

Eurographics 2011  
LLANDUDNO UK  
11-15 April

Tutorial Notes

# Scientific Evaluation in Visualization

**Camilla Forsell**

C-Research, Linköping University  
Norrköping, Sweden  
camilla.forsell@liu.se

**Matthew Cooper**

C-Research, Linköping University  
Norrköping, Sweden  
matthew.cooper@liu.se





# Scientific Evaluation in Visualization

## Tutorial instructors

Camilla Forsell

Department of Science and Technology

C-Research, Visual Interactive Data Analysis

Linköping University

Norrköping, Sweden

camilla.forsell@liu.se

<http://www.itn.liu.se/mit/research/visual-interactive-data-analysis?l=en>

Matthew Cooper

Department of Science and Technology

C-Research, Visual Interactive Data Analysis

Linköping University

Norrköping, Sweden

matthew.cooper@liu.se

<http://www.itn.liu.se/mit/research/visual-interactive-data-analysis?l=en>

## Length

Half day

2X90 minutes

## Level

Introductory

## Who should attend?

The tutorial targets PhD students, researchers and practitioners interested in usability and evaluation studies in Visualization. Attendees are assumed to be knowledgeable in Visualization but not in the field of evaluation.

## Abstract

The objective of this half-day introductory tutorial is to increase awareness of what constitutes a sound scientific approach to evaluation in Visualization and to provide basic theoretical knowledge of and practical skills in current research practice. The content presents the current challenges and trends related to how to characterize and optimize the complex interactive visual displays present in Visualization today. It will cover the most basic and relevant issues to consider during different phases of evaluation: planning, design, execution, analysis of results and reporting. The content outlines how to proceed to achieve high quality results and point out common pitfalls and mistakes which are threats to high quality results. The main focus is on quantitative experimental research but the general knowledge applies to all kinds of studies. The tutorial will present the main part of the content by means of a lecture style using power-point presentations, and will use example studies

from the tutorial leaders' own publications as well as other relevant work. There will also be demonstrations of different techniques for capturing data during an evaluation study. The participants will be given the opportunity to try out some of these methods hands-on to further facilitate a discussion of their potential suitability for different kinds of studies. Taking part in this tutorial will not train a novice participant to be fully capable of designing and conducting an evaluation study and analyzing its outcome, such a goal would require a substantially larger course. The aim is to introduce the topic, provide a general knowledge about what is important to consider and what resources are available to guide them in further study in this area. Further, participants will also learn to better judge the relevance and quality of a publication presenting an evaluation when reviewing such work since the same rules apply.

## Resume of tutorial instructors

Dr. Camilla Forsell works as an assistant professor in Information Visualization with C-Research, a group associated with the Norrköping Visualization Centre and a part of the Department of Science and Technology, Linköping University, in Sweden. She has a MSc. in Cognitive Science from Linköping University, Sweden (2003) and a Ph.D. in Human-Computer Interaction from Uppsala University, Sweden (2007). Her main research interests are in perceptually-motivated constraints on Visualization and in development and validation of new methods and metrics for evaluation in Visualization. Camilla teaches undergraduate level courses in Information Visualization and Human-Computer Interaction Interfaces and graduate level courses in Scientific Research Methodology. Camilla is also course leader and teacher at diploma courses on Visualization given to Swedish industry and the public sector by Linköping University. She is also symposium committee chair and head organizer of the International Symposium on Information Visualization Evaluation (IVE) (<http://www.graphicslink.co.uk/IV10/IVE.htm>), part of the IEEE International Conference on Information Visualization (<http://www.graphicslink.co.uk/IV10/>).

