CS 571 - Data Visualization & Exploration

Tasks, Design, and Evaluation

Instructor: Hamza Elhamdadi

UMass Amherst

Upcoming Dates

Apr 25: Homework 4 (Due at 11:59pm Eastern)

(Tentatively) Apr 20: Project Peer Feedback Groups will be assigned (Due Apr 25)

Apr 28: Final Quiz

Apr 29: Final Group Activity

Tasks Analysis

A Nested Model for Visualization Design and Validation

Tamara Munzner, *Member, IEEE*

Abstract—We present a nested model for the visualization design and validation with four layers: characterize the task and data in the vocabulary of the problem domain, abstract into operations and data types, design visual encoding and interaction techniques, and create algorithms to execute techniques efficiently. The output from a level above is input to the level below, bringing attention to the design challenge that an upstream error inevitably cascades to all downstream levels. This model provides prescriptive guidance for determining appropriate evaluation approaches by identifying threats to validity unique to each level. We also provide three recommendations motivated by this model: authors should distinguish between these levels when claiming contributions at more than one of them, authors should explicitly state upstream assumptions at levels above the focus of a paper, and visualization venues should accept more papers on domain characterization.

Index Terms—Models, frameworks, design, evaluation.

1 Introduction

Many visualization models have been proposed to guide the creation and analysis of visualization systems [8, 7, 10], but they have not been tightly coupled to the question of how to evaluate these systems. Similarly, there has been significant previous work on evaluating visualization [9, 33, 42]. However, most of it is structured as an enumeration of methods with focus on *how* to carry them out, without prescriptive advice for *when* to choose between them.

The impetus for this work was dissatisfaction with a flat list of evaluation methodologies in a recent paper on the process of writing visualization papers [29]. Although that previous work provides some guidance for when to use which methods, it does not provide a full framework to guide the decision or analysis process.

In this paper, we present a model that splits visualization design into levels, with distinct evaluation methodologies suggested at each level based on the threats to validity that occur at that level. The four levels are: characterize the tasks and data in the vocabulary of the problem domain, abstract into operations and data types, design visual encoding and interaction techniques, and create algorithms to execute these techniques efficiently. We conjecture that many past visualization designers did carry out these steps, albeit implicitly or subconsciously, and not necessarily in that order. Our goal in making these steps more systems, and compare our model to previous ones. We provide recommendations motivated by this model, and conclude with a discussion of limitations and future work.

2 Nested Model

Figure 1 shows the nested four-level model for visualization design and evaluation. The top level is to characterize the problems and data of a particular domain, the next level is to map those into abstract operations and data types, the third level is to design the visual encoding and interaction to support those operations, and the innermost fourth level is to create an algorithm to carry out that design automatically and efficiently. The three inner levels are all instances of design problems, although it is a different problem at each level.

These levels are nested; the output from an *upstream* level above is input to the *downstream* level below, as indicated by the arrows in Figure 1. The challenge of this nesting is that an upstream error inevitably cascades to all downstream levels. If a poor choice was made in the abstraction stage, then even perfect visual encoding and algorithm design will not create a visualization system that solves the intended problem.

Purpose of the Nested Model

Capture Design Decisions

- what is the justification behind your design?

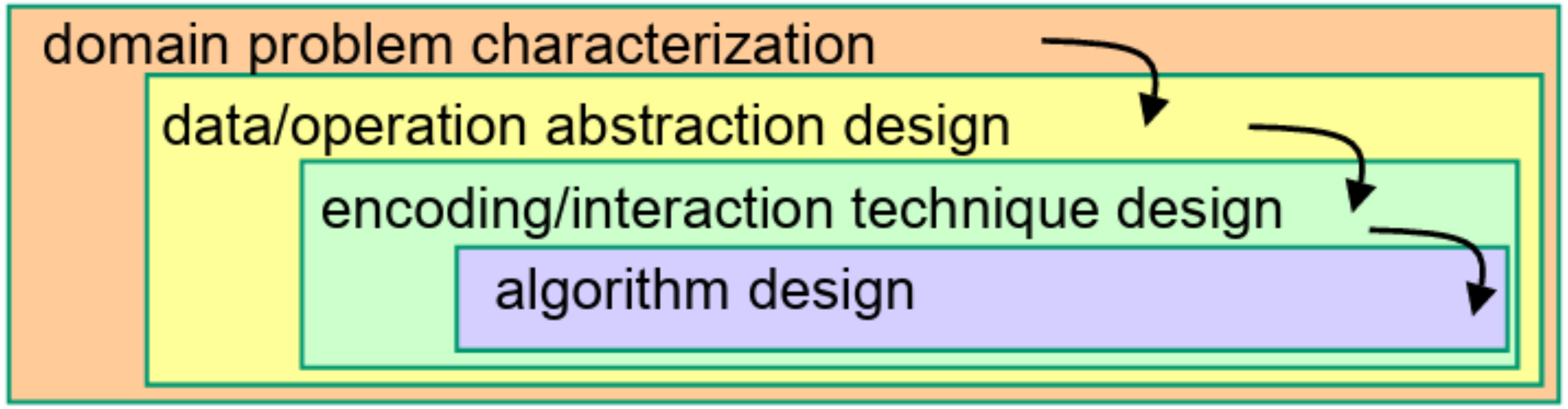
Analyze Modular Aspects of the Design Process

