

# Towards an Evaluation Methodology for the Development of Research-Oriented Virtual Communities

Mark Chignell, Janet Ho, m.c. schraefel

Knowledge Media Design Institute,  
University of Toronto  
10 King's College Road, SF4306  
Toronto, Canada  
M5S 3G4

## Abstract

A virtual community is a group of people whose shared needs and/or interests are largely communicated within, and mediated by Web-enabled interactions, e.g., within a Website. While some virtual communities (e.g., Motley Fool) have flourished, many sites designed for online collaboration have failed to succeed as virtual communities, places for people to hang out and collaborate online. In order to design effective virtual communities, we need methods to evaluate the usability and effectiveness of collaborative environments, and the degree to which they function as virtual communities (VCs). Our concern in this paper is to explore, through the discussion of a pilot study of a research-purposed VC, how to evaluate a VC infrastructure.

## Keywords

Virtual communities, evaluation methodology, distributed research communities

## 1 Introduction

A virtual community is at once both a set of technologies, processes, and associated content that provide people with the means to meet, collaborate and interact as an online community, and a collection of people who regard themselves as a virtual community because of those meetings, collaborations, and interactions. From the technological perspective, Virtual communities are a complex collection of types of media components which are pulled together to serve an imagined sense of collocated community. From the human perspective, a virtual community is a group of people whose

shared needs and/or interests are largely communicated within, and mediated by Web-enabled interactions, e.g., within a Website. Motley Fool ([www.fool.com](http://www.fool.com)), which formed around an interest in stock investing, is an example of a virtual community (Hagel and Armstrong, 1997).

We leave aside for the moment a definition of what constitutes a virtual community and what attributes must be supported to constitute a "VC". We assume that such definitions can be worked out, and we have postulated several in previous work [10].

Regardless of the specific definition of a given VC, certain elements are consistent across such environments: the application includes support for collecting and distributing multimedia content to and from users who may or may not be collocated, and who may or may not be communicating synchronously in the environment. For such a system to be successful, however, it needs not only to attract users, but to maintain their interest and willingness to continue to contribute to the VC, otherwise the VC will atrophy. Therefore, VC builders need an appropriate methodology for VC development in order to build "sticky" infrastructures. Our concern in this paper is to explore, through the discussion of a pilot study of a research-purposed VC, how to evaluate a VC infrastructure once a definition for that structure is determined, as part of a process of ongoing iterative design, evaluation, and improvement.

**Layers** In order to define a methodology for developing VC's, we begin by considering how best to evaluate existing Virtual Communities. Like any multimedia application, a VC should be subjected to some form of usability testing [1]. However, traditional usability evaluation methods originate from single user applications

and do not necessarily address issues in collaborative tasks and distributed, integrated systems like the Web and Web based VC's. Since web site development involves the close interaction of content, navigation, and appearance of the website, it is the entire site, not simply a component of it, that needs to be evaluated. The complexity of evaluating such a system, moreover, is multiplied when the web site forms the nexus for a Virtual Community, a site designed not only for viewing but for cooperative, collaborative support of its users/members. The evaluation of a virtual community therefore will likely need to consist of an agglomeration of usability inspection methods for hypermedia applications and new methods for evaluating distributed virtual environments. The approach will also need to incorporate ethnographic analysis of VC interaction that includes the concerns of usability engineering (Steed & Tromp [12]).

This paper is a first approach to such a collective approach to VC evaluation. We describe some general usability evaluation methods, and the rationale for their selection for VC evaluation. We then present an evaluation questionnaire that we have used in our pilot study to evaluate a virtual community specifically built for research collaborators.

## 2 Methods of evaluation

In our pilot study, we have employed heuristic evaluation for problem solicitation, cognitive walkthroughs for naïve user observation, and questionnaires for self reporting of user experience. We treat each of these in turn.

Nielsen's heuristic evaluation has been demonstrated to be a successful method for "discount usability testing" of interface design in particular. A small group (three to four) of expert evaluators individually assess an interface according to a set of heuristics [0]. However, Nielsen and Molich [0] also note that heuristic evaluation focuses on finding problems rather than proactive solution development. The analysis of the interface suggests how to fix errors on the proposed design rather than where to recreate a new design.

The complementary approach to the expert heuristic evaluation, then, is the cognitive walkthrough. Cognitive walkthrough is a method of analysis which focuses on exploratory learning [7]. In our context, this method seems to have significant potential since our VC has many features, but few of them have been used

by participants. Indeed, of all the tools available for synchronous communication, building of separate virtual lab spaces and document generation features, the single feature being used with any frequency is document uploading to the VC web site. Cognitive walkthroughs can help us understand how naïve users may move through the site and understand how to access (or not) the available tools and environment. These walkthroughs are particularly suited to such first time users since they measure how users explore and learn how to use the interface of the VC. The evaluation is specifically concerned with success rate and recovery from error, but is not concerned with the task performance [7].

