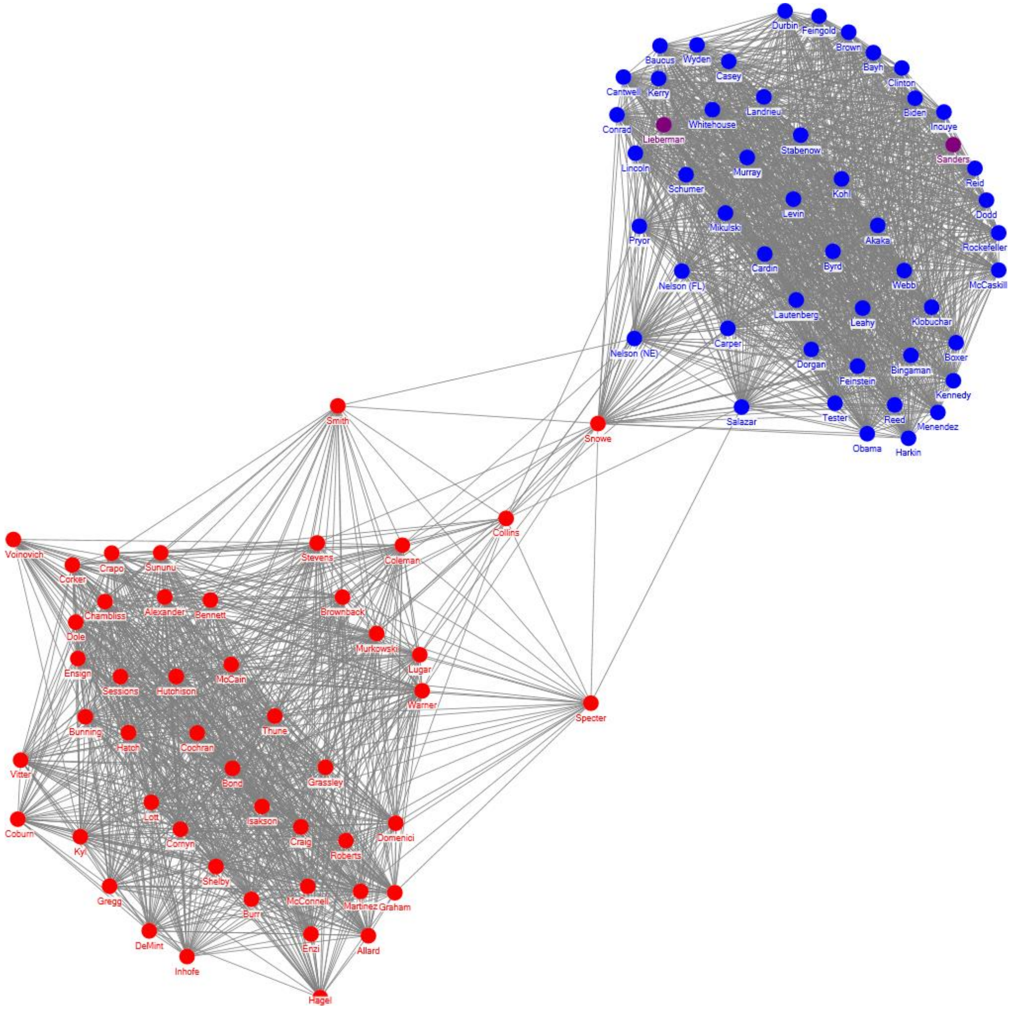
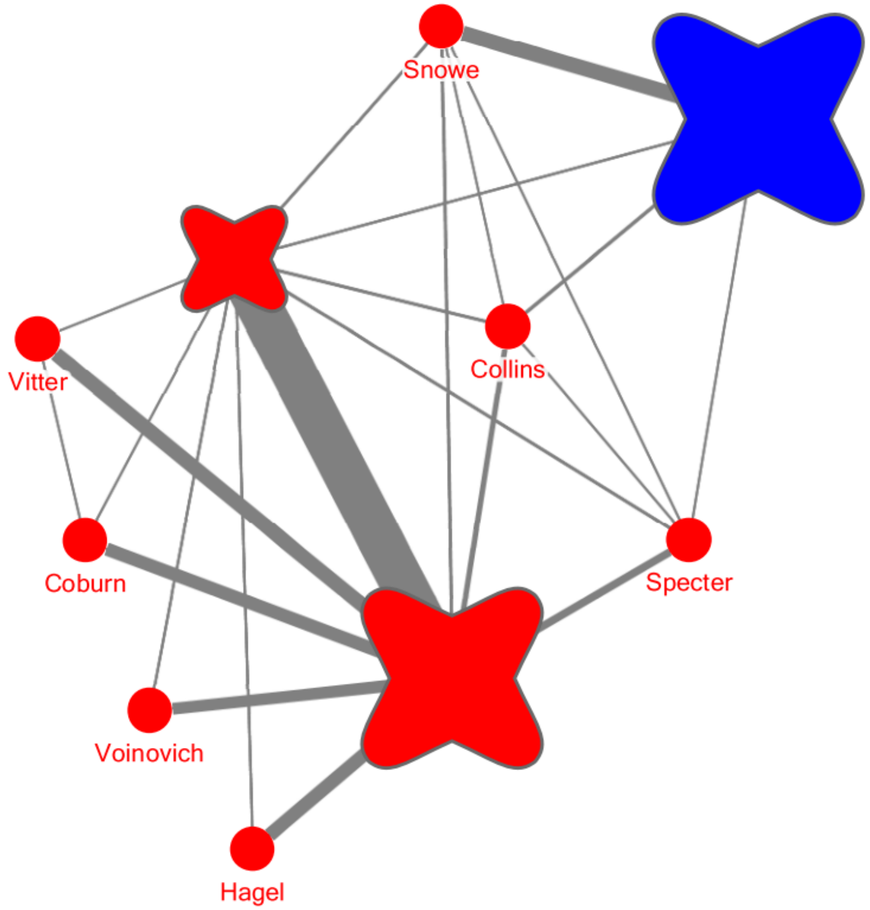