- four different concerns

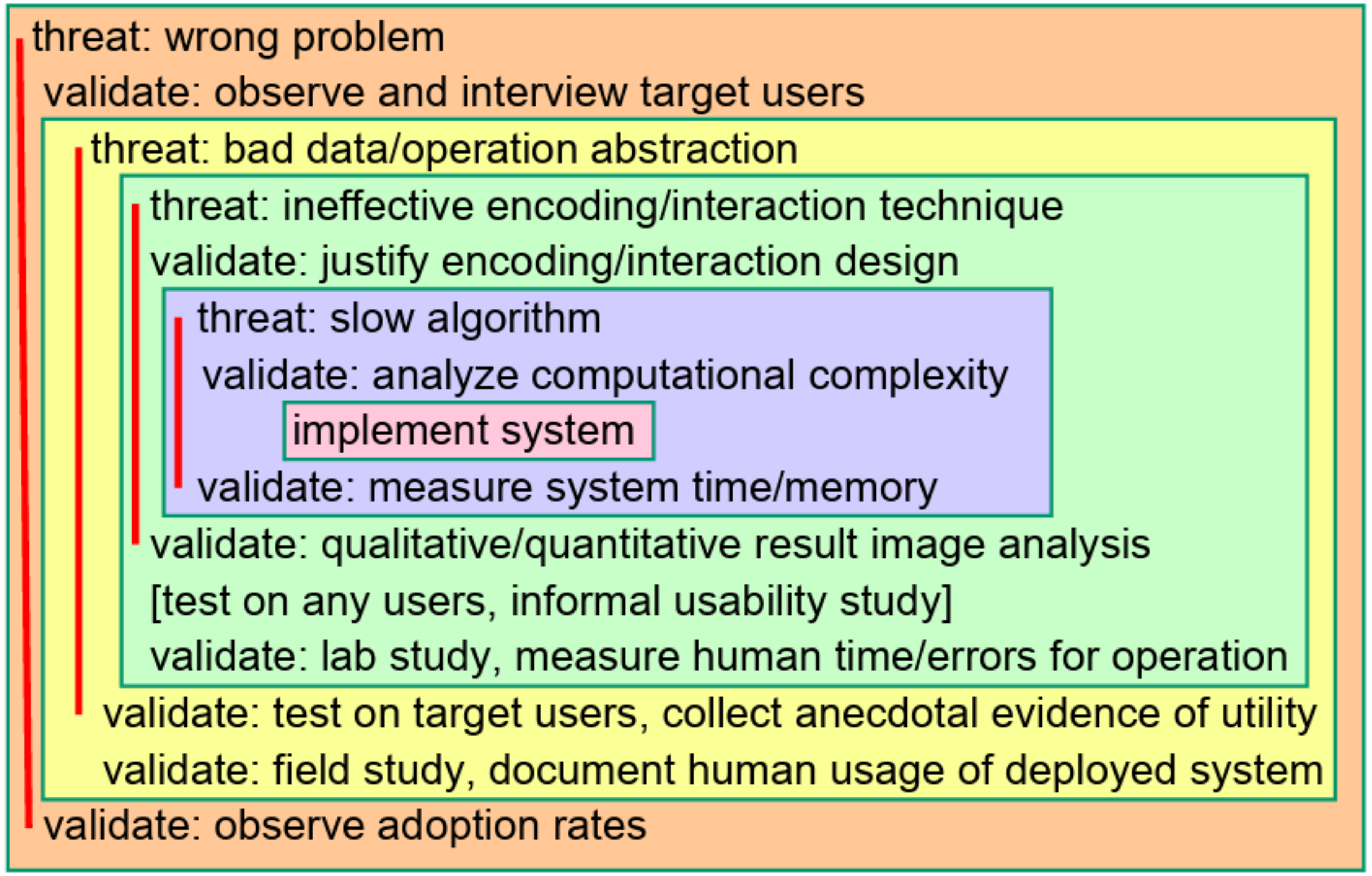
Validate/Evaluate Early and Often

- avoid making ineffective solutions

Nested Model for Visualization Design

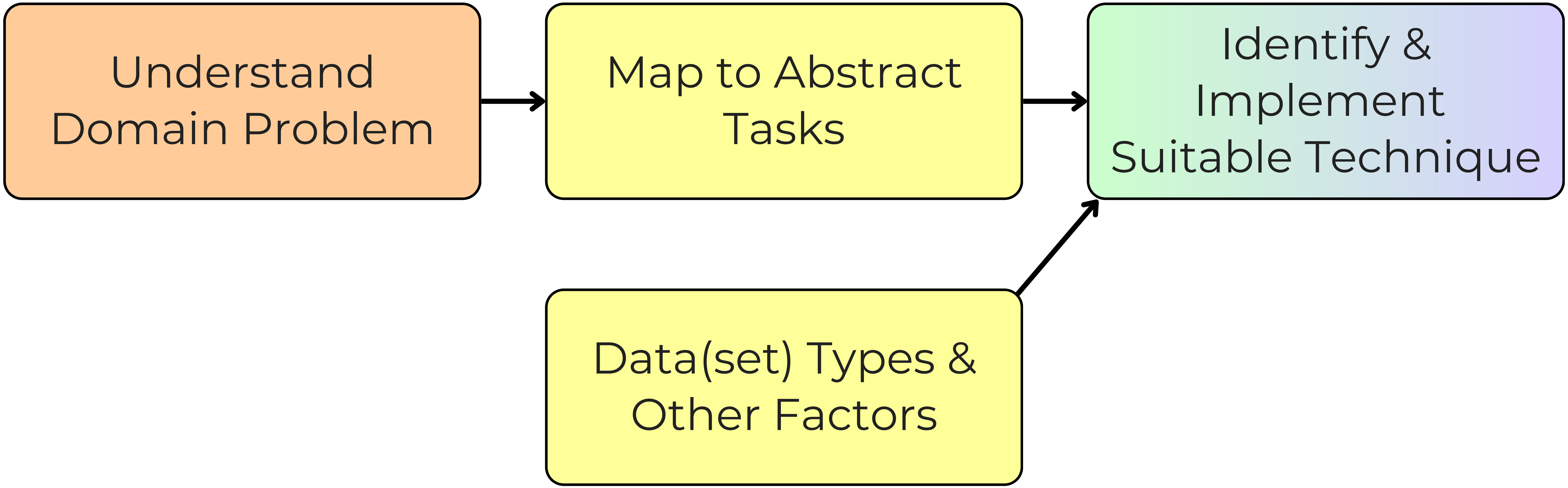


Design



Threats & Evaluation

Design Process



Domain Characterization

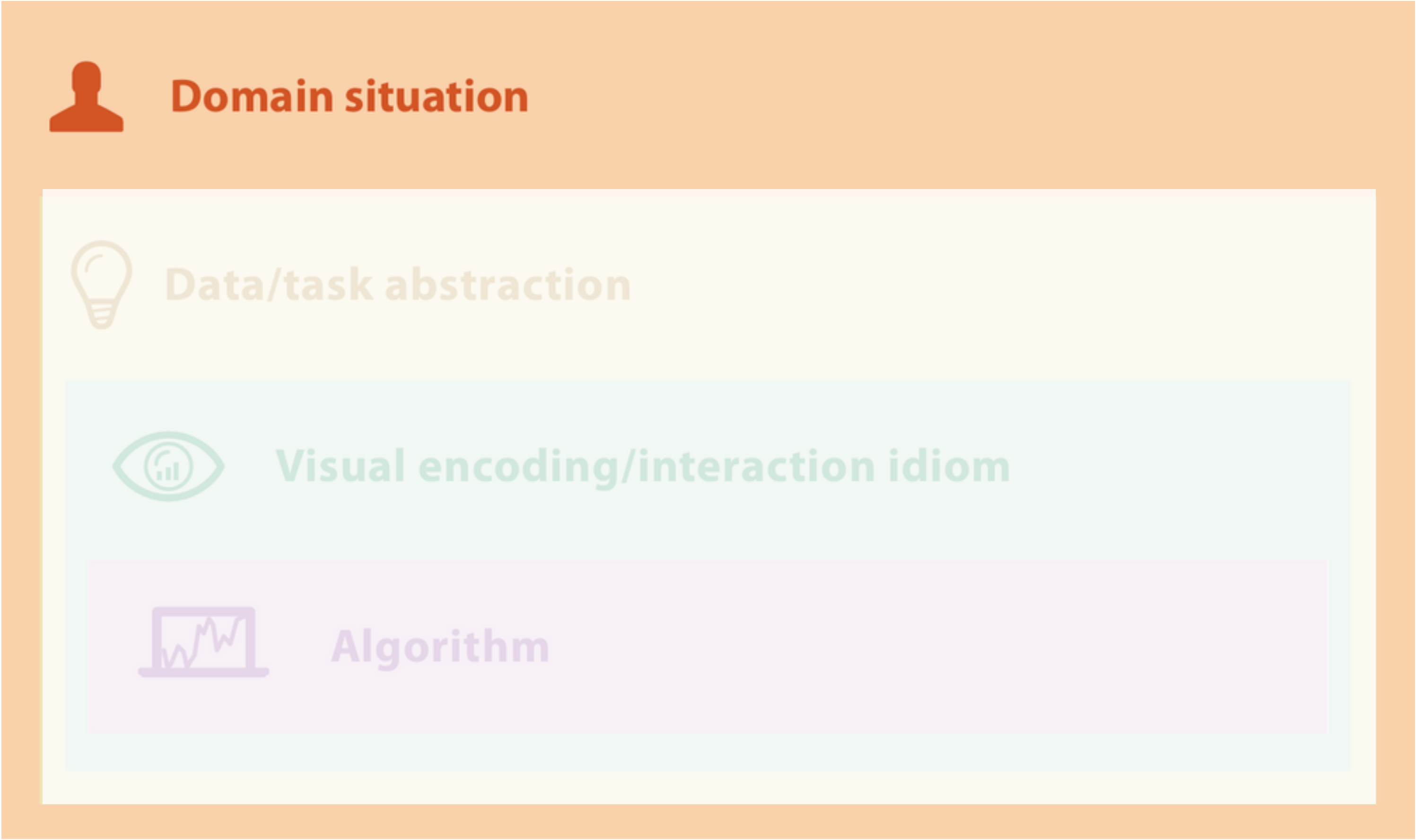
Details of an application domain

Who are your **users**? What are their **questions**?

- varies wildly by domain
- **must be specific** enough to continue to the next step

Cannot just ask people what they do

- **introspection is hard!**



Domain Problem Characterization

There are infinitely many specific domain tasks

So, we break them down into **smaller abstract tasks**

- we know how to solve these abstract tasks

If we can identify the {**task, data**} combination:

- we can probably find an existing solution

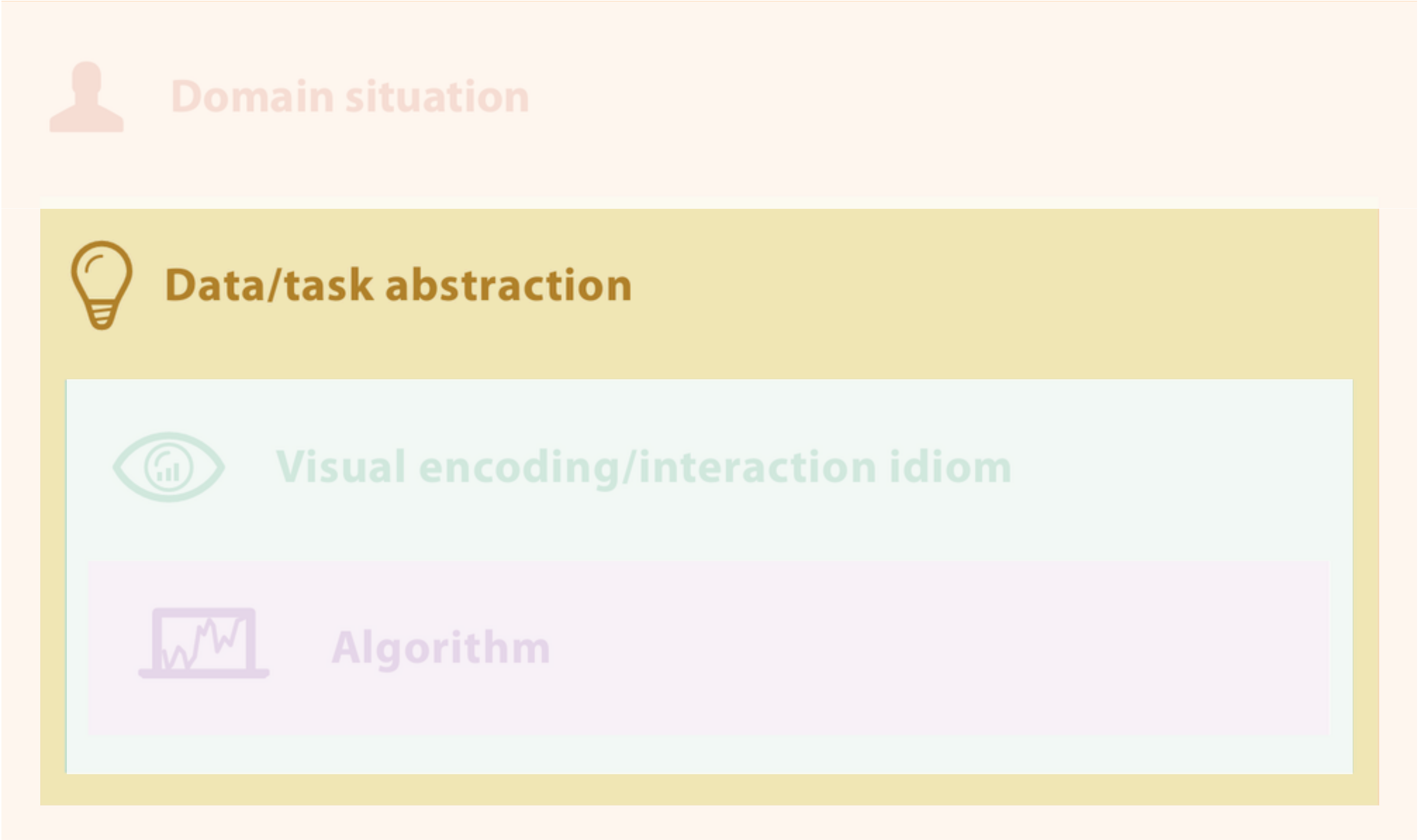
Example: Find Good Movies

"I want to identify good movies in genres that I like!"

Who are your users?

- general audiences
- movie enthusiasts

Data & Task Abstraction



Generalized terms (what & why)

Identify the **tasks** that users wish to perform (or already do)

Determine **data types** and develop a good model of the data

Sometimes, you will need to transform the data for a better vis solution

- how you do this varies depending on the specific tasks

Example: Find Good Movies

What is a good movie?

- highly rated by movie critics?
- highly rated by audiences?
- Successful at the box office?
- Similar to other movies I like?
- Specific Genres?

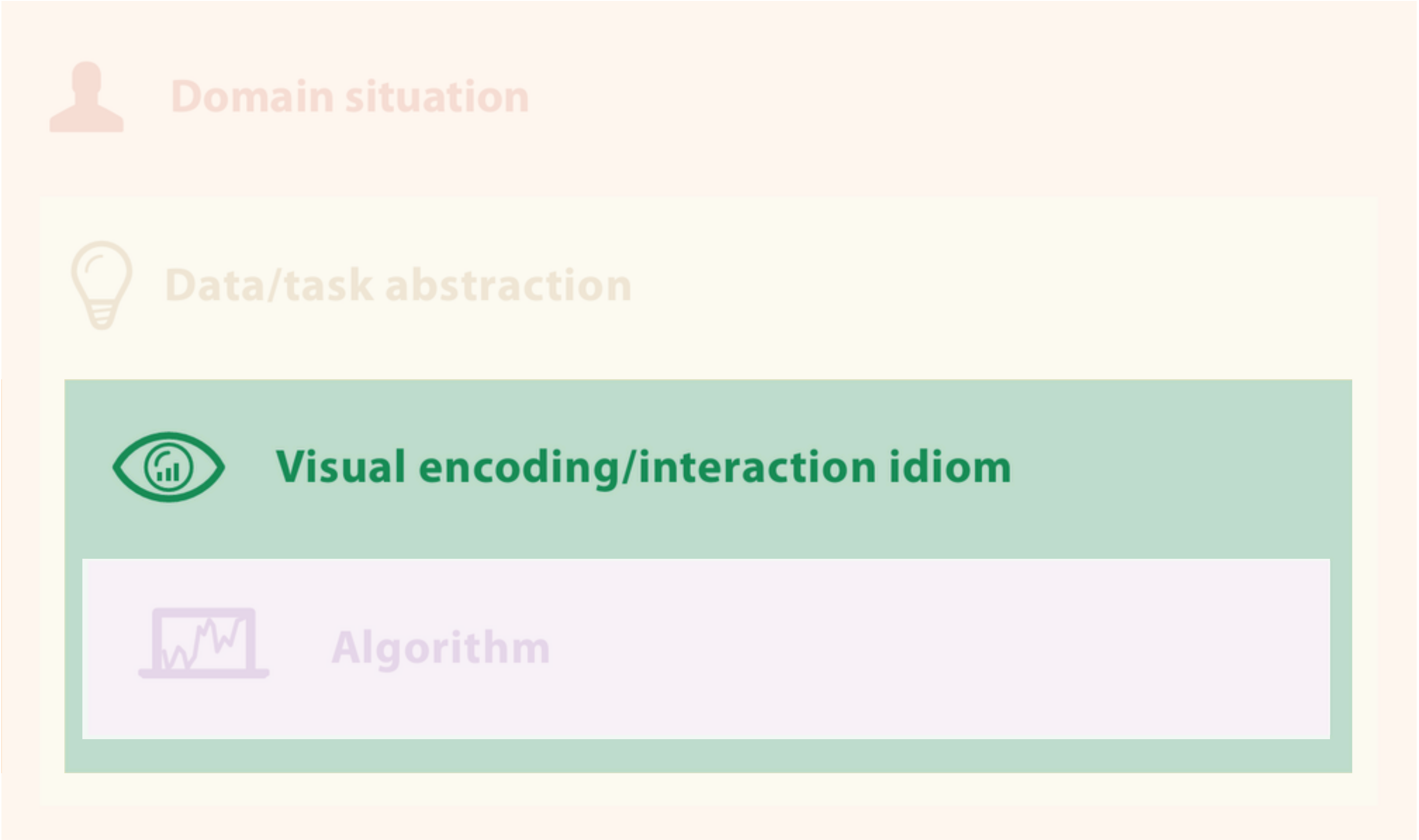
Data Sources: IMDb, Rotten Tomatoes, …

Task:

- **Identify** horror movies that are highly rated by audiences on Rotten Tomatoes
- **Compare** audience and critic ratings

Encodings & Interactions

Methods for **creating** and **manipulating** the **visual representation of data**



These decisions about **encodings** and **interactions**:

- **are driven by design principles** (e.g., perception, gestalt principles, expressiveness, effectiveness, etc)
- **may be separate or intertwined**

Example: Find Good Movies

Task:

- **Identify** horror movies that are highly rated by audiences on Rotten Tomatoes
- **Compare** audience and critic ratings

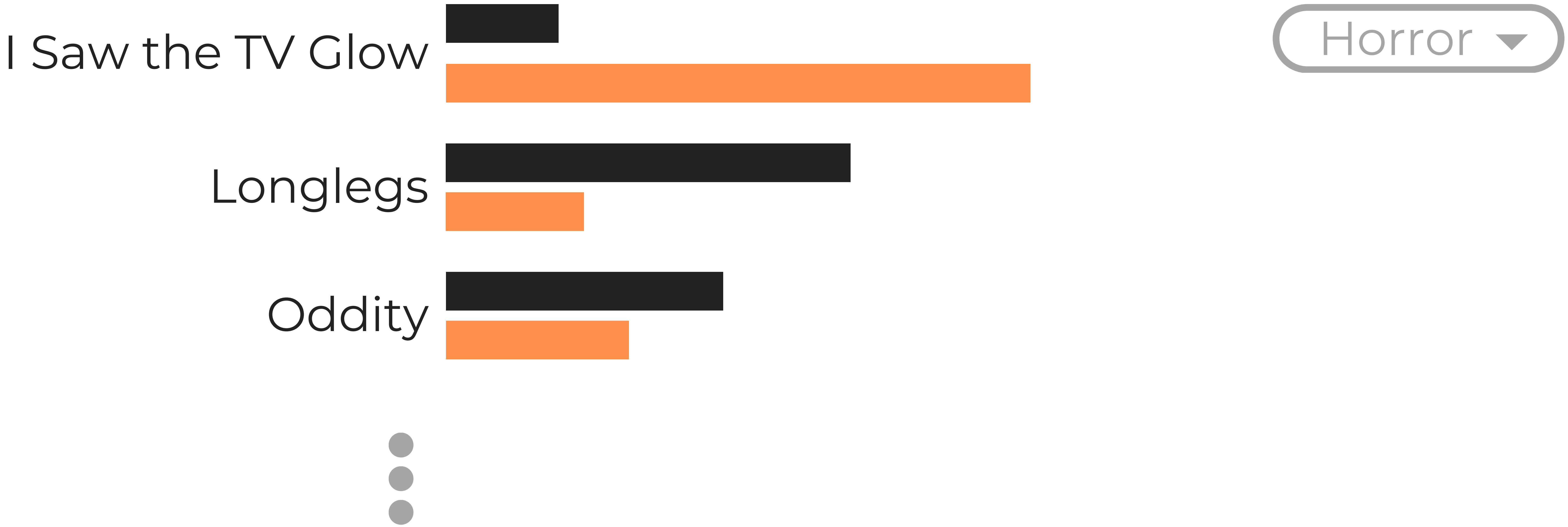
Idiom: stacked bar chart for ratings w/ filter interface for genre

