

Using and Evaluating HCI Techniques in Geovisualization: Applying Standard and Adapted Methods in Research and Education

Stephanie Larissa Marsh & Jason Dykes

giCentre, School of Informatics, City University London, EC1V 0HB

Tel. (+ 44 20 7040 8906) Fax (+44 20 7040 8845)

s.l.marshjad7 @soi.city.ac.uk, www.gicentre.org/jad7

KEYWORDS: usability engineering, evaluation, HCI, education, research, geovisualization

1. Introduction and Purpose.

Geovisualization involves the use of interactive software tools to prompt thinking about unknown datasets (MacEachren and Kraak, 2001). Successful use of visualization in research results in ‘ideation’ – the generation of ideas (MacEachren and Kraak, 1997). Successful visualization in education results in the communication of knowledge and improved understanding (Marsh and Dykes, 2005). Techniques from Human Computer Interaction Design (HCI) are becoming more prevalent in a field that relies upon rapid communication interactions between machine and user. MacEachren and Kraak (2001) describe a wide range of potential benefits of applying HCI methods to the geovisualization domain. They also describe some likely impediments. For example, HCI techniques include evaluations that are costly in terms of time and resources, particularly in the case of user testing and geovisualization may involve a small number of expert users. In most cases, the results of HCI experiments are specific to particular software / task / user configurations – indeed the standard definition of usability refers to specific users, specific goals and specific context of use (ISO 9241-11, 1998) - yet there is a need to generalize the findings of individual experiments understand geovisualization, evaluate the utility of HCI techniques and make geovisualization software more usable (Fuhrmann et al., 2005). The reliability and replicability of HCI evaluations is also questionable.

Geovisualization is not alone in using interactive graphics to explore structured data and draws upon disciplines such as information visualization (see Gahegan *et al.* 2001; Andrienko *et al.* 1999; Tobón, 2003). Many of the impediments and benefits are thus shared. Information visualization usability studies can inform geovisualization usability research, as effective evaluation in both domains is likely to require the adaptation of existing HCI techniques or the development of new paradigms and methods. Yet the dominant relationship between the locations of spatial phenomena and their representation on screen puts additional constraints on geovisualization. There is little or no geovisualization literature critiquing usability in the context of the domain’s specific requirements. Figure 1 shows these schematically by differentiating between software that is ‘standard’, for information visualization and for geovisualization using characteristics relating to users, interaction, tools and layout.

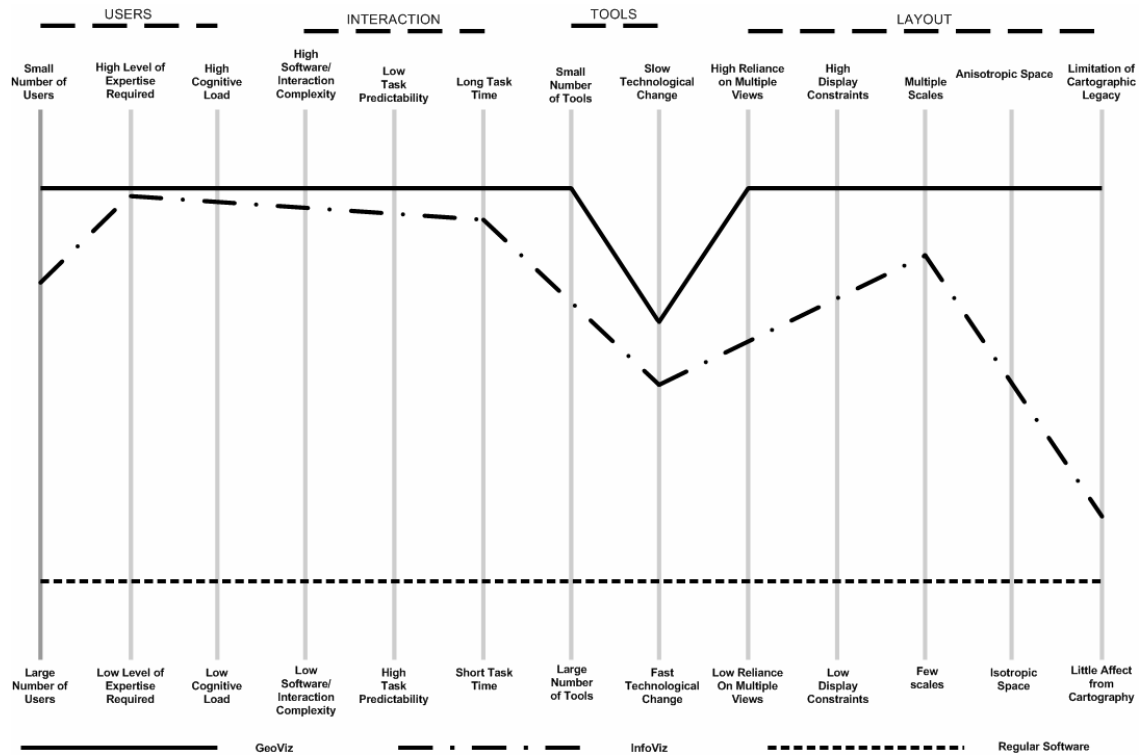


Figure 1. Qualitative representation of differences between geovisualization, information visualization and ‘regular’ software for which HCI techniques have been developed.

A variety of HCI methods can be utilised to evaluate tools, however little systematic research has been undertaken on the reliability and replicability of these methods. Appropriate evaluation solutions are likely to vary for different contexts making the selection and application of HCI techniques to geovisualization tools complex. Another challenge of applying evaluation techniques to geovisualization is the assessment of the support of ideation. How do we assess whether or not a tool supports ideation, knowledge discovery and hypothesis generation? Can ideation be measured and be an indicator of goal achievement? Such research problems are not considered widely by traditional HCI literature.

However, confidence in these techniques is crucial to the use of HCI in geovisualization and the development of geovisualization applications. Advances in geovisualization design can only be made if the evaluation findings are accurate and reliable. Consequently, we need to know the most efficient, effective and reliable ways to evaluate geovisualization tools, and the circumstances in which these techniques might be valid.

2. Approach.

An experimental framework that enable a variety of data collection techniques to be tested and compared under various conditions was developed, involving eight related experiments. This included both laboratory-based tests in which conditions were controlled and users evaluated ‘*in vitro*’ and evaluations ‘*in vivo*’ where geovisualization software was employed in real tasks (Robinson *et al.*, 2005). The framework operated in a linear fashion, enabling us to reflect upon the approaches and subsequently adjusted them, to suit the needs and requirements of geovisualization

and the situations in which they occurred. This enabled us to investigate the effectiveness and validity of candidate techniques for geovisualization in a number of contexts and draw general conclusions by triangulating evidence taken from a variety of tests. A number of candidate usability evaluation methods were considered, including: ethnographic observation, retrospective think-aloud, focus groups and questionnaires for experiments conducted in the field; onscreen interaction capture, concurrent think-aloud and task sheets for lab experiments (Buttenfield, 1999; Andreinko *et al.* 2003; Harrower *et al.*, 2000; Tobón, 2002; Zaphiris *et al.*, 2004; Bekker, *et al.* 2004; Maguire, 1998).

We evaluated geovisualization in research and education with a variety of users undertaking tasks of different type and complexity in a range of conditions. We focused on three aspects of software use in the majority of experiments: usability issues, ideation and interactions that occurred. These factors are integral to the effective use of geovisualization. Quantitative methods were used to analyse the numeric data collected, including ANOVA analysis. Qualitative data was subjected to content analysis to draw out evidence on which to derive results and confidence in them (Silverman, 2001). Triangulation was used to establish the strength of support for our findings.

3. Findings.

Our analysis allows us to make a number of inferences and draw conclusions about using and evaluating HCI techniques in geovisualization by applying standard and adapted methods in research and education. In summary, we have strong evidence that:

1. HCI techniques allow us to capture ideas that are generated in a research context and knowledge being generated in the context of learning through a process of visualization.
2. Different user groups do not perform consistently *in vitro*. *Sampling strategies have an effect on the results generated from methods involving small numbers of users and selection will thus influence some evaluation results in geovisualization.*
3. The inclusion of students in evaluation is only appropriate if evaluating pedagogic tools. *Students are often used as a substitute for experts to increase the numbers of participants in geovisualization evaluations (e.g. Andrienko et al., 2002). In comparative tests, we recorded considerable differences in behaviour. Though this behaviour was not statistically significant due to the small numbers of participants used, the statistics together with qualitative data constitute convincing evidence. Any results derived from studies that use students as expert surrogates should be treated with extreme caution.*
4. HCI techniques and methodologies (standard and adapted) are complementary and not interchangeable. *Alternative sets of methods may produce different results in geovisualization.*
5. Both standard and adapted HCI techniques successfully identified usability issues. *An overview of the results shows that 294 usability issues were identified over 3 in vivo studies, 3 in vitro studies and 1 remote study, involving 2 geovisualization tools and 1 geovisualization demonstrator. Furthermore, 143 issues (49%) were accepted for re-design by two tool developers.*
6. The setting, data collection techniques and user groups employed in the

evaluation influenced results in our comparative studies. *This is particularly true in relation to ideation – the main objective of geovisualization, but also in terms of usability responses and the interactions that occur. The effect is particularly evident when comparing in vitro and in vivo experiments.*

7. Short answer open-ended dataset directed exploratory tasks are more appropriate for facilitating ideation and identifying usability issues *in vitro* than instructional tasks. *Whilst they are less specific than many typical tasks used in HCI such tasks are more typical of geovisualization activity and result in more ideation.*
8. Onscreen interaction capture and think-aloud are the most appropriate data collection techniques in the *in vitro* setting during geovisualization usability evaluations. *These methods are least disruptive to the geovisualization process which involves a significant cognitive load and result in a higher volume of more useful data than alternatives.*
9. When conducting an evaluation with students, discursive and reflective methods yield the most useful and usable data. *Effective discursive methods include mediated focus groups and paired retrospective think-aloud. Reflective methods include structured questionnaires yield. These are particularly useful if scheduled immediately after learning activity.*
10. Satisfaction levels amongst students are affected by levels of training and preparation, motivation, time and resources. *Time allocated for geovisualization learning activity must be adequate for preparation as well as satisfactory task completion. Resources must support the tasks and learning activity fully. This has not been the case in some geovisualization user studies.*
11. Longitudinal *in vivo* studies with experts are important for understanding how geovisualization supports truly complex ideation. *The detailed case study approach may be the most appropriate as geovisualization takes time, requires expertise and the most effective solutions are likely to be data, task and user dependent.*

4. Practical Implications and Conclusions.

We have strong evidence that useful knowledge can be gained by employing HCI methods in geovisualization. Yet we find that experimental results may not be consistent between applications, users, situations and contexts. Our comparative experiments show that evaluation of geovisualization in research and education requires different methodological approaches due to the differing levels of expertise, motivation and knowledge of the participants involved. Thus, it is important that intended users are the subjects of evaluation and a small number of these is preferable to a large number of inappropriate users. Whilst statistically significant results may be achieved in the latter case these may not be relevant if test users do not behave in meaningful ways. Students are not a suitable substitute for experts in order to gain statistical significance. The evaluation of tools intended for the use in education can sample from students and visualization experts, but the evaluation of research tools should sample a range of domain and visualization experts even if the number sampled is small. This implication has a situational caveat, in that students work within their abilities in the *in vitro* environment, perhaps due its artificiality. However, *in vivo*, students will work at level that challenges their abilities, as they are motivated to learn, and this level of activity is essential for visualization to be effective. Visualization experts can simulate students with an intermediate level of expertise within an *in vitro* environment – but this is a potentially expensive way to test, that conflicts with the rationale for using students to simulate experts!

The different behaviours of different user groups in different situations adds complexity to the use of HCI techniques in geovisualization, particularly as more precision can be gained through *in vitro* than *in vivo* experiments. However, data from *in vivo* tests offer more ecological validity, particularly in the case of ideation.

Evaluation frameworks being implemented in education should include methods such as group sessions and questionnaires within both *in vitro* and *in vivo* environments. These techniques elicit comprehensive data on sophisticated visualization software use from students.

There have been several reports of the use of a variety of HCI techniques in geovisualization. However, there is no evidence of studies comparing different HCI techniques and establishing which techniques are most suitable in certain geovisualization circumstances. This research provides empirical evidence upon which geovisualization researchers may select and adapt HCI techniques for developing geovisualization applications and evaluating their use.

Whilst HCI studies in geovisualization are important and beneficial they are limited in their standard forms and may require some adaptation and coordinated use in within a well planned framework to provide meaningful results (Marsh, 2007). If we are to use HCI techniques and methodologies in geovisualization, then we must accept that results are subjective and when applied to small user groups, not statistically significant. However, this does not mean that the results are not scientifically important. The key is not to attempt to gain statistically valid results from geovisualization evaluation by using large numbers of potentially inappropriate users. Rather the focus should be on creating frameworks for conducting *in vivo* longitudinal studies with core users, when capturing data on usability, interaction and ideation. If we want to capture truly in-depth and ecological data, intensive longitudinal *in vivo* studies are essential. Finally, it is worth noting that these conclusions are not specifically geographic in nature; suggesting either that geovisualization evaluation may not differ in nature to that of evaluation in other exploratory domains in which graphical exploration plays a key role, or that the techniques utilised here do not provide the detail required to differentiate between graphical visual interfaces in which layout is constrained by geography and used to represent geographic relationships and those in which it is not.

