluator received a set of tasks that he or she had to solve using the system. All tasks had been discussed with experts in the domain prior to the evaluation to make sure that the proposed scenario reflected the user's common work tasks.

The tasks that had to be solved with the system were:

- Create one parcel on top of aerial images.
- Create at least one more.
- Display one's parcels on a map.
- Display a map of the parcels and the soils.
- Navigate to the Vully region.
- Navigate to one's village.
- Display different layers on the map (the user has the choice between a
- variety of different layers or some predefined compositions of layers).
- Select parcels on the map and save the selection.
- Find parcels at a specific altitude.
- Find parcels located on lime-containing soil.

For each task, the evaluator had to specify the level of difficulty on a scale from 1 to 5. Throughout the second part, the user was encouraged to speak aloud about what he or she was thinking and what he or she was trying to do. The evaluation expert only helped in cases where the evaluator did not manage to solve a task within approximately three minutes.

ANALYTICAL METHODS

To determine what occurred during the evaluation and also to find an ideal way to determine how the system was used in practice, we developed a methodology that captured as much of the user's physical interaction with the system as possible and also what the user thought and said.

One point of departure for measuring the user's interaction with the system was Aoidh and Bertolotto's (2007) methodology of analyzing the spatial location of the user's mouse interactions where the mouse's spatial location is considered representative of the user's interest in a specific feature. However, because RIV is a system with several menus, tools, and features, it was difficult to apply the same methodology to our evaluations. Another example was Tulis et al.'s (2002) comparison of lab and remote testing

6:18/09:19:57	Zoom out	Layers: cartes,orthophoto,parcelles
6:21/09:20:00	Zoom out	Layers: cartes,orthophoto,parcelles
6:25/09:20:04	Scalebar 1:100000	Layers: cartes,orthophoto,parcelles
6:36/09:20:15	Pan Tool	Layers: cartes,orthophoto,parcelles
12		
6:51/09:20:30	Pan Tool	Layers: cartes,orthophoto,parcelles
20		
7:14/09:20:53	Zoom in	Layers: cartes,orthophoto,parcelles
7:22/09:21:01	Zoom in	Layers: cartes,orthophoto,parcelles
7:32/09:21:11	Zoom in	Layers: cartes,orthophoto,parcelles
13		
7:47/09:21:26	Pan Tool	Layers: cartes,orthophoto,parcelles
7:55/09:21:34	Zoom in	Layers: cartes,orthophoto,parcelles
13		
8:010/09:21:49	Pan Tool	Layers: cartes,orthophoto,parcelles
8:19/09:21:58	Zoom in	Layers: cartes,orthophoto,parcelles
8:24/09:22:03	Zoom in	Layers: cartes,orthophoto,parcelles

Figure 1. Output of the log file visualization tool

We then utilized the log file visualization tool as an index for all evaluations that were recorded. It helped us to detect for each user:

- How much time it took to complete a given task;
- How many attempts where necessary to solve the task;
- What errors the user made during the evaluation (e.g., if he or she got lost in a menu that did not offer the functionality that was required to solve a task;
- What navigation tools (zoom-in, zoom-out, etc.) the user was using;
- If the user had different strategies to solve a given task (e.g., to navigate to a specific location).

where a specifically instrumented browser was used for capturing the user's interactions with the interaction of a Web site.

During the hands-on evaluation, different items were captured:

- The user through a video camera placed in front of the user
  (capturing sound and video);
- The user's screen (through a desktop streaming tool);
- The evaluation expert's notes; and
- The user's interaction with the system in a log file on the server.

Our system uses Apache as a Web server that logs each user activity into a log file. To analyze the log file, we created a tool that parses the log file and puts it into a database with every column showing a parameter. Thereafter, we created a tool that extracts the information from the database and visualizes the whole interaction session with one specific user and the system (see Figure 1). The different parameters were translated into a more human-readable format, filtering unnecessary elements and emphasizing important elements:

- The time when the interaction occurred (with an absolute time stamp and relative time stamp to analyze the log file afterwards in synchronization with the recorded screen and video);
- What tools the user was using (e.g., zoom-in, recentering);
- What layers the user was requesting;
- If there were gaps of more than ten seconds in between the different queries.

For streaming the user's desktop, we installed a VNC server on the evaluation computer that can be used for remotely controlling the computer. The signal of this VNC server is able to send the remote computer's screen to a client, but also accepts user input from the client computer's VNC client. For our evaluation, we used a tool installed on a second computer that streams this signal directly to a video file.

Before we used our log file visualization tool for measuring the user's performance, we verified its functionality with the screen captures that had been recorded during each session.