Example: Find Good Movies

Idiom: stacked bar chart for ratings w/ filter interface for genre

Example: Find Good Movies

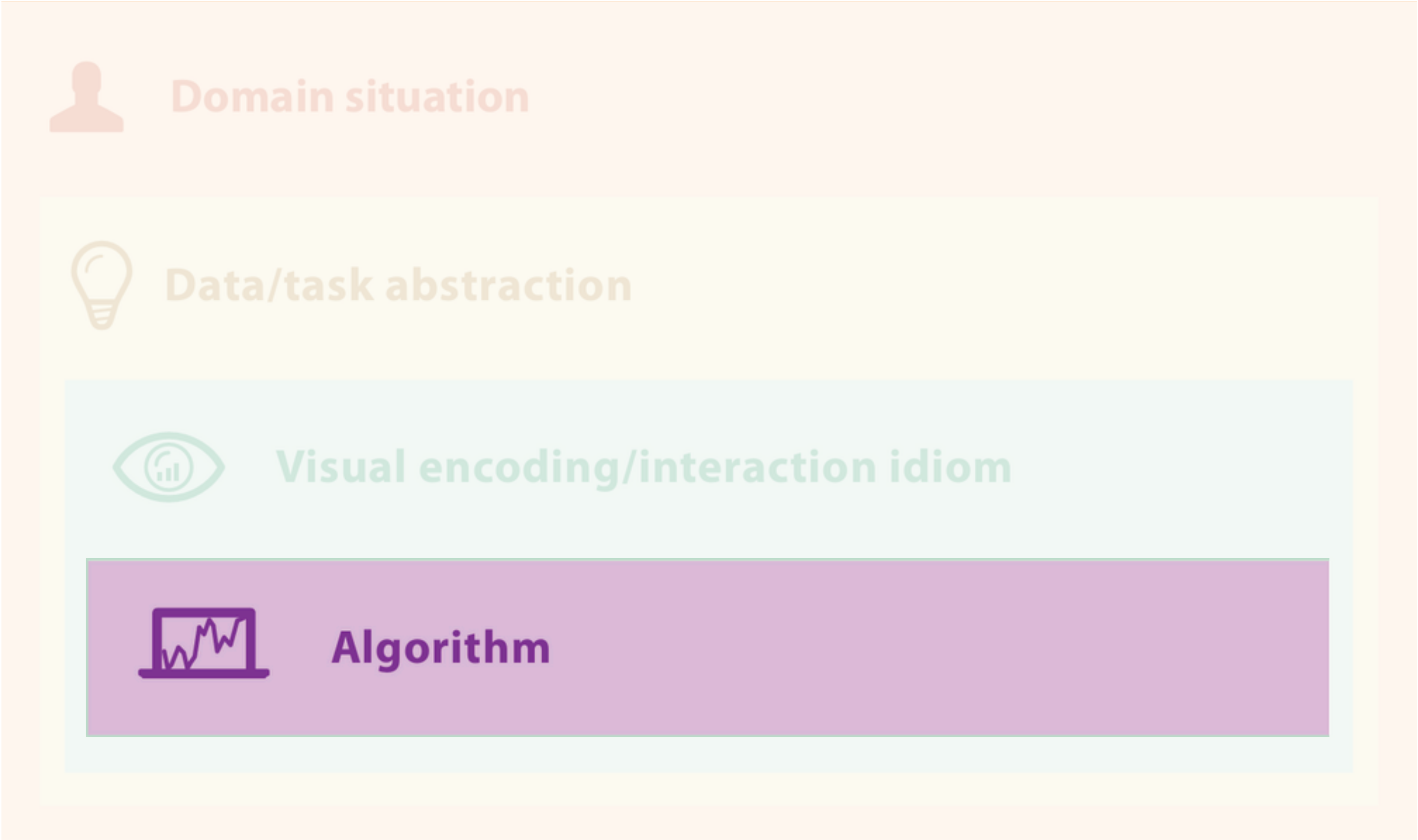
Better Idiom: dual bar chart w/ filter interface for genre



Algorithms

Details about how you will implement the visualization

D3, python, Tableau
interactive? static?



Designing Visualizations

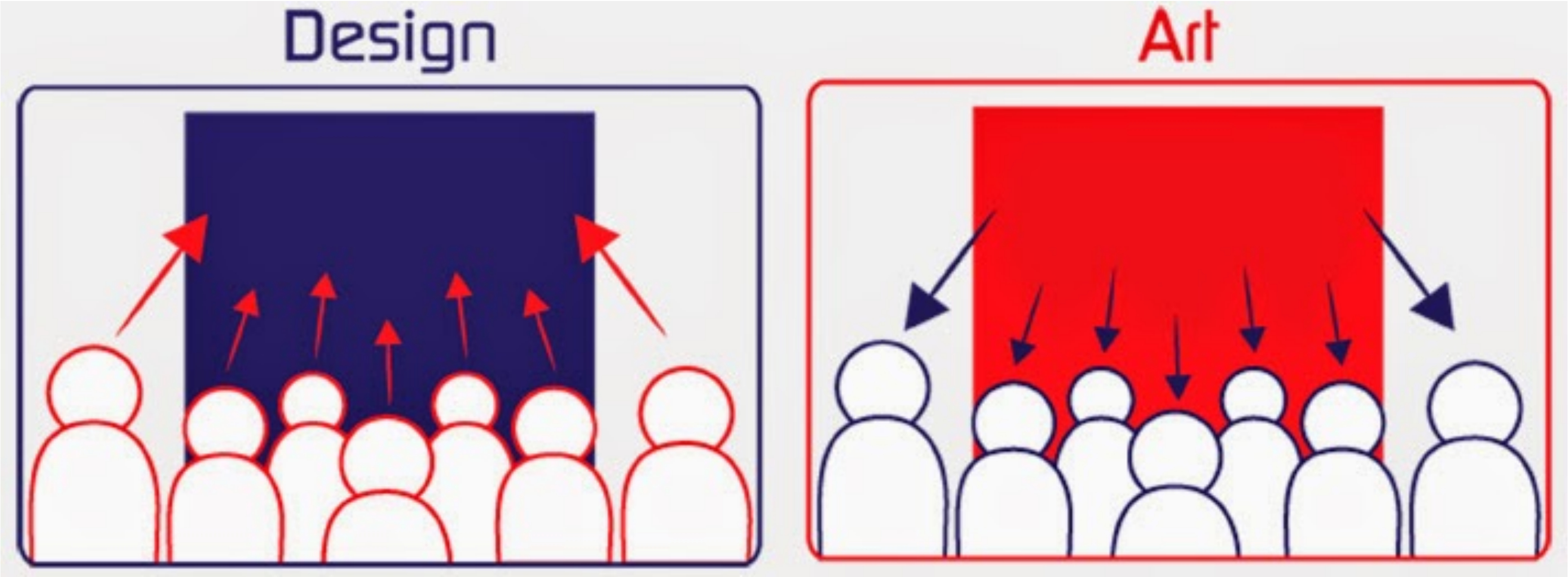
Design Is:

- creating something new **to solve a problem**
- building things with a **function** (e.g., chairs, user interfaces, user interfaces, etc.)
- used across many different fields
- for many possible users / tasks

https://www.youtube.com/watch?v=hUhisi2FBuw

Design Isn’t:

- “just making things pretty”
- without clear purpose
- building without justification or evidence



Form & Function

"form follows function"

- function constrains possible forms
- form depends on the tasks that a design aims to achieve

"the better defined the goals of an artifact, the narrower the variety of forms it can adopt" - Alberto Cairo

Function Follows Form: 18 Sculptural Home Furnishings

Turn your home into a gallery with functional furnishings that double as sculpture, like cabinets in the form of human torsos and beds shaped like roller coasters. These 18 designs blur the lines between furniture and a...

WebUrbanist / May 27, 2015

https://weburbanist.com/2015/05/27/form-follows-function-18-sculptural-home-furnishings/

Why Does Design Matter for Vis?