Dr. Matthew Cooper is Senior Lecturer in Information Visualization with C-Research, a group associated with the Norrköping Visualization Centre and a part of the Department of Science and Technology, Linköping University, in Sweden where he has been working since 2001. He holds an MSc degree in Computer Science and a PhD in Computational Chemistry from the University of Manchester and was employed at the Manchester Visualization Centre from 1996-2001 before moving to Sweden. His work in Visualization has dealt with applications in climate and weather data, medicine, the social sciences, Engineering, Physics, Chemistry, and air traffic control. Cooper currently lectures in courses in Introductory and Advanced Computer Graphics, Information Visualization, and Virtual Reality Technologies. He is an author of approximately 60 refereed journal and conference papers in the areas of Interactive Visualization and Virtual and Augmented Reality. His current research interests are in the development of new visualization methods and techniques to aid in the understanding of large and complex time-varying data, and evaluation has formed a more and more important component of this work over the years.

# Agenda

## **Part 1**

### Introduction (20 min)

- Presentation of the importance of evaluation and current trends
- Motivation and goals

### Usability centred evaluation and some basic and important concepts (30 min)

- The main aims of conducting an evaluation
- Definition and examples of methods for evaluation
- Usability dimensions and metrics

### The main phases of an evaluation study with their associated choices and operations (40 min)

- Planning and study design
- Assignment procedures, execution, analysis of results
- Importance and implications of operational definitions, reliability and validity

## **Part 2**

### Example studies from published work (60 min)

- Further explanation of the content from Part 1 that show how theory, methods and guidelines relate to practice
- Demonstrations of techniques for capturing data

### Reporting (15)

- The importance of replicable information
- Structure and content
- Review others reported work

### Closure (15 min)

- Conclusion, questions and discussion
- Short evaluation (a few questions) of the tutorial



# Scientific Evaluation in Visualization

Camilla Forsell  
C-Research, Linköping University  
Norrköping, Sweden  
camilla.forsell@liu.se

Matthew Cooper  
C-Research, Linköping University  
Norrköping, Sweden  
matthew.cooper@liu.se



## What will you gain?

- The aim is to introduce the topic, provide a general knowledge about what is important to consider and what resources are available to guide you in further study in this area
- You will also learn to better judge the relevance and quality of a publication presenting an evaluation when reviewing such work since the same rules apply

## Who should attend?

- PhD students, researchers and practitioners interested in usability and evaluation studies in Visualization
- Attendees are assumed to be knowledgeable in Visualization but not in the field of evaluation

## Agenda

### Part 1

Introduction (20 min)  
Presentation of the importance of evaluation and current trends  
Motivation and goals

Usability centred evaluation and some basic and important concepts (30 min)  
The main aims of conducting an evaluation  
Definition and examples of methods for evaluation  
Usability dimensions and metrics

The main phases of an evaluation study with their associated choices and operations (40 min)  
Planning and study design  
Assignment procedures, execution, analysis of results  
Importance and implications of operational definitions, reliability and validity

### Part 2

Example studies from published work (60 min)  
Further explanation of the content from Part 1 that show how theory, methods and guidelines relate to practice  
Demonstrations of techniques for capturing data

Reporting (15 min)  
The importance of replicable information  
Structure and content  
Review others reported work

Closure (15 min)  
Questions and discussion  
Conclusion  
Short evaluation (a few questions) of the tutorial

## Part 1

### Introduction

## Introduction

- The importance of evaluation
- Motivation and goals

## Part 1

Usability centred evaluation and some basic and important concepts

## Usability

European definition:

usability investigated by methods based on the International Organization for Standardization (ISO) standard 9241 (1998) part 11 which addresses guidance on the specification and measurement of usability.

ISO 9241-11 defines usability as:

The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.

## Usability

- Effectiveness
  - accuracy and completeness
- Efficiency
  - the resources consumed in relation to the accuracy and completeness
- Satisfaction
  - comfort and acceptability

## Usability

Please note that a basic concept of this definition is that usability is not a property in itself but dependent on:

different types of users, their different goals and the context in which they are active!