Moreover, we used the gap detection feature together with the videos (the user's screen and the user) to determine what was happening when the user was hesitating at a specific moment.

RESULTS

We structured our results according to the hypotheses and research questions mentioned previously. For each research question, we tried to find evidence in the data we collected during the evaluations.

USER BEHAVIOR/INTERFACE USE INFLUENCES

Is it possible to identify different user strategies to solve specific tasks?

To respond to the first research question, we analyzed the first task—to navigate the system's maps, to find the right spot, and then to digitize a parcel. We considered map navigation and digitization separately.

In RIV, the user had the choice of five different navigation tools (see Figure 2). Three navigation tools (zoom-in, zoom-out, and pan (to move the map)) where the selected tool is marked by

Figure 2. RIV's navigation tools: (1) zoom-in, zoom-out, and pan—the selected tool is highlighted; (2) the scale choice list with a choice of 16 scales; (3) scrollbars with direct access to the growing region and villages

a red frame, the scale choice list with 16 scales, and a menu that lets the user choose growing regions and villages. To analyze the user's strategies in the first task we verified:

- How many "navigation clicks" the user made and
- What type of clicks the user made.

At first, some differences in the frequency of use of the different tools was noticed (see Figure 3):

- One out of 20 users tried out all the tools during the first task.
- Six users used four different tools.
- Seven users used three different tools.
- Six users managed to navigate to the right spot (where the parcel had to be digitized) with only two different navigation tools.
- Eight users clearly used the pair zoom-in/zoom-out for changing the scale (however, some tried other methods as well).
- Five users used the scale choice list at least as often as the zoom-in/zoom-out pair.

Furthermore, users who used only a few navigation tools also needed only a few navigation clicks to complete the first task. On the other hand, users who used many different navigation tools also made many clicks.

In ten cases, we noticed that users tried to click on the zoom tools and expected the system to zoom in. Eight of those ten users found out later that it is necessary to click on the map to zoom in or zoom out and two users tried other navigation tools (such as the scale choice list and the recenter tool).

During the first evaluation series, all users had problems digitizing the first parcel (it took at least two and at most six attempts to digitize it correctly). The ten users of this first series had some problems in common:

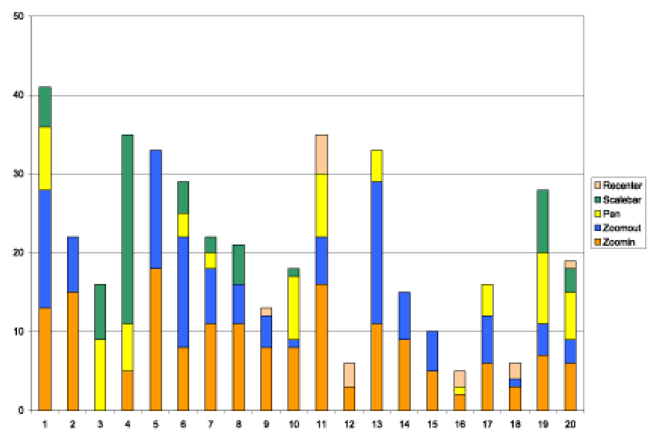


Figure 3. Number of clicks with the navigation tools to navigate to a parcel

- Five users tried to digitize the parcel as they would draw a line on a map by holding the mouse button clicked.
- Three users drew a complex polygon instead of following the outer border of the object.
- Two users tried to "paint" the parcel's interior with the digitization tool.

Because of this first result, the development team decided to help the user with that functionality and followed the suggestion of one of the first ten users to write a small note on the page (actually taken from the help pages) directing: (1) To zoom to the right place, (2) to select the digitizing tool, (3) to define the outer line of the parcel by clicking on the outer points, and (4) to click on the validation button once the parcel was finished. During the second evaluation series, six users managed to digitize the parcel at the first attempt; at most, it took three attempts to digitize it.

In the following tasks, we found no particular differences in the users' strategies compared to the first task. Here we were able to identify different user strategies for the interaction with the maps.

Which strategies result in better performance?

To analyze each user's performance during the first task, we measured:

- How much time each task took;
- Gaps of more than ten seconds between the user's clicks; and
- The number and type of errors the user made.

Navigation, the first subtask, is a continuing process where the user is:

- Considering the map and trying to put it into relation with the real world;
- Finding a strategy to change the state of the map (e.g., the map is not the right scale or is not showing the right place);
- Applying the strategy (zooming, recentering);
- Reconsidering the map, etc.