There are **many ineffective visualization combinations**

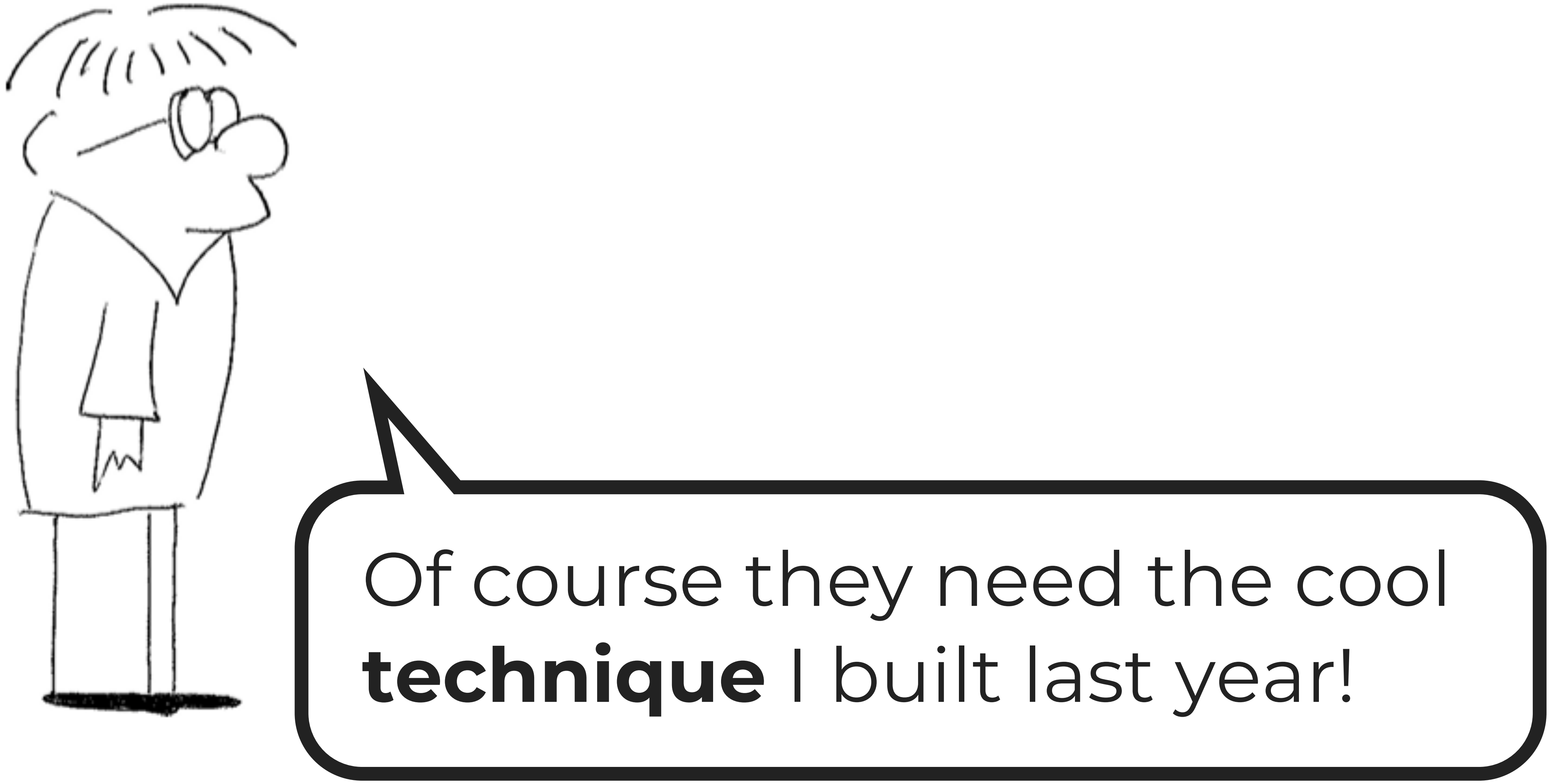
Different users have **different problems** & data

The possible **design space is large**

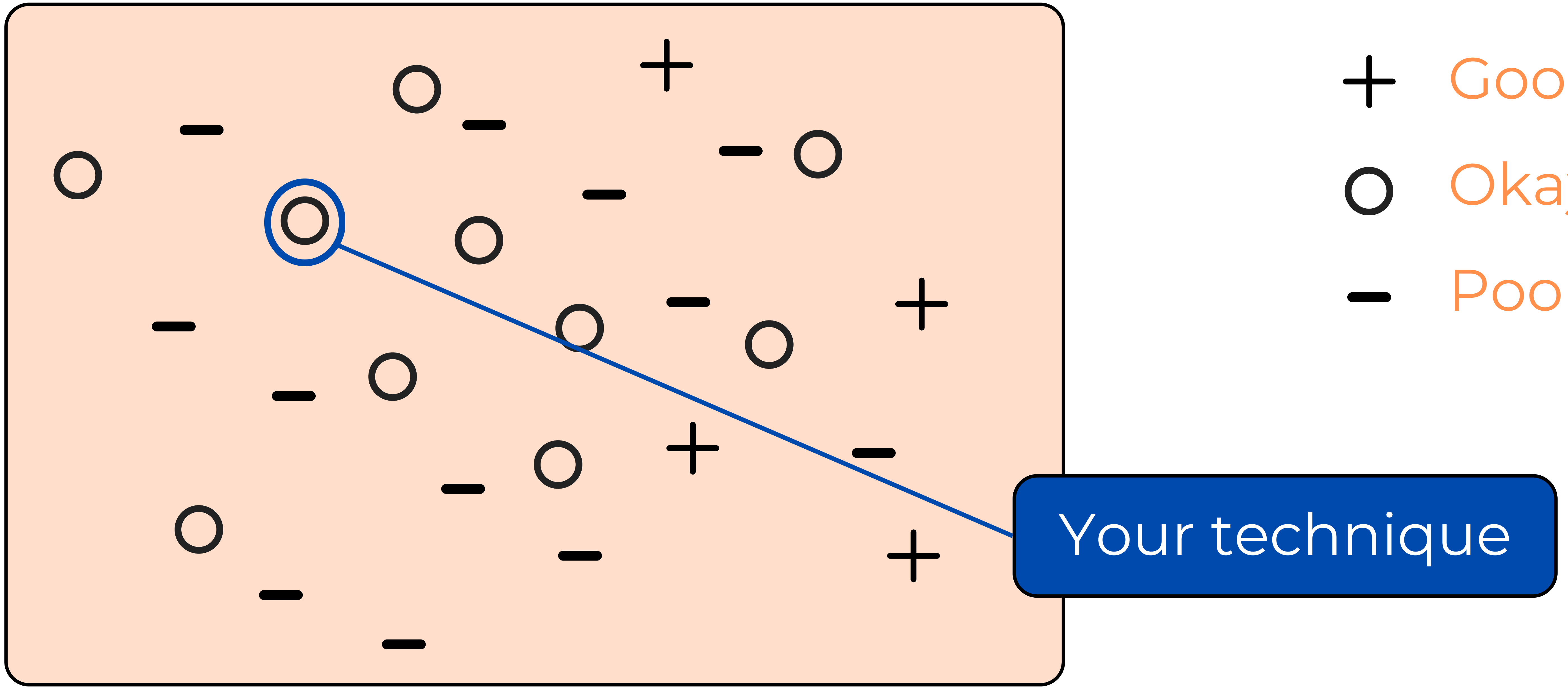
- Making thoughtful design choices is important

A Common Pitfall

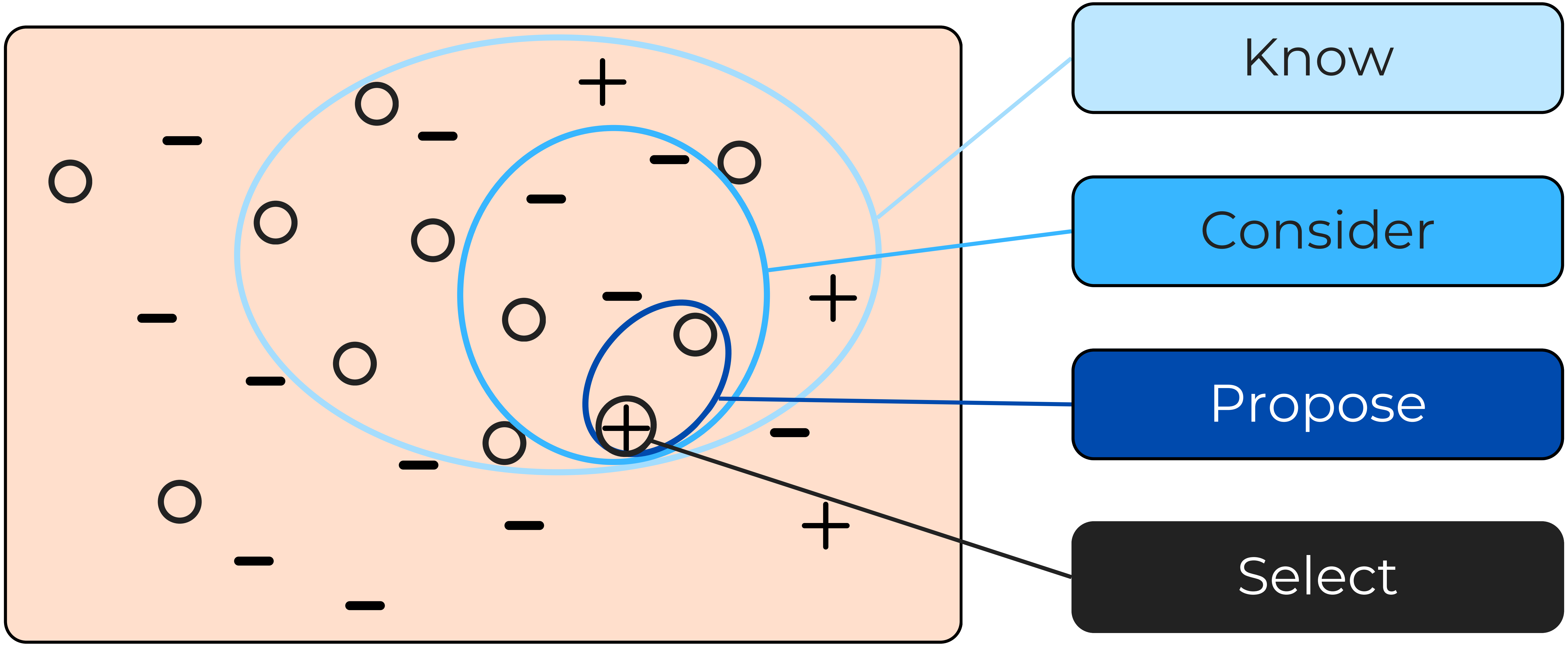
Premature Design Commitment



A Metaphor: Design Space



A Metaphor: Design Space



A Metaphor: Design Space

Design Methods

Creativity Workshops

https://ieeexplore.ieee.org/document/8440830

A Framework for Creative Visualization-Opportunities Workshops

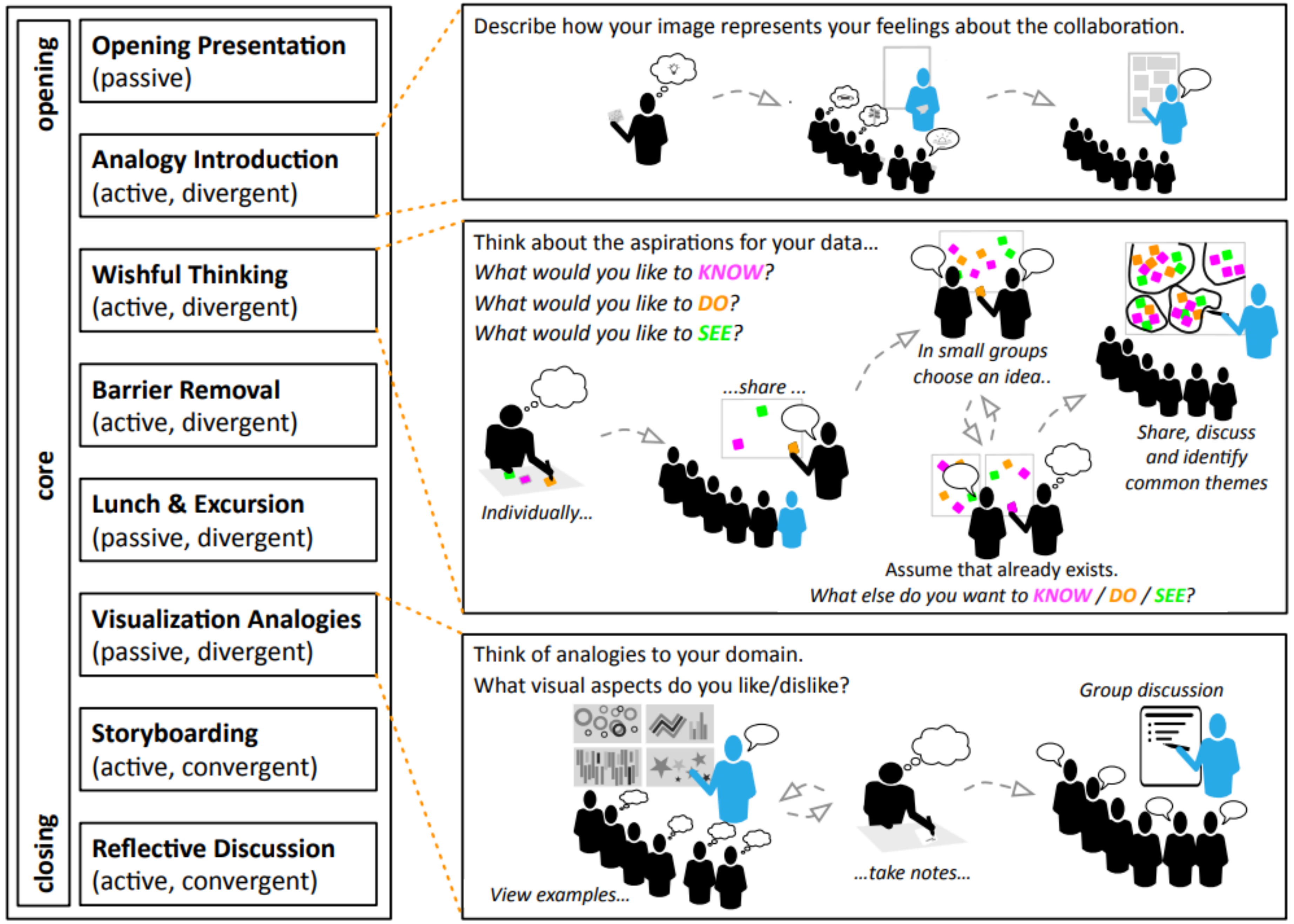
Ethan Kerzner, Sarah Goodwin, Jason Dykes, Sara Jones, Miriah Meyer

Abstract—Applied visualization researchers often work closely with domain collaborators to explore new and useful applications of visualization. The early stages of collaborations are typically time consuming for all stakeholders as researchers piece together an understanding of domain challenges from disparate discussions and meetings. A number of recent projects, however, report on the use of creative visualization-opportunities (CVO) workshops to accelerate the early stages of applied work, eliciting a wealth of requirements in a few days of focused work. Yet, there is no established guidance for how to use such workshops effectively. In this paper, we present the results of a 2-year collaboration in which we analyzed the use of 17 workshops in 10 visualization contexts. Its primary contribution is a framework for CVO workshops that: 1) identifies a process model for using workshops; 2) describes a structure of what happens within effective workshops; 3) recommends 25 actionable guidelines for future workshops; and 4) presents an example workshop and workshop methods. The creation of this framework exemplifies the use of critical reflection to learn about visualization in practice from diverse studies and experience.

Index Terms—User-centered visualization design, design studies, creativity workshops, critically reflective practice

1 Introduction

Two key challenges in the early stages of applied visualization research are to find pressing domain problems and to translate them into interesting visualization opportunities. Researchers often discover details, but with a focus on their experience using a series of workshops in a collaboration with energy analysts. In contrast, Kerzner et al. [35] summarize their workshop with neuroscientists in one sen-



Creativity Workshops

Goals:

- generate design requirements
- promote creativity

Techniques:

- wishful thinking
- constraint removal
- analogical reasoning
- storyboarding

Prototypes are measured for **appropriateness**, **novelty**, and **surprise**

Creativity Workshops

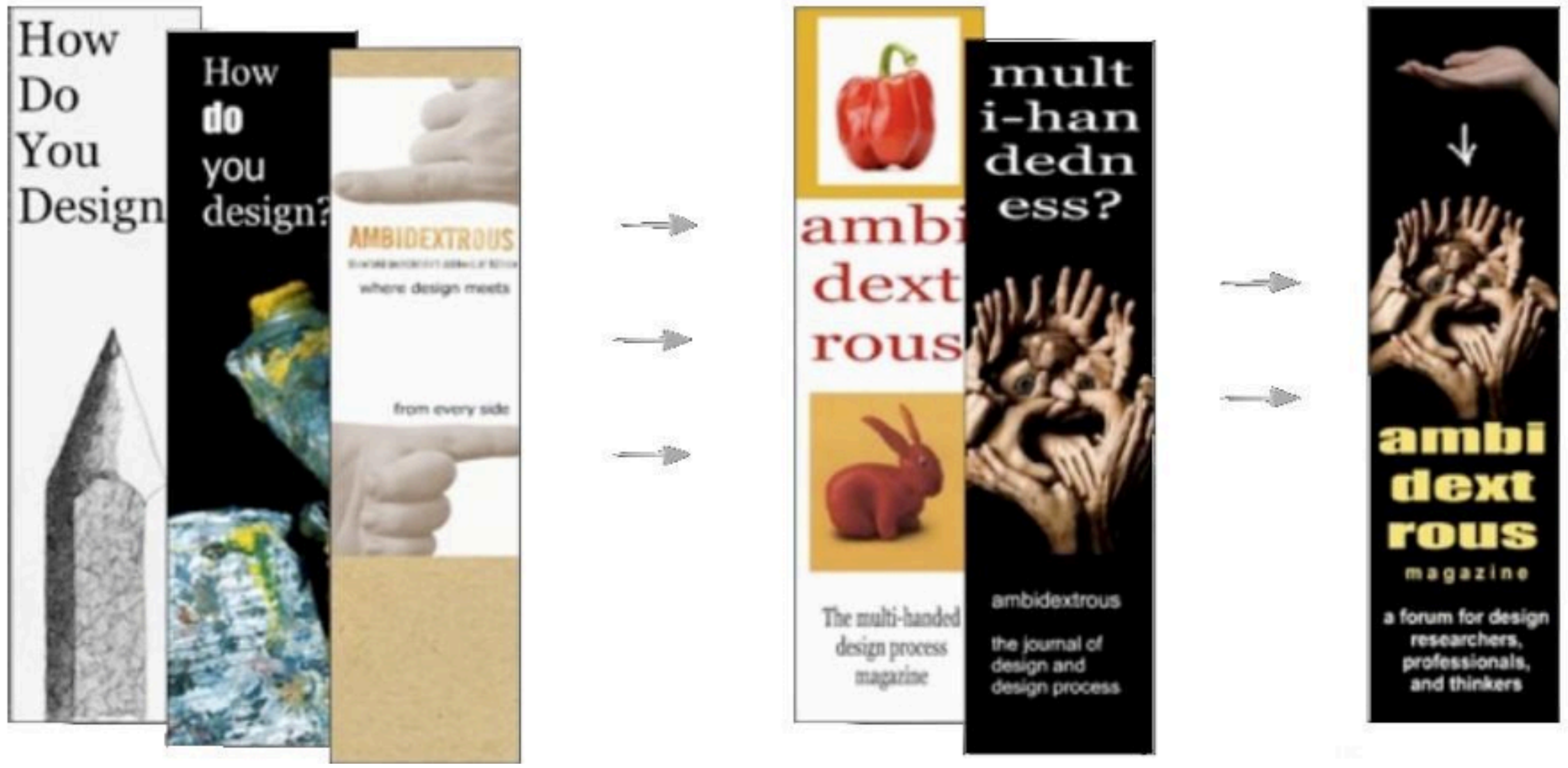
Parallel Prototyping

Develop **multiple prototypes** of your visualization in parallel

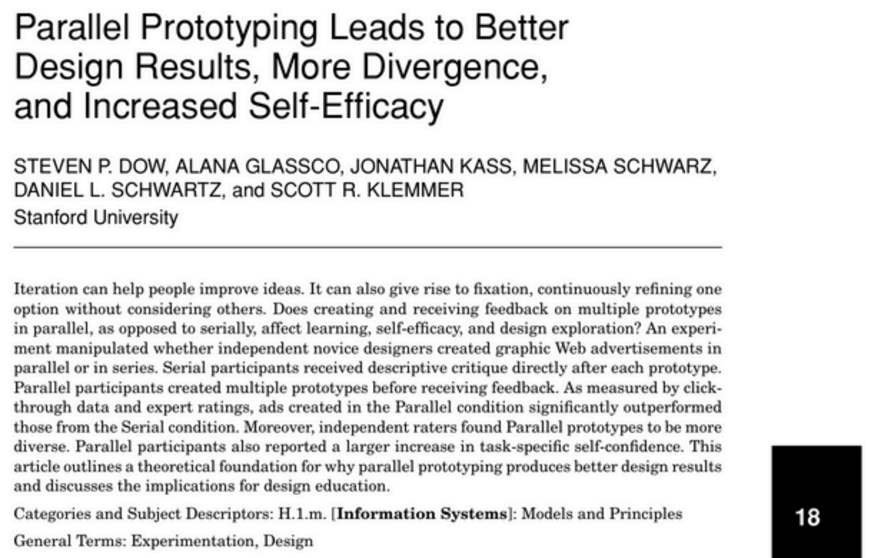
Case Study: Graphic Design

Serial Design

Parallel Design



https://dl.acm.org/doi/10.1145/1879831.1879836



Parallel Prototyping Leads to Better Design Results, More Divergence, and Increased Self-Efficacy

STEVEN P. DOW, ALANA GLASSCO, JONATHAN KASS, MELISSA SCHWARZ, DANIEL L. SCHWARTZ, and SCOTT R. KLEMMER
Stanford University

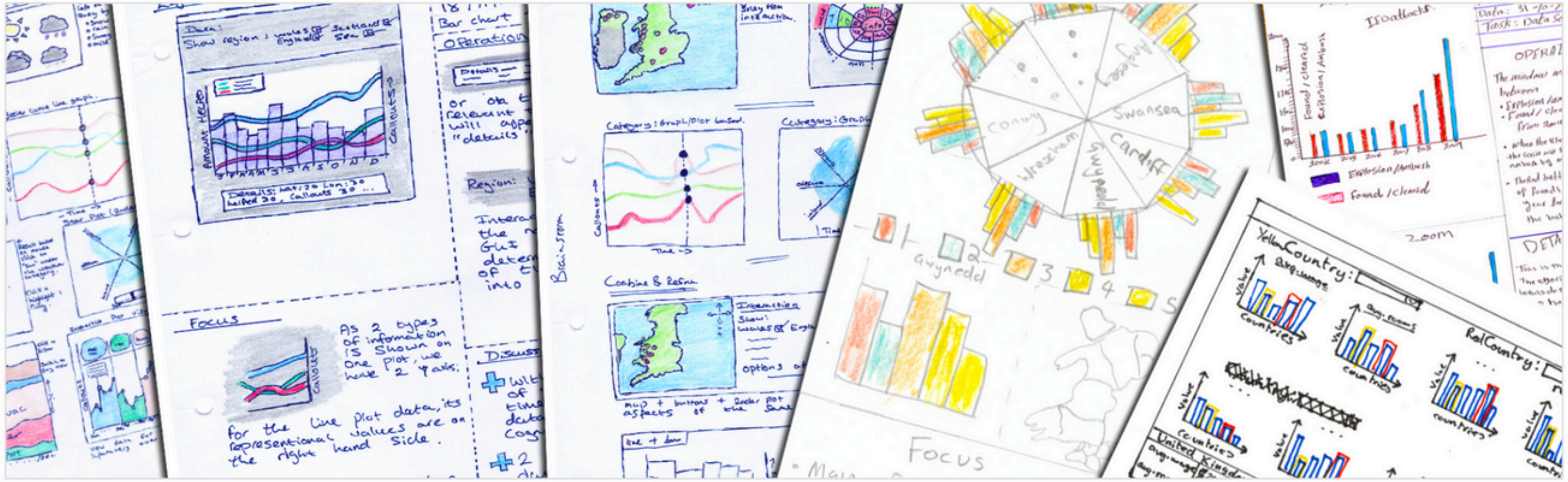
Iteration can help people improve ideas. It can also give rise to fixation, continuously refining one option without considering others. Does creating and receiving feedback on multiple prototypes in parallel, as opposed to serially, affect learning, self-efficacy, and design exploration? An experiment manipulated whether independent novice designers created graphic Web advertisements in parallel or in series. Serial participants received descriptive critique directly after each prototype. Parallel participants created multiple prototypes before receiving feedback. As measured by click-through data and expert ratings, ads created in the Parallel condition significantly outperformed those from the Serial condition. Moreover, independent raters found Parallel prototypes to be more diverse. Parallel participants also reported a larger increase in task-specific self-confidence. This article outlines a theoretical foundation for why parallel prototyping produces better design results and discusses the implications for design education.

Categories and Subject Descriptors: H.1.m. [**Information Systems**]: Models and Principles

General Terms: Experimentation, Design

18

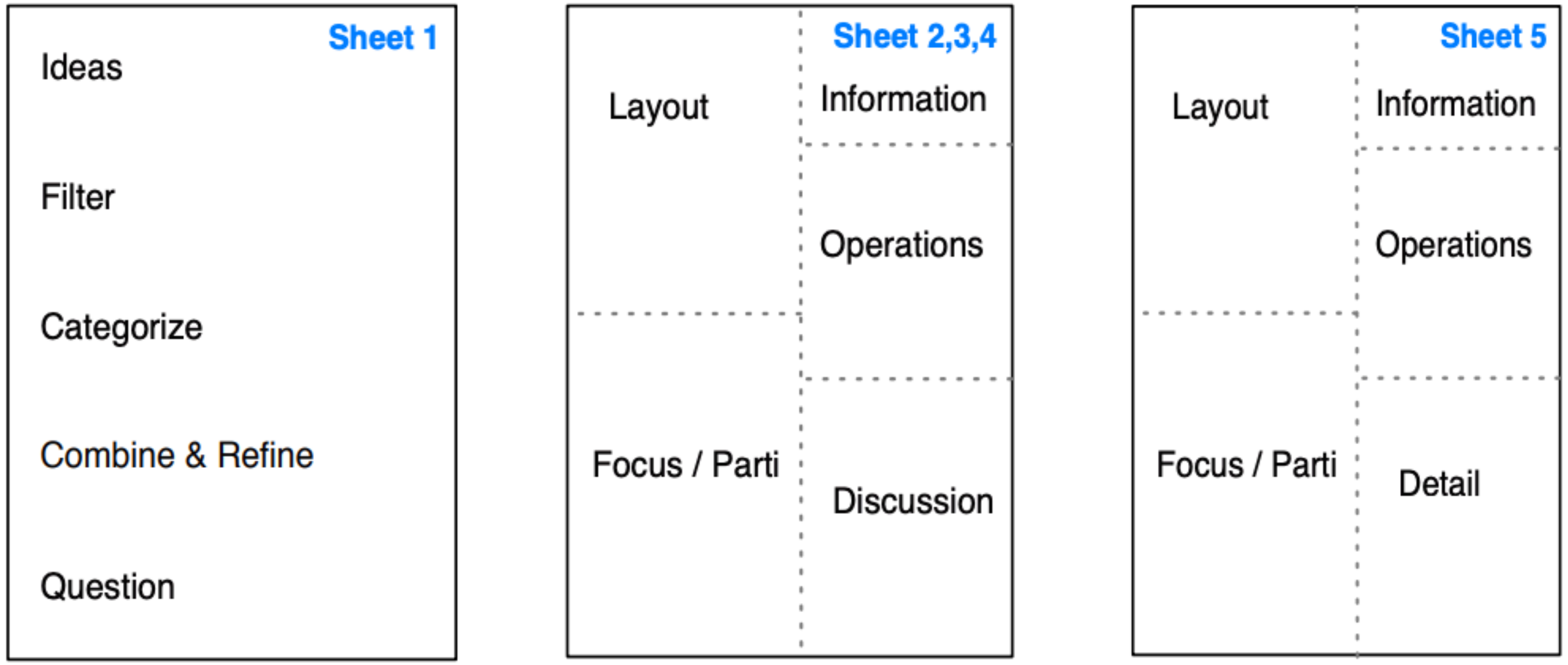
Five Design Sheets



Tailored to visualization design

Promotes **divergent** and **convergent** design

Prototyping on paper stimulates creativity and interaction



Evaluating Visualizations

Evaluating Information Visualizations

Sheelagh Carpendale

Department of Computer Science, University of Calgary,
2500 University Dr. NW, Calgary, AB, Canada T2N 1N4
sheelagh@ucalgary.ca

1 Introduction

Information visualization research is becoming more established, and as a result, it is becoming increasingly important that research in this field is validated. With the general increase in information visualization research there has also been an increase, albeit disproportionately small, in the amount of empirical work directly focused on information visualization. The purpose of this paper is to increase awareness of empirical research in general, of its relationship to information visualization in particular; to emphasize its importance; and to encourage thoughtful application of a greater variety of evaluative research methodologies in information visualization.