5. Acknowledgements

The many participants who took part in the tests reported here are gratefully acknowledged and their time effort and cooperation is much appreciated. This includes students of the Masters in Geographic Information at City University London. The work conducted by Fenia Attilakou as part of her Masters thesis is gratefully acknowledged. This paper is based upon:

Marsh, S. L. (2007). Using and Evaluating HCI Techniques in Geovisualization: Applying Standard and Adapted Methods in Research and Education. PhD Thesis. City University.

References

- ISO 9241-11 (1998) Guidance on Usability, also issued by the International Organization for Standardization.
- Andrienko, G. and Andrienko, N. (1999). Interactive Maps for Visual Data Exploration. *International Journal of Geographical Information Science*. **13(4)**. 355-374.
- Andrienko, N., Andrienko, G., Voss, H., Bernardo, F., Hipolito, J. and Kretchmer, U. (2002). Testing the Usability of Interactive Maps in CommonGIS. *Cartography and Geographic Information Science*. **29(4)**. 325-342.

- Bekker, M., Barendregt, W., Crombeen, S. and Biesheuvel, M. (2004). Evaluating Usability and Challenge During Initial and Extended Use of Children's Computer Games. In *HCI 2004*, Vol. 1 (Eds, Fincher, S., Markopoulos, P., Moore, D. and Ruddle, R.). Springer, Leeds, pp. 331-346.
- Buttenfield, B. (1999). Usability Evaluation of Digital Libraries. *Science and Technology Libraries*. **17(3/4)**. 39-59.
- Fuhrmann, S., Ahonen-Rainio, P., Edsall, R. M., Fabrikant, S. I., Koua, E. L., Tobón, C., Ware, C. and Wilson, S. (2005). Making useful and useable geovisualization: design and evaluation issues In *Exploring Geovisualization*. (Eds, Dykes, J., MacEachren, A. and Kraak, M.-J.). Elsevier, Oxford, pp. 553-566.
- Gahegan, M., Wachowicz, M., Harrower, M. and Rhyne, T. M. (2001). The Integration of Geographic Visualization with Knowledge Discovery in Databases and Geocomputation. *Cartography and Geographic Information Science*. **28(1)**. 29-44.
- Harrower, M., MacEachren, A. and Griffin, A. L. (2000). Developing a Geographic Visualization Tool to Support Earth Science Learning. *Cartography and Geographic Information Science*. **27(4)**. 279-293.
- MacEachren, A.M. and Kraak, M.-J., 1997. Exploratory cartographic visualization: advancing the agenda. *Computers & Geosciences*, 23(4): 335-343.
- MacEachren, A. M. and Kraak, M.-J. (2001). Research Challenges in Geovisualization. *Cartography and Geographic Information Science*. **28(1)**. 1-11.
- Maguire, M. C. (1998). Usability Methods: Diary Keeping. http://www.lboro.ac.uk/eusc/g_diary_keeping.html. (24/03/05).
- Marsh, S. L. and Dykes, J. A. (2005). Using Usability to Evaluate Geovisualization for Learning and Teaching. In *Proceedings of the GIS Research UK 2005 13th Annual Conference* (Eds, Billen, R., Drummond, J., Forrest, D. and Joao, E.). University of Glasgow, Glasgow, pp. 477-483.
- Robinson, A. C., Chen, J., Lengerich, E. J., Meyer, H. G. and MacEachren, A. M. (2005). Combining Usability Techniques to Design Geovisualization Tools for Epidemiology. *Cartography and Geographic Information Science*. **32(4)**. 243-255.
- Tobón, C. (2002). Usability Testing in for Improving Interactive Geovisualisation Techniques. http://casa.ucl.ac.uk/working_papers/paper45.pdf. (2003).
- Tobón, C. (2003). Evaluating Geovisualization and User Interaction: an Experimental Design Approach Based Upon User Tasks. Department of Geography, University College London. PhD.
- Silverman, D. (2001) *Interpreting Qualitative Data: Methods for Analysing Talk, Text and Interaction*, 2. Sage, London.
- Zaphiris, P., Gill, K., Ma, T. H.-Y., Wilson, S. and Petrie, H. (2004). Exploring the Use of Information Visualization for Digital Libraries. *The New Review of Information Networking*. **10(1)**. 51-69.

Biographies

Stephanie Marsh was awarded a PhD in Information Science at City University in September 2007. She now works for HSBC.

Jason Dykes was awarded his PhD in Geography at the University of Leicester in December 1999. He now works at the giCentre, City University London.