Measuring navigation time was difficult, however, because this evaluation was using verbal protocols—the user was encouraged to talk aloud while interacting with the system—thus gaps in the interaction with the system were quite frequent.

Because of these reasons, we decided to measure navigation time as follows:

- A navigation flow (many navigation clicks in a row) has less than ten seconds in between the navigation clicks; we counted the time from the beginning of the flow until the end of the flow;
- Single clicks with at least ten seconds in between.

For all users, we accumulated navigation flows and single clicks to one measure. We also counted the gap time separately.

Concerning different navigation strategies, we found out that:

- Users who only used one combination of a few navigation tools (e.g., zoom-in and zoom-out or scale choice list plus pan or recenter tool plus pan tool) needed much fewer clicks and overall time;
- Users who utilized four or more different tools needed more clicks and time to navigate to the right place.

This result is not surprising because of the system's response time on clicking and because each click produces a new map state that requires the user to realign himself or herself and to figure out the next steps.

As already mentioned, the task to digitize a parcel represented a considerable effort for most users. Users who had to start digitization over needed more time than other users.

In conclusion, users needing few tools and making few clicks needed less time than users who tried out different navigation tools and who made many clicks.

Is there an identifiable connection between the user's performance and the user's satisfaction?

Measuring the user's satisfaction is more difficult than measuring performance, for satisfaction is an individual opinion that can be divided into two statements:

- The user is satisfied with the functionality the system offers—the system enables the user to do the things he or she wants to do.
- The user is satisfied with how a feature or a tool works.

We did not explicitly ask whether the user was satisfied for the user's response might either correspond to either one or both of these statements. We therefore tried to approach the measurement of satisfaction with:

- The grade of difficulty the user had given to that task (ranging from 1 to 5) for difficulty must have an influence on the user's opinion of specific interface features and
- The comments the users had given during that task.

Comparing the grade of difficulty with the user's performance, we did not find any clear evidence for a direct connection between performance and the user's rating of the difficulty. In some cases, however, we did find a reason for the user's performance in the user's comments. Three users commented that RIV's map navigation is not functioning as in other systems the users had been using; in these cases, the users first had to figure out how RIV's map navigation works and therefore needed more time.

In the second questionnaire that was distributed after the hands-on evaluation, we used three statements (rating from 1 to 5) to indicate the user's confidence in the interaction with the system:

- "You always know where you are in the system."
- "The workflow of operations is intuitive."
- "It's always clear and comprehensible what is happening."

In the analysis of the users' responses to these statements, we assumed that users who understand where they are, see and experience what they expected, and know what they are doing are more satisfied than users who do not. Considering these statements, we could see a tendency that this assumption was true: Users who needed less time, less clicks, and few tries (one or two tries) to digitize a parcel responded to the three statements with the highest grades (three or four users).

Given the data and the parameters that we had for measuring user satisfaction, we did not find any evidence for a connection between the user's satisfaction and the user's performance; however, we did find slight evidence that users who stated that they were more confident in the interaction with the system performed better.