Related to this practice is Garzotto and Matera's hypertext usability evaluation methodology named SUE (Systematic Usability Evaluation)[2]. This model uses a number of usability attributes by which a hypermedia application can be assessed. The dimensions are structure, navigation, behaviour, user control, and presentation [2]. We can use this model to help us focus our observation protocol of the walk-through.

As suggested above, we are also concerned not only with how users interact with the VC but whether or not their experience represents a high enough value to have them continue to use a VC despite any potential limitations. We also wish to find a way to reduce the cognitive load of interaction with a VC. To that end we are seeking a metric to analyze the mental/cognitive workload to evaluate the demand a task in the VC imposes on a user's limited resources [13]. Hart and Staveland [5] proposed six factors (mental demand, physical demand, temporal demand, performance, effort, and frustration level) that associated with variations in subjective workload, and we can use these factors in both usability experiments. Before establishing an experiment, however, we can also use these factors to inform the development of a self-reporting instrument, like a questionnaire. If our concern is with enhancing stickiness, the user's own perception of their experience in an existing environment should help us design more appropriate attributes to enhance the user experience of the environment.

### 2.1 Revised evaluation methodology

Summarizing the evaluation methods described above, our first step of heuristic evaluation provides a quick feedback of the VC's performance and participants' preferences. Other forms of the evaluation methods will be

carry forward in the form of a questionnaire and personal interviews. The questionnaire is an instrument used to study a research problem with a collection of questions and statements given to a sample of individuals [6], and is used for gathering enough data to perform statistical analysis [7]. In addition, personal interviews with actual participants of the VC are required to get a hold of more personal and richer sense of their experiences in using the VC. As Newman and Lemming [7] state, interviews are a quick and common way to perform an analysis, where results are available immediately.

### 3 Questionnaire

The questionnaire used to evaluate the VC is based on Nielsen's ten heuristic for the web [0], Garzotto and Matera's SUE methodology [2], and the modified cognitive workload analysis [5].

The questionnaire is divided into six categories (see the Appendix to this paper for a list of the questions used):

- (1) structure of the content's organization
- (2) navigation of links used to explore the website's structure
- (3) behaviour of functions and links of the website
- (4) user control of available interaction components
- (5) presentation of features shown to users, e.g. layout and visual appeal, and
- (6) functionality of the component and amount of effort required

Each of the items in these categories consisted of a statement with which the respondent could either disagree or agree (a Likert type scale was used with five categories: strongly disagree, disagree, neutral, agree, strongly agree).

In future versions of the questionnaire we also plan to add questions concerning:

Degree of Personalization  
Sense of Community  
Perceived Stickiness  
Loyalty and willingness to continue using the site

These types of questions/items should be of particular interest in tracking how and why online collaborative environments do, or don't evolve into effective virtual communities.

### 3.1 Discussion of results

We performed an initial study on the evaluation of the VC in question. We handed out a questionnaire to a number of participants who are currently using the VC for a research project. The questionnaire asks the participants to rate the VC according to the six criteria: structure, navigation, behaviour, user control, presentation, and functionality. All participants have used the Internet for over 3 years, and are familiar with using Internet chat software, newsgroups or other discussion sites.

Preliminary results from the pilot study indicated that overall the VC was not satisfactory to the participants. The participants found the structure and organization of the site difficult to understand and represented inconsistently. The participants complained that there were too many links to go through before arriving at the relevant link. A moderate high amount of cognitive demand was required to complete the tasks, where most work is used on searching for the relevant information. The VC did not have much user control, and participants were unable to cancel their wrong actions. The presentation of the website was not found to be particularly appealing and help was needed to guide the participants. On the other hand, the VC's general behaviour was as expected and the functionality was familiar to participants. From these preliminary results, however, the VC did not appear to attract general use.

This pilot study now serves as a ground work for two contexts: the first is refining to have established its usefulness as an instrument for a larger study. The preliminary analysis suggests that we can proceed with the survey to collect , more data to evaluate how the site can be improved to achieve its goal: to act as a distributed virtual community of many different lab groups. We are currently engaged in gathering the user logs for the site to see how these complement the results of the self-reporting in our pilot survey.

The second context for discovery is that by creating an effective instrument for evaluating existing VC's we can move towards our actual goal: to develop a methodology based on these findings that will inform the early stages infrastructure deployment of a heterogeneous VC environment.

## 4 Conclusion

From the above evaluation, one might think that the VC we evaluated was a rough prototype. It isn't. It was developed by an experienced team from industry over the course of a year, and recently transferred to our research lab for further development. The VC itself, while having a set of compelling features for a variety of synchronous and asynchronous, secure and public exchanges, is not currently compelling as a virtual community. Since researchers rely on a heterogeneous set of tools and technology practices, research-oriented VC's must effectively integrate many complementary technologies across discrete environments. Because of the diversity of practices and technologies, developing such an integrated architecture is a non-trivial problem. The pilot study results underline our view that an appropriate evaluation methodology for VC deployment is required, not only as a cost-saving measure in the early stages of VC building (for assisting iterative design), but also as a potentially new way to gain insight into how people may be able to come to work and do research (rather than simply report on it) in the virtual domain.