CS 7280-03 Special Topics on Visualization in Network Science

Lecture 11



Professor Cody Dunne

https://codydunne.github.io/cs7280-f16/
c.dunne@northeastern.edu

Presentation Signups

https://piazza.com/northeastern/fall2016/cs728003/home

Project Discussion

https://codydunne.github.io/cs7280-f16/project

Paper Prototyping

Google for Entrepreneurs

Rapid Prototyping

Part 1: Paper Prototyping

Discussion: Design Study Methodology

Design Study Methodology: Reflections from the Trenches and the Stacks

Michael Sedlmair, *Member, IEEE*, Miriah Meyer, *Member, IEEE*, and Tamara Munzner, *Member, IEEE*

Abstract—Design studies are an increasingly popular form of problem-driven visualization research, yet there is little guidance available about how to do them effectively. In this paper we reflect on our combined experience of conducting twenty-one design studies, as well as reading and reviewing many more, and on an extensive literature review of other field work methods and methodologies. Based on this foundation we provide definitions, propose a methodological framework, and provide practical guidance for conducting design studies. We define a design study as a project in which visualization researchers analyze a specific real-world problem faced by domain experts, design a visualization system that supports solving this problem, validate the design, and reflect about lessons learned in order to refine visualization design guidelines. We characterize two axes—a *task clarity axis* from fuzzy to crisp and an *information location axis* from the domain expert's head to the computer—and use these axes to reason about design study contributions, their suitability, and uniqueness from other approaches. The proposed methodological framework consists of 9 stages: *learn*, *winnow*, *cast*, *discover*, *design*, *implement*, *deploy*, *reflect*, and *write*. For each stage we provide practical guidance and outline potential pitfalls. We also conducted an extensive literature survey of related methodological approaches that involve a significant amount of qualitative field work, and compare design study methodology to that of ethnography, grounded theory, and action research.

Index Terms—Design study, methodology, visualization, framework.

1 Introduction

Over the last decade design studies have become an increasingly popular approach for conducting problem-driven visualization research. Design study papers are explicitly welcomed at several visualization venues as a way to explore the choices made when applying visualization techniques to a particular application area [55], and many exemplary design studies now exist [17, 34, 35, 56, 94]. A careful reading of these papers reveals multiple steps in the process of conducting a design study, including analyzing the problem, abstracting data and tasks, designing and implementing a visualization solution, evaluating the solution with real users, and writing up the findings.

And yet there is a lack of specific guidance in the visualization literature that describes holistic methodological approaches for conducting design studies—currently only three paragraphs exist [49, 55]. The relevant literature instead focuses on methods for designing [1, 42, 66, 79, 82, 90, 91] and evaluating [13, 33, 39, 50, 68, 69, 76, 80, 85, 86, 95] visualization tools. We distinguish between methods and methodology with the analogy of cooking; *methods* are like ingredients, whereas *methodology* is like a recipe. More formally, we use Crotty's definitions that methods are "techniques or procedures" and a methodology is the "strategy, plan of action, process, or design lying behind the choice and use of particular methods" [18].

From our personal experience we know that the process of conducting a design study is hard to do well and contains many potential pitfalls. We make this statement after reflecting on our own design studies, in total 21 between the 3 authors, and our experiences of reviewing many more design study papers. We consider at least 3 of our own design study attempts to be failures [51, 54, 72]; the other 18 were more successful [4, 5, 10, 40, 43, 44, 45, 46, 52, 53, 67, 70, 71, 73, 74, 75, 77, 78].

In the process of conducting these design studies we grappled with many recurring questions: What are the steps you should perform, and in what order? Which methods work, and which do not? What are the potential research contributions of a design study? When is the use of visualization a good idea at all? How should we go about collaborating with experts from other domains? What are pitfalls to avoid? How and when should we write a design study paper? These questions motivated and guided our methodological work and we present a set of answers in this paper.

We conducted an extensive literature review in the fields of human computer interaction (HCI) [7, 8, 9, 12, 16, 19, 20, 21, 22, 25, 26, 27, 28, 29, 30, 31, 38, 47, 57, 63, 64, 65, 83] and social science [6, 14, 18, 24, 32, 62, 81, 87, 93] in hopes of finding methodologies that we could apply directly to design study research. Instead, we found an intellectual territory full of quagmires where the very issues we ourselves struggled with were active subjects of nuanced debate. We did not find any off-the-shelf answers that we consider suitable for wholesale assimilation; after careful gleaning we have synthesized a framing of how the concerns of visualization design studies both align with and differ from several other qualitative approaches.

This paper is the result of a careful analysis of both our experiences in the "trenches" while doing our own work, and our foray into the library "stacks" to investigate the ideas of others. We provide, for the first time, a discussion about design study methodology, including a clear definition of design studies as well as practical guidance for conducting them effectively. We articulate two axes, *task clarity* and *information location*, to reason about what contributions design studies can make, when they are an appropriate research device, and how they are unique from other approaches. For practical guidance we propose a process for conducting design studies, called the *nine-stage framework*, consisting of the following stages: *learn*, *winnow*, *cast*, *discover*, *design*, *implement*, *deploy*, *reflect*, and *write*. At each stage we identify pitfalls that can endanger the success of a design study, as well as strategies and methods to help avoid them. Finally, we contrast design study methodology to related research methodologies used in other fields, in particular those used or discussed in HCI, and elaborate on similarities and differences. In summary, the main contributions of this paper are:

- definitions for design study methodology, including articulation of the task clarity and information location axes;
- a nine-stage framework for practical guidance in conducting design studies and collaborating with domain experts;
- 32 identified pitfalls occurring throughout the framework;
- a comparison of design study methodology to that of ethnography, grounded theory and action research.

We anticipate that a wide range of readers will find this paper useful, including people new to visualization research, researchers expe-

- *Michael Sedlmair and Tamara Munzner are with the University of British Columbia, e-mail: [msedl,tmm]@cs.ubc.ca.*
- *Miriah Meyer is with the University of Utah, e-mail: miriah@cs.utah.edu.*

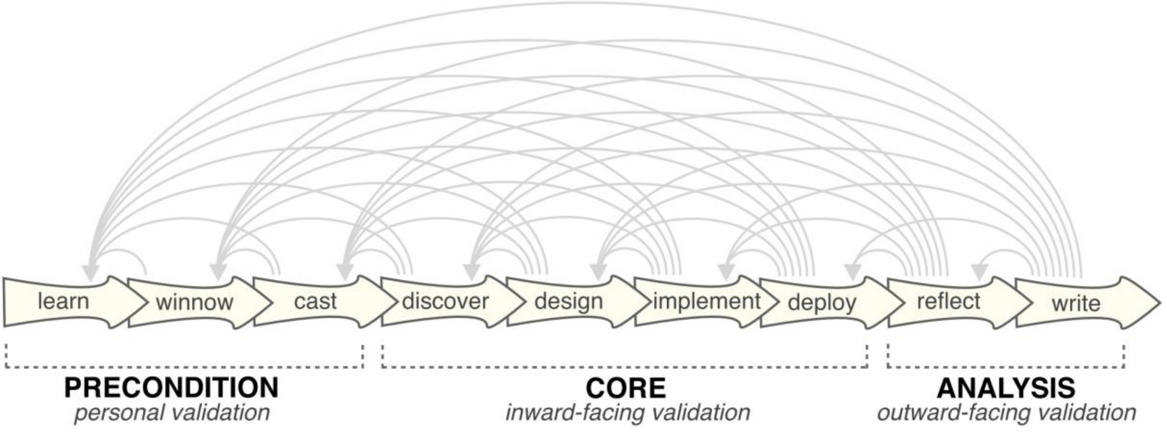
Manuscript received 31 March 2012; accepted 1 August 2012; posted online 14 October 2012; mailed on 5 October 2012.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

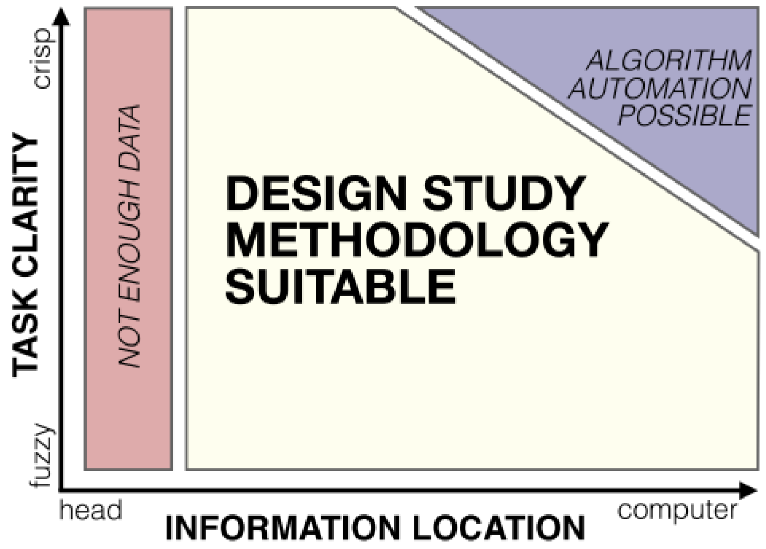
Design Study Research Contributions

- Problem characterization and abstraction
  - Shared understanding
  - Requirements
  - Automation
- Validated visualization design
- Reflection
  - Lessons learned

Transferability, not reproducibility

Sedlmair et al., 2012





Sedlmair et al., 2012

- Precondition
  - Learn – literature
  - Winnow – **data**, engagement, intellectual
  - Cast – roles
- Core
  - Discover – characterize & abstract problem, observe & inquire
  - Design – broad (parallel) -> narrow
  - Implement – rapid prototype, usability test
  - Deploy – real users, tasks, data, validation
- Analysis
  - Reflect – value for field
  - Write – lots of work, part of research

Sarah's 5 stages

- Before designing a study think carefully about what you hope to accomplish and what approach you need. (Describe axes as a tool for doing this).
- Think about what data you have and who needs to be part of the conversation. (pre-condition phase)
- Design and implement the study (core phase)
- Reflect and share your results (analysis phase)
- Throughout the process, be sure to think carefully about goals, timelines and roles (pitfalls)

Sedlmair et al., 2012

1 premature advance: jumping forward over stages
2 premature start: insufficient knowledge of vis literature
3 premature commitment: collaboration with wrong people
4 no real data available (yet)
5 insufficient time available from potential collaborators
6 no need for visualization: problem can be automated
7 researcher expertise does not match domain problem
8 no need for research: engineering vs. research project
9 no need for change: existing tools are good enough
10 no real/important/recurring task
11 no rapport with collaborators
12 not identifying front line analyst and gatekeeper before start
13 assuming every project will have the same role distribution
14 mistaking fellow tool builders for real end users
15 ignoring practices that currently work well
16 expecting just talking or fly on wall to work
17 experts focusing on visualization vs. domain problem
18 learning their problems/language: too little / too much
19 abstraction: too little
20 premature commitment: consideration space too small
21 mistaking technique-driven for problem-driven work
22 non-rapid prototyping
23 usability: too little / too much
24 premature end: insufficient deploy time built into schedule
25 usage scenario not case study: non-real task/data/user
26 liking necessary but not sufficient for validation deploy
27 failing to improve guidelines: confirm, refine, reject, propose
28 insufficient writing time built into schedule
29 no technique contribution 6= good study
30 too much domain background in paper
31 story told chronologically vs. focus on final results
32 premature end: win race vs. practice music for debut

Discussion: Insight-Based Study

An Evaluation of Microarray Visualization Tools for Biological Insight

Purvi Saraiya[1], Chris North[2]
Department of Computer Science
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061 USA

Karen Duca[3]
Virginia Bioinformatics Institute
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061 USA

ABSTRACT
High-throughput experiments such as gene expression microarrays in the life sciences result in large datasets. In response, a wide variety of visualization tools have been created to facilitate data analysis. Biologists often face a dilemma in choosing the best tool for their situation. The tool that works best for one biologist may not work well for another due to differences in the type of insight they seek from their data. A primary purpose of a visualization tool is to provide domain-relevant insight into the data. Ideally, any user wants maximum information in the least possible time. In this paper we identify several distinct characteristics of insight that enable us to recognize and quantify it. Based on this, we empirically evaluate five popular microarray visualization tools. Our conclusions can guide biologists in selecting the best tool for their data, and computer scientists in developing and evaluating visualizations.

CR Categories: H.5.2 [Information Interfaces and Presentation]: User Interfaces – Evaluation/Methodology, I.6.9 [Visualization] – Information Visualization, Visualization Systems and software, Visualization techniques and Methodologies

Keywords: Data visualization, empirical evaluation, insight, high throughput experiments, microarray data, bioinformatics

1 Introduction

Biologists use high-throughput experiments to answer complex biological research questions. Experiments, such as gene-expression microarrays [8], result in datasets that are very large. Due to its magnitude, microarray data is prohibitively difficult to analyze without the help of computational methods.

The advent of high-throughput experiments is causing a shift in the way biologists do research, a shift away from simple reductionist testing on a few variables towards systems-level exploratory analysis of 1000s of variables simultaneously. Hence, they use various data visualizations to derive biological inferences. The main purpose in using these visualizations is to gain insight into the extremely complex and dynamic functioning of living cells. In response to these needs, a large number of visualization tools targeted at this domain have been developed [2], [19] and [26].

However, in collaborations with biologists, we received mixed feedback and reviews about these tools. First, with so many tools to choose from, there is significant confusion among the biologists about which tool to use. Second, because of the open-ended and exploratory nature of the tasks, it is unclear how and if these tools are meeting their needs in providing insight.

The main goal of the research reported in this paper is to evaluate some of the most popular visualization tools for microarray data analysis, such as Spotfire® [30]. The key research questions are: How successful are these tools in assisting the biologists in arriving at domain-relevant insights? How do the various visualization techniques affect users' perception of data? How does the user's background affect the tool usage?

Typically, visualization evaluations have focused on controlled measurements of user performance and accuracy on predetermined tasks. However, to answer these research questions requires an evaluation method that more closely matches the exploratory nature of the biologists' goals. We devise and deploy an insight-based approach to visualization evaluation that we believe can be generally applied in other data domains.

2 Related Work

A large number of studies have been conducted to measure effectiveness of visualizations using different evaluation methods.

Controlled experiments: Many studies have evaluated visualizations through rigorous controlled experiments [4], [5]. In these studies, typical independent variables control aspects of the tools, tasks, data, and participant classes. Dependent variables include accuracy and efficiency measures. Accuracy measures include precision, error rates, number of correct and incorrect responses, whereas efficiency includes measures of time to complete predefined benchmark tasks. E.g., [18] compares three different visualization systems on different tasks in terms of solution time and accuracy.

Usability testing: Usability tests typically evaluate visualizations to identify and solve user interface problems. Methods involve observing participants as they perform designated tasks using a 'think aloud' protocol, noting the usability incidents that may suggest incorrect use of the interface, and comparing results against a predefined usability specification [14]. Refer to [24] for a professional example.

Metrics, Heuristics, and Models: Different from empirical evaluations are inspections of user interfaces by experts, such as with heuristics [21]. Examples of specific metrics for visualizations include expressiveness and effectiveness criteria [20], data density and data/ink [31], a variety of other criteria for representation and interaction [10], as well as high-level design principles [28]. Cognitive models, such as CAEVA [17], can be used to simulate visualization usage and thereby study the low-level effects of various visualization techniques.

Longitudinal and Field Studies: A longitudinal study of information visualization adoption by data analysts is presented in [13]. Their work suggests advantages when visualizations are used as complementary products rather than stand alone products. [25] examines users' long-term exploratory learning of new user interfaces, with 'eureka reports' to record learning events.

Thus, a range of evaluation methods have been used to measure effectiveness of visualizations [22]. In the literature, controlled experiments are the most prevalent for identifying and validating more effective visualizations. Unfortunately, these studies evaluate visualizations based only on a set of predefined tasks.

[1] email: psaraiya@vt.edu
[2] email: north@cs.vt.edu
[3] email: kduca@vbi.vt.edu

Insight-based studies

- Goal: compare tools
  - Identify insight occurrences
  - Measure overall learning
- Procedure
  - Tutorial
  - Initial questions
  - Examine data, estimating potential insight found
  - Assess experience
- Dependent variables
  - Questions
  - Time
  - Amount learned
  - Insights & characteristics
  - Techniques used
  - Usability issues
  - Demographics

[Saraiya et al., 2004]

Characteristics of insight

- Fact
- Time
- Domain value – rating
- Hypotheses
- Breadth vs. depth – coded by domain expert
- Directed vs. unexpected
- Correctness – coded by domain & vis expert
- Category – coded

- Dependent variables
  - Questions
  - Time
  - Amount learned
  - Insights & characteristics
  - Techniques used
  - Usability issues
  - Demographics

[Saraiya et al., 2004]

Limitations

- Labor intensive to capture & code
- Requires domain expert
- Requires motivated subjects
- Training and trial time low – Self reporting? Diary? Automated capture?

Empirical Studies in Information Visualization

Empirical Studies in Information Visualization: Seven Scenarios

Understanding data analysis by:

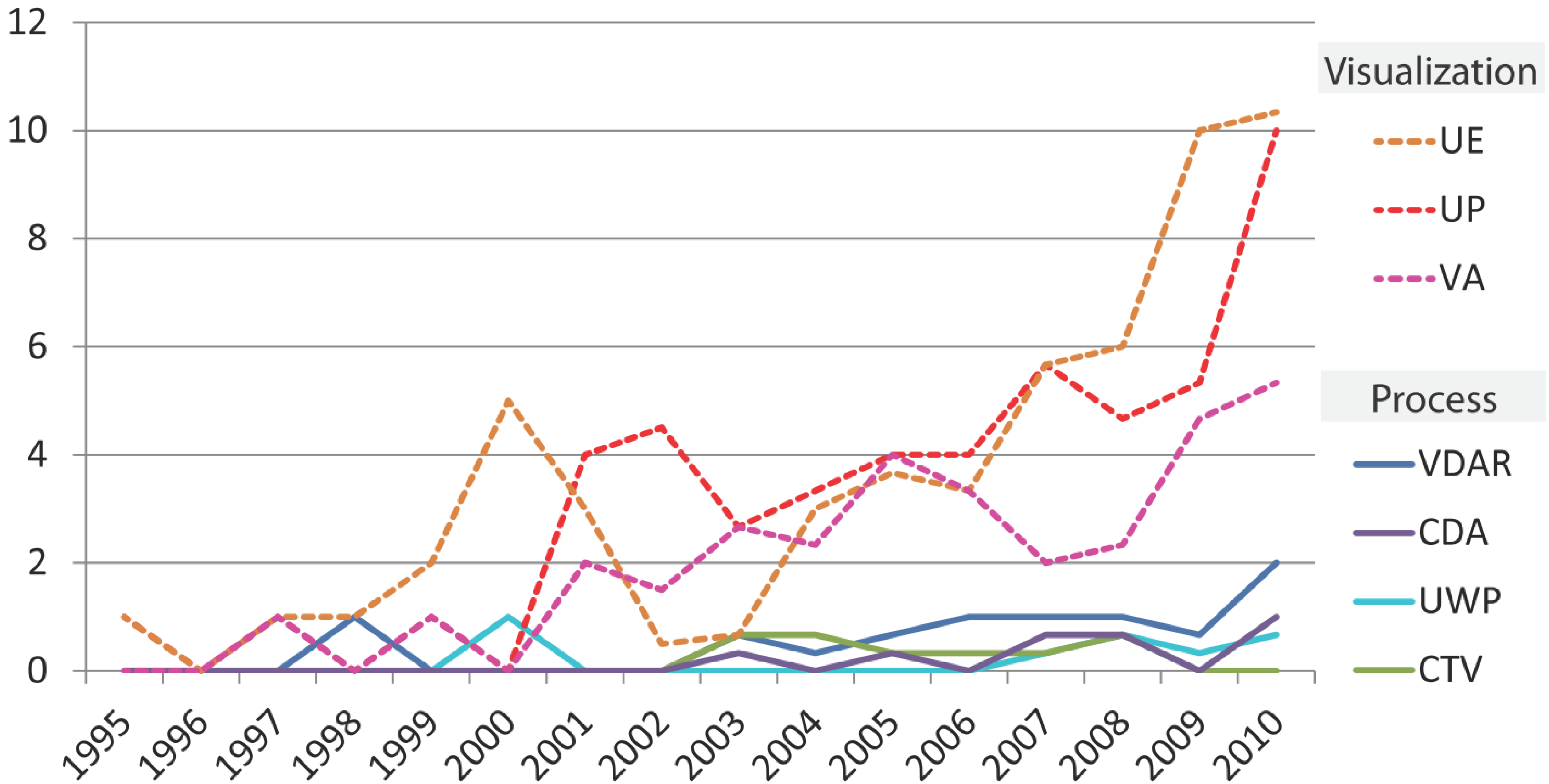
- understanding environments and work practices,
- evaluating visual data analysis and reasoning,
- evaluating communication through visualization, and
- evaluating collaborative data analysis.

Understanding visualizations by:

- evaluating user performance,
- evaluating user experience, and
- evaluating visualization algorithms.

Lam et al., 2012

Empirical Studies in Information Visualization: Seven Scenarios



[Lam et al., 2012]

Understanding environments and work practices

- Goals & outputs
  - Understand work, analysis, or information processing practices of people
  - Without software in use: inform design
  - With software in use: assess factors for adoption, how appropriated for future design
- Evaluation Questions
  - Context of use?
  - Integrate into which daily activities?
  - Supported analyses?
  - Characteristics of user group and environment?
  - What data & tasks?
  - What visualizations/tools used?
  - How current tools solve tasks?
  - Challenges and usage barrier?

Understanding environments and work practices

- Methods
  - Field Observation
    - Real world, free use of tool
    - Derive requirements
  - Interviews
    - Contextual inquiry: interview then observe in routines, with little interference
    - Pick the right person
    - Laboratory context w/domain expert
  - Laboratory Observation
    - How people interact with each other, tools
    - More control of situation

Evaluating visual data analysis and reasoning

- Goals & outputs
  - Assess visualization tool’s ability to support visual analysis and reasoning
    - As a whole! Not just a technique
  - Quantifiable metrics or subjective feedback
- Evaluation Questions: Does it support…
  - Data exploration?
  - Knowledge discovery?
  - Hypothesis generation?
  - Decision making?

Evaluating visual data analysis and reasoning

- Methods
  - Case studies
    - Motivated experts with own data in own environment
    - Can be longitudinal
    - Insight-Based (Saraiya et al., 2004)
      - Unguided, diary, debriefing meetings
    - MILCS: Multidimensional In-depth Long-term Case studies (Shneiderman & Plaisant, 2006)
      - Guided, observations, interviews, surveys, automated logging
      - Assess interface efficacy, user performance, interface utility
      - Improve system during
  - Lab observations and interviews
    - Code results
    - Think aloud
  - Controlled Experiment
    - Isolate important factors

Evaluating communication through visualization

- Goals & outputs
  - How effectively is a message delivered and acquired
- Evaluation Questions
  - Quantitative: learning rate, information retention and accuracy
  - Qualitative: interaction patterns
- Methods
  - Controlled experiments
  - Field observation & interviews

Evaluating Collaborative Data Analysis

- Goals & outputs
  - Evaluate support for taskwork and teamwork
  - Holistic understanding of group work processes or tool use
  - Derive design implications
- Evaluation Questions
  - Effective and efficient?
  - Satisfactorily support or stimulate group sensemaking?
  - Support group insight?
  - Is social exchange and communication facilitated?
  - How is the tool used? Features, patterns…
  - What is the process? User requirements?

Evaluating Collaborative Data Analysis

- Methods
  - Context critical, but early formative studies less dependant
  - Heuristic evaluation
    - Heuristics: actions, mechanics, interactions, locales needed
  - Log analysis
    - Distributed or web-based tools
    - Combine with questionnaire or interview
    - Hard to evaluate unlogged & qualitative aspects
  - Field or laboratory observation
    - Involve group interactions and harmony/disharmony
    - Combine with insight-based?

Evaluating User Performance

- Goals & outputs
  - Measure specific features
  - Time, accuracy, and error; work quality (if quantifiable); memorability
  - Descriptive statistics results
- Evaluation Questions
  - What are the limits of human perception and cognition?
  - How do techniques compare?
- Methods
  - Controlled experiment -> design guideline, model, head-to-head
    - Few variables
    - Simple tasks
    - Individual differences matter
  - Field logs
    - Suggest improvements, recommendation systems

Evaluating User Experience

- Goals & outputs
  - Inform design: uncover gaps in functionality, limitations, directions for improvement
  - Subjective: user responses
    - Effectiveness, efficiency, correctness, satisfaction, trust, features liked/disliked
  - Objective: body sensors, eye tracking
- Evaluation Questions
  - Features: useful, missing, to rework?
  - Are there limitations that hinder adoption?
  - Is the tool understandable/learnable?

Evaluating User Experience

- Methods
  - Informal evaluation
    - Demo for domain experts (usually) and collect feedback
  - Usability test
    - Watch (video) how participants perform set of tasks to perfect design
    - Take note of behaviors, remarks, problems
    - Carefully prepare tasks, interview script, questionnaires
  - Field observation
    - Understand interaction in real setting
  - Laboratory questionnaire
    - Likert scale
    - Open ended


FIND...

Evaluating Visualization Algorithms

- Goals & outputs
  - Quantitatively judge generated output quality (metrics) & performance
  - How scores vs. alternatives
  - Explore limits & behavior
- Evaluation Questions
  - Which shows interesting patterns best?
  - Which is more truthful?
  - Which is less cluttered?
  - Faster, less memory, less money?
  - How does it scale?
  - Extreme cases?

Evaluating Visualization Algorithms

- Methods
  - Visualization quality assessment
    - Readability metrics, image quality measures
  - Algorithmic performance
    - Varied data, size, complexity, corner cases
    - Benchmark data sets