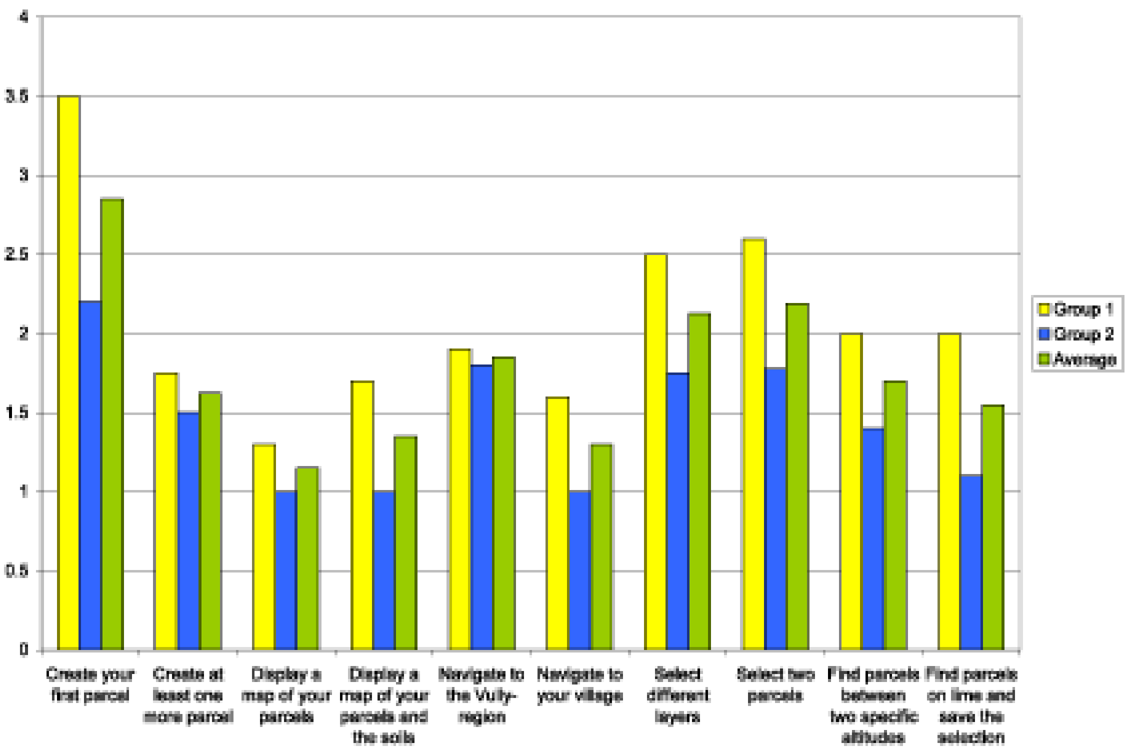


Figure 4. The difficulty of the different tasks rated by the users

SPECIFIC INTERFACE FEATURES INFLUENCE

Is there any evidence that some interface features cause a higher cognitive load?

To identify interface features that cause a higher cognitive load, we considered the following elements as indicators:

- The user's rating of the difficulty of a specific task;
- The user's comments on specific features; and
- The user's slower performance because of the features that the user had difficulties with.

At first, we analyzed the difficulty level the users had given to each task. The tasks with the highest difficulty level were: creating the first parcel, selecting two parcels (with the tool to select parcels), selecting different layers, and navigating to the Vully region.

Interface features that the users commented on were the navigation tools: to zoom in, to zoom out, and to move the map (pan). These tools were configured to work in the following manner:

- The user selects a tool and interacts with the map; if the user wants to change from zoom-in to zoom-out, the user has to select the tool by clicking on it—the user navigates the map by clicking on it with the selected tool;
- The user uses the zoom-in tool by drawing a rectangle on the region that the user wants to zoom in on;
- The user uses the pan tool by clicking on the map and moving the map—when the user stops clicking, the map is updated.

Some users commented that the zoom-in tool was difficult to understand. In fact, only five of 20 users used the zoom tool as intended (drawing a rectangle of the region to be zoomed in) and of the 15 users who did not draw a rectangle, ten accidentally drew a small rectangle (by clicking on the map a little bit longer) and zoomed in to the maximum level (1:100). The result of this problem was that:

- Users had to regain orientation on the map;
- Users had to use different methods to zoom out again;
- Users had either the impression that they did something wrong or that the system's zooming function is bad.

Moreover, ten users had a problem with the zoom-out tool (18 of 20 users tried to use it). All ten users clicked on the tool and expected the map to change (the idea was, as mentioned, to select the tool and then to interact with the map only). Thus, the users had to figure out a solution:

- Three users chose a different solution of zooming out (two the scale choice list, one the recenter tool).
- The other users discovered after a while that a click on the map was necessary in order to zoom out. However, two users each time they wanted to zoom out clicked first on the tool and then on the map.

The pan tool was used by only 12 users; two users commented that they did not understand how it was supposed to work and another user said that he would have preferred small arrows around the map in order to move it.

As a result, we can say that the manner in which the stan-

Figure 5. Layer management in ArcGIS, MapInfo, and Manifold

dard navigation tools are implemented in RIV can cause a high cognitive load.

In parallel to the problem with the zoom-in tool, the users had the same problems with the tool when selecting several parcels. The user could either use it to select one parcel by clicking on it or draw a rectangle to select several. As the task was to select several parcels, the user had to find a solution.

Only six out of 20 users knew or found out how to use this tool (by drawing a rectangle). The other 14 users were using either the menu "Parcel Groups," where the user simply could choose the parcels from a list, or the tool to query parcel by selecting different criteria (both offer the possibility to save the selection of parcels).

The task to digitize a parcel was an interesting issue for the discussion of the cognitive load of specific interface features for half of the users had a small text that gave a hint on how to use the parcel digitizing tool; therefore, we analyzed these tasks for both groups of users (first ten users, last ten users) separately.

The first group (who did not have the hint) needed much more time to navigate and to digitize than did the second group. Only the digitizing part took almost twice as much time for the first group than for the second group. Surprisingly, when both groups were asked to digitize the second parcel, the first group digitized and navigated quicker than did the second group.