## 5 References

1. Chignell, M & Waterworth, J (1997) 'Multimedia', in G. Salvendy (ed's), *Handbook of human factors and ergonomics*, 2nd ed. , John Wiley & Sons, Inc., USA, pp. 1808-1861.
2. Garzotto, F., and Matera, M. (1997) 'A systematic method for hypermedia usability inspection' *The New Review of Hypermedia and Multimedia*, p.39-65.
3. Garzotto, F., Paolini, P. & Schwabe, D. (1993) 'HDM- A Model-based Approach to Hypertext Application Design' *ACM Transactions on Information Systems*, Vol. 11, No. 1, January, pp. 1-26.
4. Hagel, John III and Armstrong, Arthur G., 1997. "Net Gain". Cambridge, Mass.: Harvard Business School Press.
5. Hart, S. & Staveland, L (1988) 'Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research', in P.A. Hancock & N. Neshkati (ed's), *Human Mental Workload*. Elsevier, pp. 139-183
6. Mantel and Marshall (1998) 'Introduction to Questionnaire Design' In *MIE 1403 Course Lecture Notes*, Toronto.
7. Newman, W.M., and Lamming, M.G. (1995) *Interactive System Design*. Great Britain: Addison Wesley.
8. Nielsen, J. (1992) 'Finding usability problems through heuristic evaluation' in *Proc. CHI'92*

*Human Factors in Computing Systems*, Monterey CA, May 3-7, pp. 373-80, New York: ACM/SIGCHI.

9. Nielsen, J., and Molich, R. (1990) 'Heuristic evaluation of user interfaces'. *Proc. ACM CHI'90*. Seattle, WA, April, pp. 249-256
10. schraefel, m.c., Janet Ho, Mark Chignell, Michael Milton "Building Virtual Communities for Research Collaboration" *AIWoRC'00 An International Working Conference and Industrial Expo on new advances and emerging trends in Next Generation Enterprises*, Buffalo, NY, April 27 - 29, 2000.
11. Scholtz, J., & Downey, L 'Methods for Identifying Usability Problems with Web Sites' September 30, 1998  
<<http://www.itl.nist.gov/iaui/vvrg/jpaper/final.htm>> (December 2, 1999).
12. Steed, A.& Tromp, J., (1998) 'Experiences with the Evaluation of CVE Applications', *Proceedings of Collaborative Virtual Environments*, 17-19th June 1998, Manchester, UK.
13. Wickens, C.D. (1992) *Engineering Psychology and Human Performance*, 2nd ed., New York: HarperCollins Publishers.

## 6 Appendix: Website Evaluation Questionnaire

### I. Structure

1. It is easy for me to find desired items on the website:
2. The organization of materials on the website is easy to understand:
3. There are too many levels of nesting on the website, requiring me to drill down to items I want:
4. The maps or overviews help me to locate where I am within the website:
5. I am able to visualize the whole structure and layering of the website:
6. The organization of materials on the website is represented consistently:
7. This structure of the website and how things are organized makes sense to me:
8. The information on most pages appears in a logical and natural order:

### II. Navigation:

1. The links to other pages are clearly marked:
2. I can log out of (leave) the website quickly:

3. There are enough links to jump to relevant area in the website without going through too many links:
4. It is possible to return to bookmarked pages:
5. I got disoriented at times and wasn't sure where I was in the website:
4. Rate the amount of work spent on looking and searching for the relevant pages:
5. I accomplished my goals in using the website:
6. I was able to perform tasks and accomplish my goals on the website efficiently:

### **III. Behaviour:**

1. The text used on the website's pages is generally understandable:
2. I am familiar with most of the functions on this website because I have used them on other websites:
3. The links go to where I expect them to go:

### **IV. User Control:**

1. The website supports undo and redo functionality:
2. I can cancel my previous action and proceed to another page:

### **V. Presentation:**

1. The mapping between buttons and functions is consistent throughout the website:
2. The titles and headers are worded consistently throughout the website:
3. There were irrelevant and extraneous materials on the website that I found distracting:
4. The screen layout (e.g. button appearance, font size, font type, colour) of the website is appealing:
5. More help functionality and information is needed on the website:
6. The help function provided is useful:
7. I can locate specific help items:
8. There are unknown error without proper messaging, e.g. Javascript error:
9. In the event of an error, there are helpful solutions to solve the issue:

### **VI. Functionality:**

1. Please rate the amount of effort you used to perform tasks in the website:
2. Please rate the difficulty in performing tasks in the website:
3. Please rate your level of annoyance or discouragement caused when using the website: