Usability Evaluation of a Simulated Vascular Reconstruction System

H.S.M. Cramer[1], V. Evers[1], E.V. Zudilova[2] and P.M.A. Sloot[2]

[1]Section Social Science Informatics, University of Amsterdam, The Netherlands
[2]Section Computational Science, University of Amsterdam, The Netherlands

1. Introduction

In order to ensure successful adoption of Virtual Reality applications, usability and the context in which a system will be used has to be considered in system design and development. Both quantitative and qualitative methods are available for system development (see e.g. Preece et al., 2002). This paper focuses on a combination of qualitative methods that have been applied in the development of a prototype of a VR system for medical diagnosis and planning for vascular disorders, the Virtual Radiology Explorer (VRE). In this case, heuristic evaluation was applied to assess usability of the VRE prototype and to improve usability of future versions of the VRE. Besides the application to the VRE, many issues identified in this study pertain to development of virtual reality applications in general.

Usefulness of any system largely depends on how user needs for both utility and usability are supported (Grudin, 1992; as quoted by Hilbert & Redmiles, 2000). Both should be taken into account in design and evaluation of a system. For our purposes we comply with the definition of usability as put forward in the ISO 9241-11 (1998) standard: the extent to which a product can be used by specific users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use. The heuristic evaluation did not only result in identification of usability problems of the prototype but also in the realization that a closer look at context of use and desired utility of the system would be necessary in further development.

A contextual analysis was conducted into the process of medical diagnosis and treatment planning of vascular disorders to gain insight in functional and non-functional requirements for systems supporting these processes. This study is briefly described and a more elaborate discussion of the study can be found in Cramer et al., (2004b). This paper focuses mainly on the issues and questions that arose from the evaluation of the VRE prototype. We will describe the VRE evaluation and its results, and will highlight some generic issues surrounding VR development.

2. The Virtual Radiology Explorer (VRE)

The Virtual Radiology Explorer (VRE) is a multi-modal VR system for 3D visualization of medical scans, simulation and visualization of possible treatments for vascular disorders. The goal of the VRE is to assist in the pre-surgical planning and VR training for abdominal vascular bypasses. The potential users that have been considered are surgeons, radiologists, medical lecturers and students. 3D reconstructions of scan data can be explored interactively and measurements can be conducted if necessary. Virtual objects, including the vascular structures, are visualized in a virtual environment. This environment can be freely navigated through by the user. Its virtual objects can be moved and manipulated. Both the artery geometry and the simulated blood flow can be visualized in 3D and in time (4D, see Belleman and Sloot, 2001). The working prototype of the VRE system provides visualization of the results of blood flow simulation before and after a simulated bypass procedure. The results of the blood flow simulation (velocity, pressure and shear stress) are visualized by means of glyphs and streamlines (Sloot et al., 2003). The VRE is an immersive application, in principle it can run in a CAVE. The evaluated VRE version (fig. 1) currently runs on the semi-immersive Distributed Real-time Interactive Virtual Environment (DRIVE) system. With this system, stereoscopic images are offered on one large screen to the user, who wears shutter glasses and whose position is tracked. Navigation and selection are achieved by manipulating a 'wand' (3D mouse) and by speech recognition. Extensive information on the VRE and the DRIVE can be found in Sloot et al., (2003), Belleman (2003), Belleman and Sloot (2001), Zudilova and Sloot (2003) and Zudilova et al., (2002).

3 Problem statement

Due to the nature of the VRE's development as a technological test-bed, as is the case with many VR systems, the development had focused more on technological feats than on the practical implementation for actual users. A number of usability problems of the VRE prototype were apparent to the development team, but a more complete overview of potential usability problems was needed in order to inform development of the VRE. Subsequently, the need has been acknowledged for studying usability criteria for such a system

for medical scan visualization and vascular surgical planning. Thus the main question: what are the potential usability problems of the VRE prototype?