In conclusion, the feature to draw a rectangle on a map (either with the zoom-in tool or the select-object tool) does cause a higher cognitive load. In the case of a problem, the user had to find a different solution to solve a task that involved map navigation or the selection of objects. Moreover, the parcel digitizing tool did cause a high cognitive load for the user for the user had to discover its functionality during the task.

Were there differences in performance that were associated with features of RIV that are common with conventional GIS versus less conventional/interactive tools?

To define features that are typical for conventional GIS, we first analyze which interface features exist in both RIV and conventional GIS. As examples for conventional GIS, we refer to the GIS most commonly used at our laboratories: ArcGIS, MapInfo, and Manifold:

- A map that is built up by different layers;
- Navigation tools to navigate the map;
- Tools to create and to manipulate georeferenced data; and
- Tools to query this data.

Map and Layer Model. The map is the central part in all conventional GIS. It is composed of layers that are organized in a hierarchical manner. The user can display different layers in different modes and reorganize them. In all three conventional GIS we considered, layer management is done very differently (ArcGIS—a layer tree lets the user organize and display the different layers; Mapinfo—a toolbox, accessible through the main menu, displays the layer management; and Manifold—a map window where the layers' vertical position symbolizes the layers' hierarchical position and layers can be deactivated by clicking on the layer's name (see Figure 5)).

In RIV, the ArcGIS layer management system was adopted (see Figures 5 and 6) with the extension that the user could chose between different layer categories.

Moreover, the display of some layers changes automatically from scale to scale (e.g., at a scale of 1:1,000 the user sees aerial images; at a scale of 1:25,000 the user sees a map that is optimized for that scale; and at a scale of 1:100,000 the user sees a different map).

Despite the fact that users had problems finding some layers because of the selection of the different layer categories, all users performed well while using this feature.

Navigation Tools. As mentioned previously, the following problems occurred for many users while they were using the navigation tools:

- Drawing a rectangle to zoom in and to select features and
- "Selecting" a tool that manipulates the map.

Both interaction methods are standard in all conventional GIS we considered.

Tools to Create and Manipulate Georeferenced Data. In conventional GIS, a huge variety of tools exists to create and manipulate georeferenced data, such as different digitizing modules, tools to create buffers, etc.

The digitizing tool that was implemented in RIV was conceived with a conventional GIS in mind: In all conventional systems, digitizing is always performed by clicking on the vertices of the object—a click on the next vertex adds a new segment of the polygon until the polygon is closed. In RIV, the user has to

Figure 6. Layer management in RIV

In the task where the users had to display different layers, all users understood that clicking on a layer's checkbox did display the layer. However, because of the number of layers that the user can choose, the development team chose to group layers (28) into categories and subcategories.

All users needed some time to figure out that a certain layer that they had to select was hidden behind a scrollbar (where the user could chose the category of layers). Another feature that some users did not see was a small "+" icon in the layer's tree that visualized the legend of a specific layer.

click on the first vertex (after adding more vertices) to close the polygon.

As mentioned previously, many users had problems digitizing a parcel in RIV. Users made many errors and needed much time and effort to digitize a first parcel. We believe that this problem also is related to a conflict of interaction methods (in a Web context versus a GIS context) and that there was no possibility the user could know how the tool worked. Moreover, we discovered that some users hatched their parcels with the digitizing tool or just drew each line by using the mouse button clicked. We believe that this fact is related to the following issues:

Digitizing in a GIS context means creating vector data that is represented by points and lines. However, a parcel representation in terms of lines and points is not necessarily the representation a winegrower has in mind when thinking of his or her parcel.

The closest we can come to a digitizing task on a computer is probably to draw a parcel on paper. Drawing on paper is done by drawing a pen along a line. Many users probably considered the digitizing task as related to drawing on paper.

Unfortunately, we were not able to implement a better way to help the user digitize, so we decided to explain it by using a short text.

Tools to Query Georeferenced Data. In conventional GIS, a variety of different possibilities exists to query georeferenced data. The most common way to query georeferenced data is through the query language SQL that permits running all kinds of different scripts. Some GIS (such as MapInfo or ArcGIS) have toolboxes that let the user chose possible operators and data from scrollbars to form an SQL or SQL-like query string.

In RIV, we decided to implement a task with only one spatial operator (within) and one nonspatial operator (is equal to) for a menu called "Search by Criteria." The user could chose between a variety of different spatial and nonspatial attributes that should be true. Moreover, we tried to form a natural sentence "Search all parcels for the following attributes" and then for each attribute a scrollbar.

