POSITION STATEMENT
Evaluating Information Visualizations: Issues and Opportunities

John Stasko
College of Computing/GVU Center
Georgia Institute of Technology
Atlanta, GA 30332-0280
1 404 894 5617
stasko@cc.gatech.edu

ABSTRACT
The evaluation of information visualization techniques and systems is important in order to understand how to improve the systems and create innovative, useful new techniques. This article describes some of the challenges and issues in information visualization evaluation, and it discusses potential opportunities to improve the process. By better identifying the tasks people perform using information visualization tools as well as the insights gained therein, one can conduct more rigorous and comprehensive assessments of the tools.

Categories and Subject Descriptors
H.5.2 [**Information Interfaces and Presentation (e.g., HCI)**]: User Interfaces.

General Terms
Human factors, experimentation, measurement.

Keywords
Information visualization, evaluation, user tasks, knowledge discovery.

1. INTRODUCTION
The early years of research in the area of information visualization were ripe with new and innovative techniques being proposed. Many visualization and interaction techniques were introduced that now are accepted as fundamental components of the area. As information visualization has matured, however, researchers within the discipline have become more reflective and questioning of some of the area's basic principles.

It is still relatively common to see new information visualization techniques proposed that have high "wow" value: being visually striking and products of impressive technical engineering. Unfortunately, the potential value of many of these techniques in actually helping end-users solve problems and perform tasks is dubious in many cases.

I argue that while all newly proposed techniques and visualizations need not be accompanied by evaluation studies investigating their value, each should at least be accompanied by argumentation and discussion of why the technique may be useful and how people may benefit from it. Why is it important to discuss the utility of information visualization systems? Fundamentally, we want to ensure that the methods being proposed are improving the state-of-the-art and adding value. We want to learn what aspects of a visualization system provide value and how and why. It is extremely important to know both the strengths and weaknesses of a technique because this helps to inform us when to use that technique. Finally, we want to insure that new techniques minimally have the potential to help people and are not just "cool."

How do we evaluate information visualization techniques and systems? Plaisant provides a good overview of the potential evaluation techniques to apply [4]. Controlled experiments are good for measuring performance and comparing multiple techniques. Metrics such as correct task completion, completion time, and errors made can be investigated. Subjective assessments that uncover people's opinions and views about a system can be valuable too. Did a person find the use of a system enjoyable, confusing, fun, difficult? Personal opinion and judgment strongly influences use and adoption, sometimes even overcoming performance deficits. Qualitative, observational evaluations also can provide much insight about the utility of an information visualization system. Such studies are helpful to determine whether a system is being used as expected and can suggest new designs and improvements to the technique or system being examined.

2. ISSUES AND OPPORTUNITIES
Evaluating an information visualization technique or system is a challenging task. In assessing a system, one must investigate both the usability and the utility of the system. Over the years, the field of human-computer interaction has provided many good methods for evaluating usability and I feel that we have a reasonable handle on doing that. Assessing the utility of a system, however, is a much more challenging task, particularly in the field of information visualization.

Why is assessing the utility of information visualizations so difficult? Many factors contribute to this challenge. Plaisant has identified difficulties in 1) Matching tools with users, tasks, and real problems; 2) Improving user testing; 3) Addressing universal usability [4]. Here, I build on these challenges and provide more

reasons why evaluating information visualization systems is so difficult.

Information visualization systems are often developed for use by experts and analysts in specific fields of study. For instance, a visualization system may be designed to aid financial analysts modifying investment portfolios or to help microbiology researchers investigating new drug treatments. Such fields are highly specialized and the people working within them have a great deal of domain knowledge. Thus, understanding the potential benefits and insights provided by a system may require a relatively deep understanding of the discipline of study. Furthermore, simply accessing people who can serve as participants in evaluation studies of the systems may be challenging.

A second reason why the evaluation of information visualization systems is so challenging is that relatively few systems in the area are designed to address the exact same problem. When two systems have differing goals and objectives, even if only slightly different, it is difficult to perform direct comparative evaluations of them. Successful evaluations include representative tasks for participants to perform. Clearly, the same tasks must be performed with each of the systems being compared in order for the evaluation to be fair. When some of the tasks being assessed are not exactly the kind of tasks a system was built for, then the evaluation simply will not be balanced and fair. Simply put, it may be very difficult to "compare apples to apples and oranges to oranges."

Evaluating information visualization techniques is also challenging because it is difficult to separate a technique from its specific system implementation. That is, the accompanying user interface and set of basic user interaction capabilities of a system can strongly influence the utility of said system and people's perceptions of the value of the system. For instance, a visualization technique that is very flexible and powerful can be undermined by a poorly designed user interface, one with confusing commands, poor graphic design, and difficult-to-learn options. Kobsa found such issues in his study of commercial information visualization tools [3]. Fundamentally, the usability of a system can strongly affect its perceived utility and thus cloud evaluations of the true intrinsic merit of the techniques being provided.

The lack of standard, accepted evaluation tests and techniques in the field of information visualization is yet another challenge. Such tests and techniques would help researchers who are not so experienced in evaluation to assess and evaluate their new systems. The variety and diversity of information visualization techniques may make standardized evaluations unlikely and impractical, however.

Finally, I feel that the evaluation of information visualizations is challenging primarily because of the actual target purpose of systems in this area. When a software tool addresses a problem with concrete, easily-quantifiable outcomes, then evaluation of the tool is relatively straightforward. For instance, when we can clearly determine if the user of a system has achieved a desired outcome, then evaluation of the system is made easier. In the field of information visualization, however, outcomes of system use are much more nebulous and difficult to clearly articulate and quantify. Often, information visualizations are best-suited for browsing and general data exploration rather than solving specific problems.

When it is difficult to specify the task being performed or the problem being solved by a system, then it is very challenging to identify the criteria for evaluating the system and the values to be measured for evaluation.

My research group has worked the past few years on better understanding the tasks performed by people when using information visualizations. End-user tasks may range from high-level, difficult-to-quantify tasks like learning a new domain, predicting the future, and discovery of unknown relationships down to more low-level, specific user tasks such as correlation and outlier detection.

In a recent study of students answering questions and learning about data sets from five different domains [2], we identified the fundamental low-level user analytic tasks:

- Retrieve Value
- Filter
- Compute Derived Value
- Find Extremum
- Sort
- Determine Range
- Characterize Distribution
- Find Anomalies
- Cluster
- Correlate

Each of these tasks can be considered a different dimension upon which to evaluate an information visualization system. Moreover, in the absence of a formal evaluation study, the developers of systems can discuss how their system supports each of these then analytic inquiries. Thus, this set of low-level tasks provides a baseline or common language to use when discussing information visualization utility.

The benefits of information visualization systems on broader, higher-level user analytic activities such as knowledge discovery, complex decision making, and identifying the nature of trends is even more challenging. Here, evaluations that focus on insights gained while using a system appear much more valuable. A good step in this direction is the recent study by Saraiya, North, and Duca identifying the insights about a microarray biology data set brought about by different systems [5]. One challenge, however, is the subjective nature of insight. Whether a finding is a new insight clearly could be debatable. I feel that this is a key future research area in the evaluation of information visualization. We must develop better techniques for identifying a system's benefits for knowledge discovery, understanding, and insight.

My research group has argued that information visualization systems often are lacking in support for such high-level analysis and discovery processes, and systems need to provide a more comprehensive level of "analytic primacy" [1]. This term characterizes a system's primary focus on helping people with their analyses, as opposed to simply providing accurate depictions of data which we term "representational primacy." We have identified a number of methods for achieving higher analytic primacy in information visualization systems such as

- Determine domain parameters
- Expose multivariate explanation
- Facilitate hypothesis testing
- Expose uncertainty
- Concretize relationships
- Expose cause and effect

Another potentially fruitful avenue of evaluation, therefore, is to assess the level that information visualization systems support each of these objectives. Again, a key notion here is to provide relatively specific dimensions or metrics against which information visualization techniques and tools can be judged.

Finally, we must always be careful in evaluating information visualization, however, to not impose too rigid evaluation requirements that could stifle innovation. For example, a new technique may not fare well in an initial study of its utility, but ideas from that technique may influence other designers who influence other innovators who eventually produce an impactful system. Such chains of influence are extremely common in the research community, and it is important to allow such processes to occur in the future.

3. CONCLUSION

In this article I have identified the importance of information visualization evaluation as well as the challenges and issues in this area. I described potential ways to improve system evaluation and opportunities for new research to be conducted. By better understanding the tasks and goals of people and by better identifying the insights information visualization systems facilitate, we will move our discipline forward to richer and broader levels of understanding.

4. ACKNOWLEDGMENTS

This research has been supported in part by the National Science Foundation under awards IIS-0118685 and IIS-0414667. I would like to thank the other members of the Information Interfaces Research Group for their help in developing these ideas.

5. REFERENCES

[1] Amar, R. and Stasko, J. Knowledge Precepts for Design and Evaluation of Information Visualizations. *IEEE Trans. on Visualization and Computer Graphics*, *11*, 4 (Jul/Aug. 2005), 432-442.

[2] Amar, R., Eagan, J., and Stasko, J. Low Level Components of Analytic Activity in Information Visualization. In *Proceedings of the IEEE Symposium on Information Visualization (InfoVis '05)*, (Minneapolis, MN, October 2000), 111-117.

[3] Kobsa, A. An Empirical Comparison of Three Commercial Information Visualization Systems. In *Proceedings of the IEEE Symposium on Information Visualization (InfoVis '01)*, (San Diego, CA, October 2001), 123-130.

[4] Plaisant, C. The challenge of information visualization evaluation. In *Proceedings of the 2004 ACM Working Conference on Advanced Visual Interfaces (AVI '04)*, (Gallipoli, Italy, May 2004), 109-116.

[5] Saraiya, P., North, C., Duca, K. An Insight-Based Methodology for Evaluating Bioinformatics Visualizations. *IEEE Trans. on Visualization and Computer Graphics*, 11, 4 (Jul/Aug. 2005), 443-456.