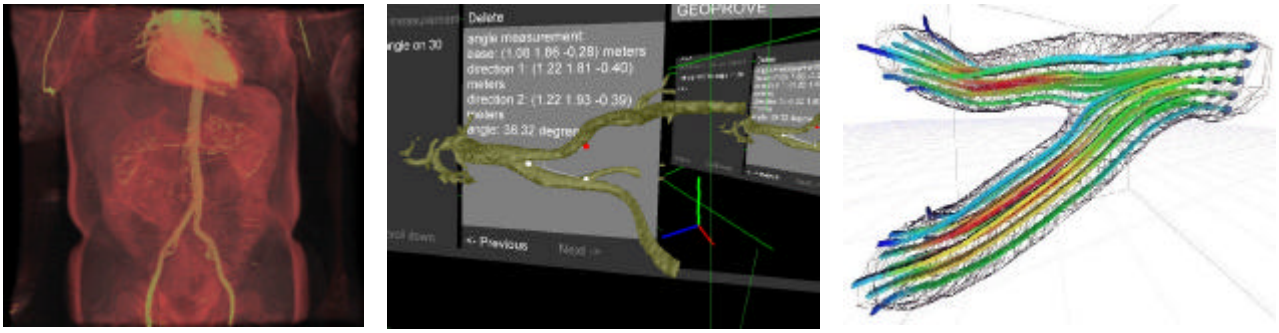

Fig 1. Immersive version of the VRE. Left: Visualization of patient data (Zudilova and Sloot, 2003, image courtesy by D. Shamonin). Middle: Visualization of a vessel, GEOPROVE measurement system menus (Belleman et al., 1999). Right: 3D flow visualization (Belleman, 2003, image courtesy by H. Ragas)

4 Heuristic evaluation

A heuristic evaluation of the VRE prototype has been conducted to gain insight into the potential usability problems of the system. Heuristic evaluation by experts is a reasonable quick and cheap method for evaluating usability of early prototypes (Nielsen, 1994). Staff working with the early prototype already identified some usability problems and it was doubted whether it would be robust enough for an evaluation with end-users. The observations relating to the prototype stage of the VRE, combined with the relative low cost of carrying out a heuristic approach balanced against the cost of recruiting representative users of this very early prototype, made qualitative heuristic evaluation of the system a more viable option at this stage of development. It has been realized that for a complete identification of all usability problems heuristics have to be adapted to context and system. Dix et al.,(1998) for example point out that usability metrics rely on measurements of very specific user actions in very specific situations. It should be understood why and how particular usability metrics are applicable in certain situations and how they provide for greater usability for real users. Before a complete usability evaluation with users can be a meaningful instrument, it has to be known who these prospective users of a system will be, what tasks a system should support, and what the context(s) of use will be. This is not the case in an early design stage such as the VRE is in. Therefore this first evaluation has focused on finding the main usability issues of the current interface and functionality of the VRE. It does not yet give an overview of issues that might come into play outside a lab setting and it is acknowledged that subsequent evaluations with users should be performed. Future quantitative testing of new VRE versions is planned. The initial evaluation does however yield insight in the relatively context independent issues that come into play in using a VR interface for certain applications. Usability issues, identified during the evaluation, also gave insight into issues that are related to the future context of use of the system and need to be studied.

It has been pointed out in literature that traditional usability methods do not take into account factors which come into play in VR systems, such as presence, physiological effects (e.g. motion sickness), issues of locating and manipulating objects, or disorientation and navigating in 3D worlds. Bowman et al., (2002) give an overview of usability methods applied to virtual environments. They notice that a lack of well-formed guidelines and heuristics for virtual environments user interface design and evaluation make heuristic evaluation challenging. Indeed a set of guidelines had to be adapted in this study. However, valuable insights and guidance into heuristics and specific evaluation of VR applications are provided by for example Kalawsky(1997), Steed and Tromp(1998), Kaur(1999), Suttcliffe et al., (2000), Neale and Nichols (2001), Willans et al., (2001). These guided us to a concise set of heuristics that have been used to identify problems during system walkthroughs using tasks representative of the VRE's current functionality. Tasks included placement of a virtual bypass on a virtual vascular structure and exploring its results.

