

Visualization Viewpoints

Toward Measuring Visualization Insight

Editor: Theresa-Marie Rhyne

Presented By:

- **Dipkumar Patel**
- **Noopa Jagadeesh**
- **Prasanth Varma**

Agenda

- Introduction
- Defining insights
- Controlled experiments on benchmark tasks
- Eliminating benchmark tasks
- Conclusion and Future Scope

Introduction

- 2005 IEEE Visualization conference had many of the presentations related to evaluating visualizations.
- There were also papers related to Boolean usability studies that offered two alternatives: either the users liked the visualization tool in question, or they did not.
- There was papers related to variety of rigorously controlled experiments.

Introduction



What is the ultimate purpose of visualization and how it should be evaluated?

Defining insight

- Insight has been commonly stated as the broader purpose of information visualization by many authors.
- However, for the most part, the definition of insight remains fairly informal, making success difficult to evaluate.
- A default and implicit definition is to equate insight with user tasks, such as finding extreme values.

Defining insight

- The essential characteristics of insight are:

Complex

Deep

Qualitative

Unexpected

Relevant

Evaluating visualizations

A variety of evaluation methods exist:

- Controlled experiments
- Usability testing
- Longitudinal studies
- Heuristic evaluation
- Cognitive walkthroughs

Evaluation of Visualization

- What is a controlled experiment?
- What is a benchmark test?



Characteristics

- Predefined Tests
- Definitive Completion Times
- Definitive and Simple Answers

Drawbacks of Benchmark Tasks

- Predefined tasks can be biased. The experiment results are limited to only the tasks that evaluator chose.
- A further confusion in the interpretation of benchmark experiment results is the tradeoff between performance and accuracy.


Toward insight: more complex benchmark tasks

- To include benchmark tasks of greater complexity. For example, a task to characterize the distribution of data values (normal/ uniform/ linearly increasing).
- Another possibility is estimation tasks.

Difficulties in Complex benchmark texts

- Greater variability in task times and correctness.
- Greater difficulty in designing isomorphic tasks.
- Together, these problems generally mean that we must test more participants to get statistically significant results.

Toward insight: eliminating benchmark tasks

-  - change the benchmark tasks from an independent to a dependent variable
- Instead of instructing users in exactly what insights to gain, researchers observe what insights users gain on their own.
- Open ended protocol - Users explore the data in a way that they choose

Toward insight: eliminating benchmark tasks(contd...)

- Users are instructed to explore the data and report their insights until they feel that they have learned all that they can from the data
- Coding method metrics:
 - Insight category
 - Complexity
 - Time to generate
 - Errors
 - Depth - [coded 1 to 5]
- Users should be from the **target domain**

Toward insight: eliminating benchmark tasks(contd...)

Advantage: It reveals what insights visualization users gained

Researchers can then compare visualizations on insight related measures such as,

- **Number** of insights
- **Speed** at which insights were generated
- Domain value of the insights
- Insights gained by users vs. insights **expected**
- Which features of the visualization **helped** achieve insight, and which caused problems for the users?

This directly leads to visualization refinement and improvement

Toward insight: eliminating benchmark tasks(contd...)

Problems:

- Long insight **time** depending on data and domain complexity
- **More effort** by the experimenters to capture and code results
- Motivated, domain knowledgeable users who will not merely follow instructions but generate insight in a self-directed manner
- Domain experts to assist in coding results along with visualization experts

Conclusion

- Both types of controlled experiments are needed:
 - Benchmark tasks experiments: Identifies low-level effects
 - Eliminating Benchmark Tasks: Provide **richer view** of the broader insight of visualization
- Expects a better results, If both approaches are combined into single ✓

Future Steps:

- Try running the same experiment with each method to see how they differ
- Examine and adapt other uncontrolled evaluation methods to better gauge insight

Ultimately determine whether visualizations are achieving their grand purpose

Resources

1. R. Amar, J. Eagan, and J. Stasko, “Low-Level Components of Analytic Activity in Information Visualization,” Proc. IEEE Symp. Information Visualization, IEEE Press, 2005, pp. 111-117
2. C. Plaisant, “The Challenge of Information Visualization Evaluation,” Proc. Working Conf. Advanced Visual Interfaces (AVI), ACM Press, 2004, pp. 109-116.
3. M. Tory and T. Möller, “Human Factors In Visualization Research,” IEEE Trans. Visualization and Computer Graphics, vol. 10, no. 1, 2004, pp. 72-84
4. C. Chen, and Y. Yu, “Empirical Studies of Information Visualization: A Meta-Analysis,” Int’l J.Human–Computer Studies, vol. 53, no. 5, 2000, pp. 851-866
5. P. Saraiya, C. North, and K. Duca, “An Insight-Based Methodology for Evaluating Bioinformatics Visualizations,” IEEE Trans. Visualizations and Computer Graphics, vol. 11, no. 4, 2005, pp. 443-456

Thank You!

Any Questions?