## Usability

- Should not be mixed up with functionality
  - only the *presence* of functionality
  - greater functionality does not imply better usability

*"It really does not matter if a product, hypothetically, can do what you want from it if you cannot do it since it is too hard to understand and operate. Also, it does not matter if something is easy to do if that's not what you want to do with it".*

## Usability is about

- Easy to learn
  - Efficient to use
    - time to do specific tasks
  - Few errors
    - amount, severity, time to recover
  - Easy to remember
  - Subjectively pleasing
- } ISO efficiency
- } ISO effectiveness
- } ISO subjectivity



## Metrics

- Performance metrics
- Preference metrics
- Predictive metrics or design metrics

## Metrics

- Metrics for Visualization

## Evaluation - definition

Evaluation is the systematic assessment of the worth or merit of some object

Evaluation is the systematic acquisition and assessment (of information) to provide useful feedback about some object

## Evaluation - goal

The general goal of an evaluation is to get "useful feedback"

- Find (potential) problems that users will encounter in "real use"
- Measure usability towards usability goals

## Evaluation - method

- Observation
- Think aloud
- Interview
- Questionnaire
- Inspection
- Measurements (performance related)

## Evaluation - classifications

- Qualitative
- Quantitative

## Evaluation - classifications

- Analytical
- Empirical

## Evaluation - classifications

- Formative
- Summative
- Explorative

## Part 1

The main phases of an evaluation study  
with their associated choices and operations

## Phases

- Plan
- Design
- Conduct
- Analyse
- Report

## Control

- Level-of-constraints
  - range from no control (naturalistic observation)
  - to a high degree (experimental research)
- Variables
  - manipulate one (or more) independent variable
  - investigate its effect on one (or more) dependent variable
- Confounding variables/factors

## Control

- Reliability
- Validity

## Data

- Real data
- Synthetic data

## Design

- What design to use depends on:
  - the aim of the study
  - resources / availability
  - confounding factors

## Design

- Examine the same participants or different groups of participants:
  - within-subjects design (within-groups, related, repeated measures)
  - between-subjects design (between groups, unrelated)
  - mixed design

## Participants

- Selection
- Number of participants

## Participants and assignment

- Assignment
  - in to groups
  - over conditions
- Randomization
- Counterbalancing
  - complete
  - partial

## Participants and assignment

- Assignment
  - in to groups
  - over conditions
- Randomization
- Counterbalancing
  - complete
  - partial

1	2	3
2	3	1
3	1	2

## Conducting

- Perform a pilot study
- Written instructions
- Training
- Collect relevant background information



## Analyse results

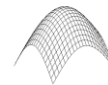
- Statistical test
  - follows from the design

## Analyse results

- Choose appropriate test/s
  - must be adequate to the nature of the data
  - parametric vs nonparametric tests
- Explore data
  - check assumptions
- Level of significance
  - post hoc

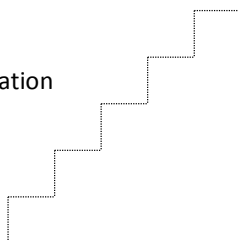
## Part 2

### Example studies from published work



## Part 2

### Demonstration



## Part 2

### Reporting



## Reporting

- How and what to report?
  - structure
  - content
- How to handle the trade-off between limited space and level of detail?

## Reporting

- The method section
  - should describe all relevant details about what you have done and how you proceeded when doing it
- Consider it as a *recipe*
  - enables replication

## Reporting (15 min)

- Fundamentals that need to be covered
- Some typical subsections are:
  - Stimuli (or material)
  - Apparatus
  - Participants
  - Experimental design
  - Procedure
  - Results

## Reporting

- Stimuli (or material)
  - details about stimuli and other materials used, data if not described earlier
- Apparatus and conditions
  - equipment, setup and usage conditions
- Participants
  - essential information about the people taking part
- Experimental design
  - structure and nature of the evaluation
- Procedure
  - how the evaluation was conducted in a practical sense
- Results
  - the results from applying the method
  - the results should stand for themselves i.e. no interpretation and/or discussion here in most cases

## Reporting

- The text should be presented in a way that is:
  - scientific
  - unambiguous
  - useful
- Terminology is highly important

Part 2

Closure