Results evaluation

The walkthroughs and heuristic evaluation helped to identify the main context-independent usability problems of the current VRE prototype. The main problems of the prototype are highlighted in this section; a complete description of the results are given in Cramer (2004a). While some issues result from the prototype stage of the VRE and technological challenges of VR development in general, some problems however also seem to result from using the VR paradigm for this type of application. The former will be discussed first, then the latter seemingly paradigm-related problems will be highlighted. It should be noted that due to use of this heuristic evaluation, findings are subjective and can merely be seen as insight into the potential usability problems that may arise. However, an in-depth qualitative analysis of potential problems that might arise is useful to guide further research and development. Identifying potential user problem areas in the interface design can highlight the need to address these issues in development of the VRE and other VR systems. It can also provide insight in possible principal issues that have not yet been satisfactory addressed in VR research. To eliminate the subjective aspect of this study, further research will include quantitative analysis of the issues highlighted during the heuristic evaluation and user profiling (Zudilova and Sloot 2003).

A number of problems in using the functionality of the VRE resulted from limitations of the current hardware. Problems were apparent in the magnetic tracking system used. The magnetic properties of an assortment of items in the testing lab ranging file cabinets to metal in the walls themselves resulted in an intense jitter which made pointing and selection in the VRE very difficult. Some lag in real-time feedback resulted from limitations of computational resources in for example running and visualizing extensive simulation of blood flow. The jitter and lag at times rendered working with the VRE very difficult. A hardware related issue is the discomfort which users can experience while working with the prototype, not only in regard to simulator sickness, but also to the ergonomic qualities of the wand, headgear and the shutter glasses. Improvements in these areas are possible and it can be expected that usability of VR systems in general would benefit from hardware improvements (e.g. Belleman, 2003). The different input modalities were in an early prototype stage, an example how both this test-bed stage as well as more context related issues affect usability can be illustrated by considering speech recognition. The voice recognition system did not always correctly interpret commands from all users and not all relevant possible synonyms were included. In a real context of use speech recognition might be less desirable. The recognition ideally should be robust enough to handle environment noises and interruptions. Some users might find using speech recognition embarrassing (talking to a screen, being heard by others) or might be uncomfortable when other can hear that they make mistakes (irrespective whether these are actually a consequence of a system design problem) and the system 'does not understand'. This would certainly impair the social acceptability of a system. However, voice recognition could be a good alternative for users whose hands are occupied or find manual interaction in VR cumbersome. Context of use and user characteristics might be very important factors in usability of speech recognition.

Other usability problems seemed to result from the choice of a VR paradigm for the VRE application and may be important in (not) choosing a VR type application. When analysing the various VR evaluation methods it was noticed that most of these methods focus on applications offering a 'virtual world'. Examples of these are architectural walkthroughs, which aim to immerse the user in a complete environment. This fully immersive focus seems to make some of the heuristics in the literature (e.g. presence in the virtual world) less relevant to applications such as the VRE. Systems such as the VRE primarily support inspection and manipulation of virtual objects rather than exploration of a virtual world. The free exploration possibilities offered by the VRE, while system functionality focuses on 3D objects and not world exploration result in a number of usability problems. First of all navigation issues arose, due to the free exploration offered, users could get lost in the environment and lose sight of the vascular structure object of interest. This while the environment around the structures did not yield any information for a diagnostic or planning task. The combination of 2D and 3D objects in 3D space posed some problems as well. Virtual menus were modelled as 2D objects in a 3D world. This adheres to the paradigm of a virtual environment, but just as other virtual objects the menus could be moved out of sight, or could be explored from the side

Fig 2: Bypass placement in desktop VRE (Zudilova and Sloot 2003, image courtesy by D. Shamonin)

and back. This made a selecting a menu option from a selection task into a elaborated task that could include extensive navigation of the environment. Research into more optimal ways of offering menus in virtual environments could be of value. This also goes for offering of other 2D information that might be of use, e.g. 2D images or text. This is especially relevant for medical applications which aim to support clinicians who use 2D scans.

Solving these problems of free navigation and combining 2D and 3D in virtual environments might indicate a necessity to switch to another paradigm, perhaps moving away from VR or moving to hybrid variations. Certainly hybrid forms of interface paradigms are possible and known (e.g. CAD design systems, desktop 3D applications) and using fixed menus and restricting navigation within a more desktop type application might solve some navigation issues. A desktop VR version of the VRE (fig.2) is being developed to explore these issues.

5 Gaining insight in user requirements

Usability evaluations often also address utility issues, as noted as well by e.g. Hilbert and Redmiles (2000). A need to look into issues surrounding utility has indeed in this case been identified during usability evaluation as well. The heuristic evaluation resulted not only in the identification of potential usability problems, but also in the insight that a look at the context of use to gain insight in both desired functionality and possible context-dependent usability problems of the VRE would be necessary. The choice was made to conduct a small exploratory study as a first step to gain insight in the context of use of the VRE. In this paper only a limited number of aspects of the context analysis are highlighted to illustrate the value of supplementing and iterating usability evaluations with other methodologies. The context analysis study itself is extensively described in Cramer et al., (2004b). For this analysis, seven (interventional) radiologists and seven vascular surgeons at nine Dutch hospitals participated. Users, tasks, technical, social, organizational and physical environment that could potentially affect system use were studied using interviews and observation methods.

Usability problems were identified in the heuristic evaluation, but this did not yet show the applicability of a system's utility in a specific context of use. Like the heuristic evaluation of the VRE, the context analysis gave indication of a number of issues that have not yet been sufficiently addressed in VR development, but also to issues more specific to the VRE. An example of the changes needed to the simulation offered by the VRE is highlighted here, other results can be found in Cramer et al. (2004b). The VRE visualizes simulation of blood flow before and after a simulated intervention with placement of a stent prosthesis/ bypass. The simulation offered focused on the intervention result right after treatment, participants indicated in contrast that this direct result might be important but that a reliable prediction of long-term results of an intervention in terms of re-stenosis and blood flow would be much more valuable. Some findings suggest that blood flow simulation would only be used at all in an adapted form used in prediction of wall stress, calcification, or survival of a stent prosthesis over time. Currently, detailed blood flow information is often not used in advanced treatment planning, a reduction in diameter and thus flow in vessels is assessed, but no numerical data on flow is usually considered after an assessment of early tests. Simulation in the VRE would only be used if its added value to the quality of a treatment and the efficiency is clear to the prospective user, which is not the case for its current functionality. Possibilities have been identified as well for improved automatic segmentation of images and automatic measurements of artery properties following standards set by medical practitioners themselves.

It was possible to gain insights into necessary changes to the VRE's functionality during the heuristic usability evaluation. Apart from assessing the usability of existing functionality, it was also possible to evaluate whether the system's current functionality would match user needs. Contextual analysis of the environment of use provided insight into these needs. . Evaluating whether a system caters for the real-life demands and needs is especially important for systems that have originally been developed as a technological test-bed. It was concluded that qualitative methods are very suited to achieve such an in depth understanding of the context of use. It should be noted for future research that such an approach would benefit from iteration and supplementing with quantitative approaches as recommended by e.g. Nielsen (1994).

6 Conclusion and discussion

Even limited studies such as these bring to the attention the importance of VR related issues that are not yet fully understood. Future research should investigate which heuristic evaluation and contextual analysis methods are most suited to virtual reality systems. Attention needs to be given to usability issues relatively independent to context of use, as those that do depend on users, tasks and context. These issues and identified usability problems are not only relevant to the VRE, but may also yield insight in the issues that might come into play in development of all sorts of virtual environments. In this case, the prototype evaluation highlighted issues for future research for VR; limitations of current hardware (both technical and ergonomic), difficulty of use of input modalities available, navigation issues and difficulties related to the combining 2D and 3D objects were the most apparent. These issues are not yet resolved. Findings of both the heuristic evaluation and the contextual analysis were valuable not only in respect to the development of medical systems similar to the VRE, but also for more general application, e.g., when the choice between fully-immersive VR and non-immersive desktop solution has to be made. In this case, the VRE development team is now experimenting with various VR environments (Zudilova and Sloot, 2003).

Future research will be needed to gain detailed insight into the usability criteria and measurement methods of usability relevant for VR applications. Qualitative methods were found to be very useful to the development process of a virtual environment. It should be noted that communicating the relevance of qualitative results to a technically oriented development team may prove challenging. Iteration and combination of multiple research methods is desired in studying issues related to interaction with virtual environments and in validating of systems. The findings of the study described in this paper suggest that qualitative research methods can provide valuable insight into usability issues to inform the development of Virtual Reality applications.