The task to query all their parcels and to search for parcels with specific attributes or that were within specific regions worked very well for all users; we believe that the fact that we tried to form a natural sentence on the page helped the users to use this feature.

USER EXPERIENCE AND TRAINING INFLUENCES

Are there identifiable differences between users?

In our first questionnaire, we used four statements that we consider indicators of the user's experience with computers and cartography:

- The user has taken courses in informatics (e.g., word processing, etc.).
- The user has a high-speed connection to the Internet.
- The user is a frequent user of paper maps.
- The user previously has used cartographic systems.

We assume that users who have taken courses in informatics know how to use a computer's input devices (keyboard and mouse) and know how a computer reacts on this interaction. Furthermore, we presume that users who have high-speed Internet connections are more likely to spend more time on the Internet (and probably also are using applications that require a high bandwidth) than users who do not. We think that users who utilize paper maps frequently have an idea how the reality is depicted on such maps, what a map scale is, and how cartographic symbols are used.

Mapping System	Click On Zoom Tool => Zoom	Scale-bar Zoom	Rectangular Zoom	Arrows	# Users
GéoPlaNet	No	Yes	Yes	Yes	9
Swissgeo	No	Yes	No	Yes	5
Mappy	Yes	Yes	No	Yes	5
Google Earth	Yes	Yes	No	Yes	4
Michelin	Yes	No	Option	No	3
CFF	No zoom tools	Yes	No	Yes	3
Google Maps	Yes	Yes	No	yes	3
Swissinfo	No	Bug	Yes	no	3
Map Search	Yes	Yes	No	no	2
Map24	Yes	Yes	Yes	no	1
Has feature	6	8(9)	3(4)	6	

Figure 7. Different systems that users had employed before the evaluation

Finally, we argue that the interaction with cartographic systems also has a significant impact on the use of RIV.

We found that out of 20 users:

- Seven users had answered yes to all four statements;
- Eight users to three statements;
- Four users to two statements; and
- One user to only one statement.

Is there any evidence to suggest that a user's geospatial technology expertise has an influence on the user's performance and on the way the user interacted with the interface?

We think that geospatial technology expertise is a combination of mainly two different elements: computer expertise and experience with geographical information.

Previously, we have detected measures that we judge relevant for both of these elements. Our assumption is that users who have answered yes to all four statements are likely to perform better than users who have answered no to at least two statements.

We used the following measures from the first task (which was the first time the users had to use RIV's mapping system) as indicators for performance:

- The time to navigate to the parcel;
- The number of navigation clicks to navigate to the parcel;
- The time to digitize a parcel (considering the fact that the second group had an explanation of how to digitize); and
- The number of tries to digitize a parcel (considering the fact that the second group had an explanation of how to digitize).

We found the following evidence in our data:

- The one user who had answered yes to only one statement (that he has a high-speed connection) performed worst in navigation time and navigation clicks.
- The four users who performed best in navigation had answered yes to three (two users) or four (two users) statements. All these users had used cartographic systems before.
- The user who performed worst in digitizing (from the first group) had answered yes to only two statements.

All but two users had used online mapping systems before this evaluation; many users had used two or three different systems. GéoPlaNet, the canton's official mapping system, was the most known of those ten alternatives we had listed. Moreover, two users indicated that they had used some installable mapping systems before (Twixtel and ArcPad) and one user mentioned that he had used a GPS before.

All these systems are either way-finding systems (Mappy, Michelin, Map24); pure mapping systems (GéoPlaNet, Swissinfo); hybrid systems (Google Maps); systems with specific information such as railway stations (CFF) and infrastructure (Map Search) and a virtual globe (Google Earth). All systems have different navigation methods with different tools that are used to zoom in and zoom out or to pan.

Because we found that navigation tools in RIV are especially problematic for many users, we analyzed how map navigation works in these systems (see Figure 7). We analyzed if:

- The system reacts directly on clicks on the navigation tools.
- There is a scale choice list that permits zooming.
- The zoom-in tool supports a "rectangular zoom."
- There are arrows around the map that the user can click on to move the map.

We found that four of the ten systems do support a "rectangular zoom"; however, all but one of these systems do support zooming through the scale choice list (one system's scale choice list partly worked). We can further see that six systems are offering map navigation that does not involve clicking directly on the map—these systems support arrows for moving the map in each direction and a scale choice list.

We compared the manner each user interacted with the system and analyzed:

- If users who used the rectangular zoom in RIV were in

contact with systems who also support this feature;
- If users who had problems with that feature were in contact with systems that support other navigation methods;
- If users who tried to click on the zoom tools without afterwards clicking on the map were using systems that support that feature; and
- If users who frequently used the scale choice list were in contact with systems that support that feature.