7 Acknowledgements

The authors would like to thank all participants in our contextual study, the VRE development team and the DiME project team. We also appreciate the invaluable help of all those who engaged in useful discussions and offered their advice and help. The research was sponsored partially by the EU CrossGrid Project IST-2001-32243 and the Token 2000 project Distributed Interactive Medical Exploratory for 3D Medical Images.

8 References

Belleman, R.G. (2003). *Interactive Exploration in virtual environments*, PhD thesis, University of Amsterdam, The Netherlands.

Belleman R.G., Sloot, P.M.A. (2001). Simulated Vascular Reconstruction in a Virtual Operating Theatre. *Proceedings of the CARS 2001 Conference*, Berlin, Germany

Bowman, D.A., Gabbard, J.L. & Hix, D. (2002) A Survey of Usability Evaluation in Virtual Environments: Classification and Comparison of Methods.*Presence*, 4(11), 404-424

Cramer, H. (2004a) Usability evaluation and contextual analysis to support development of virtual reality systems; the Virtual Radiology Explorer, *Masters thesis* (forthcoming), Faculty of Science, University of Amsterdam, The Netherlands.

Cramer, H. (2004b) Context analysis to support development of virtual reality applications, submitted to *Virtual Reality*.

Dix, A., Finlay, J., Abowd, G., Beale, R. (1998) *Human-Computer Interaction*, 2nd Edition, Prentice Hall, UK.

Grudin, J. (1992) Utility and usability: research issues and development contexts , *Interacting with Computers*, 4(5).

Hilbert, D.M. & Redmiles D.F. (2000) Extracting Usability Information from User Interface Events, *ACM Computing Surveys*, 32 (4), 384-421.

Hughes J, O. B. J., Rodden T, Rouncefield M, Sommerville I (1993) Presenting ethnography in the requirements process, *Proceedings of the IEEE International Symposium on Requirements Engineering*, San Diego, USA.

International Standards Organization (1998) *ISO 9241-11:Ergonomic requirements for office work with visual display terminals (VDTs)* - Part 11: Guidance on usability.

Kalawsky,R.S. (1997) VRUSE - a computerised diagnostic tool: for usability evaluation of virtual/synthetic environment systems, *Applied Ergonomics*, 30 (1999), 11-25.

Kaur K., Sutcliffe, A.and Maiden, N. (1999) A design advice tool presenting usability guidance for virtual environments *Workshop on user centred design and implementation of virtual environments*, York, UK.

Neale and Nichols, S. (2001) Theme-based content analysis: a flexible method for virtual environment evaluation, *International Journal of Human-Computer Studies*, 55, 167-189.

Preece, J., Rogers, Y, Sharp, H. (2002) *Interaction Design*, John Wiley & Sons, New York, USA.

Sloot P.M.A. et al., (2003) Grid-based Interactive Visualisation of Medical Images, *Proceedings of the First European HealthGrid Conference*, 57 – 66 , Brussels, Belgium.

Steed, A., Tromp, J.G., (1998) Experiences with the Evaluation of CVE Applications, *Proceedings of Collaborative Virtual Environments Conference (CVE'98)*, Manchester, UK.

Sutcliffe, A.G., Kaur, K.D. (2000) Evaluating the usability of virtual reality user interfaces. Behaviour and Information Technology, 19(6), 415-426.

Willans, J.S. and Harrison, M.D. (2001) A toolset supported approach for designing and testing virtual environment techniques, International Journal of Human-Computer Studies, 55, 145-165.

Zudilova, E.V., Sloot, P.M.A., Belleman R.G. (2002) A Multi-modal Interface for an Interactive Simulated Vascular Reconstruction System, *Proceedings of the 4^{h} IEEE Int. conference on multimodal interfaces*, Los Alamitos, USA.

Zudilova E.V., Sloot P.M.A (2003) Virtual Reality and Desktop as a Combined Interaction-Visualisation Medium for a Problem-Solving Environment, *Proceedings of the International Conference on Computational Science*, Melbourne, Australia.