We found the following evidence in our data:
- Four out of five users who used the rectangular zoom stated that they had been in contact with systems supporting this feature;
- Nine out of ten users who had problems with the rectangular zoom had used systems that support other possibilities to zoom than a rectangular zoom (the one user who had the problem had not used any cartographic systems before);
- Eight out of ten users who initially expected the system to react after clicking on the zoom tool had used systems before that support a direct click;
- Seven out of nine users who used the scale choice list had been in contact with systems supporting this feature (the two other users had not been in contact with any cartographic system before).

We found out that there seems to be a connection between the user's background, in terms of expertise in computers in general and experience with geographical information and the way they interacted with RIV. Moreover, there is a likely connection between the geospatial systems the users had used before and the methods the users performed in this evaluation.

DISCUSSION

In our evaluation, we found some evidence that supports our hypotheses. We focused mainly on the first tasks in which the users interacted directly with the map.

We could see that the 20 users in our evaluation did have different methods to interact with the system and that some methods resulted in different performances.

The problems that occurred were not only linked to one cause but to many specific reasons. We found evidence that users are highly influenced by their previous experience with geographical information and computer systems. Users who did not find the interaction method they were used to were forced to find different ways to interact with the system and thus performed worse than users who found the interaction methods they were used to. This is especially the case using the navigation tools.

The problems that occurred with these navigation tools were mainly because of two reasons:
- RIV is a system that is used in a Web context. In a Web context, a click on almost all standard elements (buttons, checkboxes, radio buttons, links, menus, etc.) has a direct effect on the display's state (e.g., one click on a hyperlink makes the system navigate to the next page). Using the conventional GIS interaction method of selecting a tool and then interacting with the map is a rupture in the Web context. Also drawing a rectangle is a rupture of this context for all standard objects that are clickable in a Web context usually are points—the user is used to moving the mouse to one feature (button, link, etc.) and then clicking on that feature. Drawing a rectangle implies holding the mouse button for a longer time and moving the mouse while clicking.
- In RIV, there was no possibility the user could have known that the navigation tool could have been used in this manner; neither the icon of the zoom-in tool suggested it, nor did a text indicate it.

Regarding our hypothesis that there is a connection between performance and satisfaction, we only found very poor evidence with the measures we had taken and methods we had chosen.

As previously mentioned, satisfaction is an individual opinion; possibly, not only the user's performance but also the user's experience can have a positive or negative influence on the user's performance and satisfaction. For instance:
- Users who were less used to computer systems and geographical information and needed more time may have been surprised to see the possibilities of such systems and thus were satisfied with the systems.
- Users who were more used to these elements and performed better could have expected more or different functionality and thus were less satisfied.
-

An interesting point was the task to create the first and the second parcel. Although the second group was provided a small explanation in the system indicating how to digitize a parcel, the performance for the second task was slightly lower than the first group's performance when doing the same task. We believe that this fact could have a relation to Bloom's Taxonomy of Educational Objectives, which says that there are different levels in the cognitive domain: knowledge (low level), comprehension, application, analysis, synthesis, and evaluation (high level):
- The first group was forced to understand and to learn how to digitize a parcel by trying out. The cognitive level thus was higher; but once the users understood and tested the task several times, it was easier to do the same task again.
- The second group "simply" followed instructions while they were doing the task and succeeded after one or two tries. However, when creating the second parcel, the second group only had a lower-level cognitive picture of the task and less practice in performing this task and thus performed slightly lower.

CONCLUSION AND FUTURE RESEARCH

We found that the interaction of a real-world user with a specific WebGIS complex compared to the interaction with a "normal" homepage with hyperlinks and content. It requires other cognitive strategies such as navigation in a virtual space using specialized tools and the interpretation of specific maps.

Despite the fact that we only had 20 persons who evaluated the interface, we could see that users performed very differently. These differences were caused by a set of factors that can be considered from the point of view of:

- The interaction between the user and the system (user strategies, performance, and satisfaction);
- The system (interface features);
- The user (experience and training).

We established hypotheses depending on these categories and we tried to accept or reject these hypotheses by answering the research questions.

Although we could not entirely accept or reject these hypotheses because of the limited number of users, we found evidence about the links between these categories. Moreover, we were able to find some tendencies and trends and to detect interface features that did cause problems for many users.

The methodology we have developed to answer the research questions is capable of delivering evidence about the level of connection between these factors. We believe that this methodology can be used for the evaluation of most WebGIS, even if they are based on different technologies.

In the future, we will refine our evaluation methodology with further tests involving different systems, users, and evaluation tasks.

Within our research program, the long-term goal is to analyze the importance of these factors; for instance:

- Which user experience and training factors are more important?
- Which factors do have an influence on the user's satisfaction?

Based on this research, we want to establish guidelines for developers and system designers on how to build specific systems for particular groups of users (e.g., it is better to develop a system with the zoom-box feature for users with much experience in the field of geoinformation for they would expect the system to have it, while it is more reasonable to develop a system without this feature for less experienced users).

Acknowledgments

The authors would like to thank Dr. Nick Hedley from the Simon Fraser University (SFU) in Vancouver, Canada, for his precious input.

About the Authors

Dr. Jens Ingensand is a researcher at EPFL's GIS Research Lab, with research interests in improving the usability of Web-based cartographic and decision-support tools. He is also a lecturer and instructor for Master's courses on geographic information technology. He holds a Master's degree in applied informatics with a major in human-computer interaction from Chalmers, University of Technology, Gothenburg, Sweden, and a Master's in geography from the University of Gothenburg, Sweden.

GIS Research Laboratory
Swiss Federal Institute of Technology, Lausanne (EPFL)
Institute of Urban and Regional Planning and Design – Geomatics
Lausanne, Switzerland
jens.ingensand@epfl.ch

Dr. Francois Golay has been a professor of Geographic Information Systems and Science with EPFL (Swiss Federal Institute of Technology - Lausanne) and director of the GIS Research Laboratory since 1994. His research interests are focused on improving the value and empowering the users of geographic information, for environmental and land-management purposes as well as for related research activities. His research encompasses work on spatial decision support systems, human-computer interaction, geographic data infrastructures and information sharing, and promoting and assessing the value and usability of GIS.

GIS Research Laboratory
Swiss Federal Institute of Technology, Lausanne (EPFL)
Institute of Urban and Regional Planning and Design – Geomatics
Lausanne, Switzerland
francois.golay@epfl.ch

References

Aoidh, E. M., and M. Bertolotto. 2007. Improving spatial data usability by capturing user interactions. Proceedings of the AGILE 2007 Conference, Aalborg, Denmark.

Beer, T., T. Anodenko, and A. Sears. 1997. A pair of techniques for effective interface evaluation: Cognitive walkthroughs and think-aloud evaluations. Proceedings of the Human Factors and Ergonomics Society, 41st Annual Meeting, 380-84.

Dumas, J. S., and J. C. Redish. 1999. A practical guide to usability testing. 2nd Ed. Kirkland, WA: Intellect Publishing.

Ericsson, K., and H. Simon. 1993. Protocol analysis: Verbal reports as data. 2nd Ed. Boston: MIT Press.

Haklay, M., and C. Tobón. 2003. Usability evaluation and PPGIS: Towards a user-centered design approach. International Journal of Geographical Information Science 17(6): 577-92.

Haklay, M., and A. Zafiri. 2008. Usability engineering for GIS: Learning from a screenshot. The Cartographic Journal 45(2): 87-97.

Ingensand, J. 2006. Developing Web-GIS applications according to HCI guidelines: The Viti-Vaud project. In Geographic hypermedia: Concepts and systems. Springer, 409-20.

Ingensand, J., R. Caloz, and K. Pythoud. 2005. Creating an interactive network for wine-cultivation. Presented at the Inaugural Nordic Geographers Meeting, Lund, Sweden, May 11-13, 2005.

Jankowski, P., and T. Nyerges. 2003. Towards a framework for research on geographic information supported participatory decision making. URISA 15: 9-17.

Jacobsen, N. E., M. Hertzum, and B. E. John. 1998. The evaluator effect in usability tests. ACM CHI 1998 Conference Summary, Los Angeles, California, April 18 to 23, 255-56. New York: ACM Press.

Preece, J., Y. Rogers, and H. Sharp. 2002 interaction design: Beyond human-computer interaction. New York: John Wiley and Sons.

Pythoud, K., and R. Caloz. 2001. Etude des terroirs viticoles vaudois—rapport d'avancement. Lausanne, Switzerland : Laboratoire de SIG–EPFL.

Traynor, C., and M. Williams. 1995. Why are geographic information systems hard to use? CHI 1995 Proceedings. ACM Press.

Tullis, T. S., S. Fleischman, M. McNulty, C. Cianchette, and M. Bergel. 2002. An empirical comparison of lab and remote usability testing of Web sites. Usability Professionals Association Conference, Orlando, Florida.