One reason that it may be important to discuss the evaluation of information visualization, in general, is that it has been suggested that current evaluations are not convincing enough to encourage widespread adoption of information visualization tools [57]. Reasons given include that information visualizations are often evaluated using small datasets, with university student participants, and using simple tasks. To encourage interest by potential adopters, information visualizations need to be tested with real users, real tasks, and also with large and complex datasets. For instance, it is

Why Evaluate / Validate?

- **Avoid** ineffective visualizations
- **Justify** visualization design choices
- **Understand** visualization perception
  - e.g., is size a better visual channel than angle?
- **Compare** new visualizations to previous vis
  - Is my vis one of a kind? or can I compare it to others to draw inspiration?
- Usability **test**ing
  - check for problems with the vis we've designed

What Evaluation Methods Are There?

- **Controlled Experiments**
  - Laboratory, crowd-sourced
- **Interviews / questionnaires**
  - structured, unstructured, semi-structured
- Field **observation**, lab observation
  - video / audio analysis
  - think aloud
  - coding/classification of user behavior (speech, gestures)
- **Case Studies**

More certain about what actually causes differences

BUT less ecological validity

More ecologically valid (realistic)

BUT less certainty about what actually causes differences

Quantitative vs. Qualitative Evaluation

Quantitative Methods:

- objective metrics / **measurements**
- **stats** and **numbers** used to interpret results

Qualitative Methods:

- subjective metrics
- **detailed descriptions** of situations, events, people, interactions, and observed behaviors
- use **direct quotes** from people about their experiences, attitudes, beliefs, and thoughts
- understand **how people make meaning** of their experiences of their environment and the world

Common Metrics

Objective metrics:

- task completion time
- physiological data
- error when performing tasks (e.g., numeric scores)
- ratio of successes to failures
- ...

Subjective metrics:

- ratings
- rankings
- comments and feedback
- ease of use
- ...

Internal vs. External Validity

Internal Validity (can you trust your experiment?)

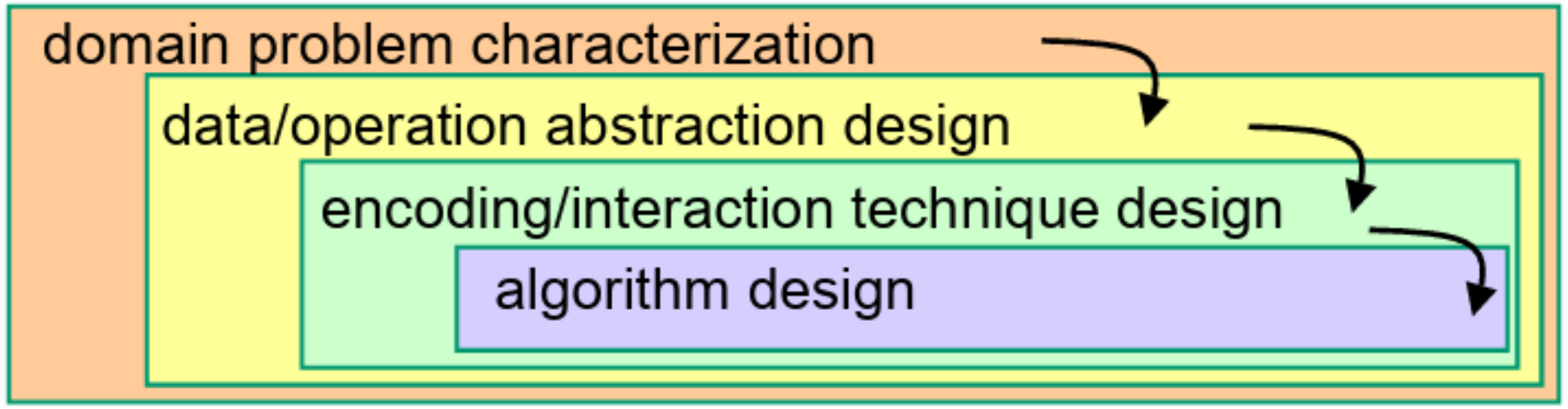
- high when tested under controlled lab conditions
- observed effects are due to the test conditions (not random variables)

External Validity (is experiment representative of the real world?)

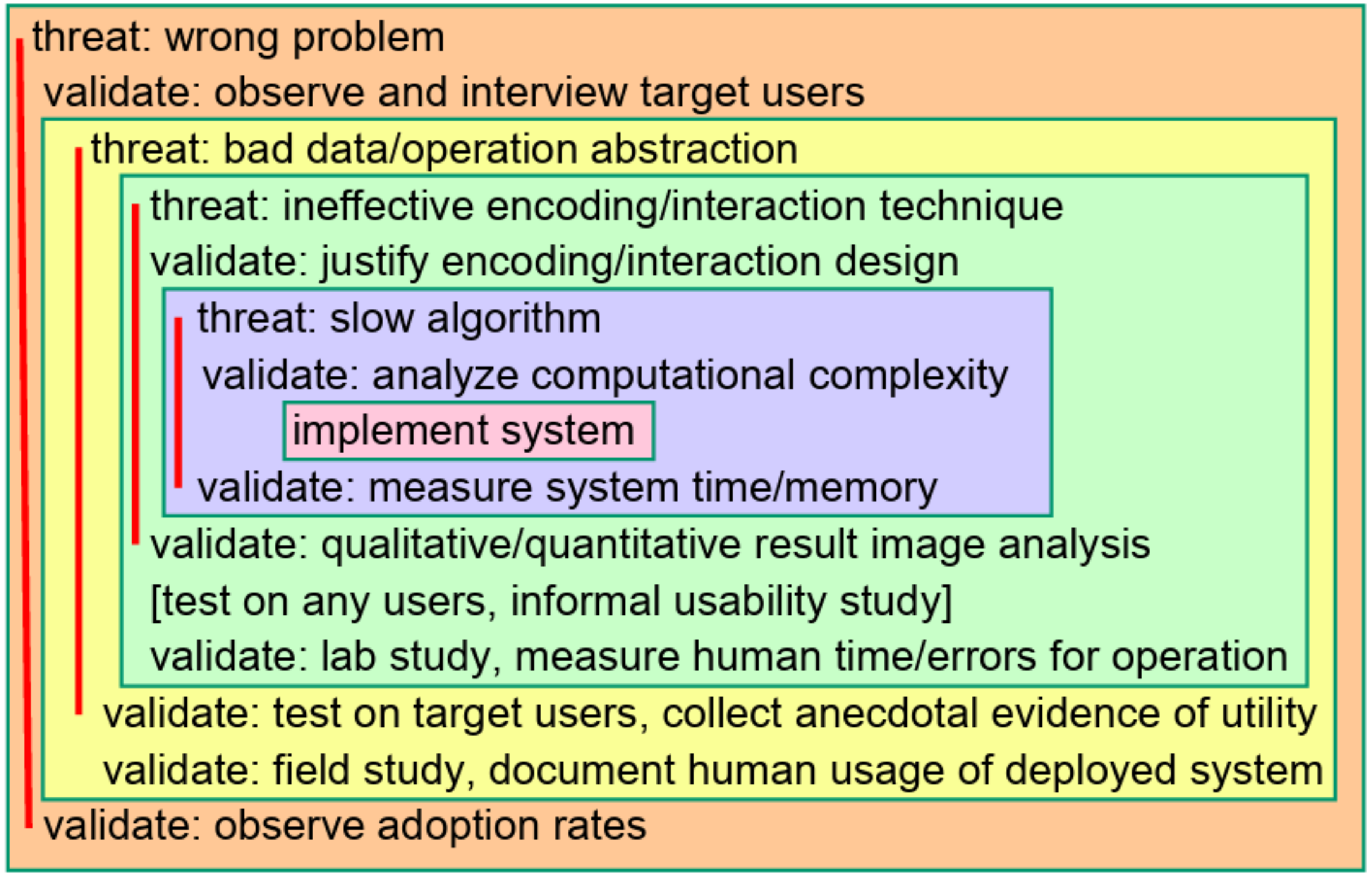
- high when tested in the field
  - e.g., collecting real users responses to visualizations on a news website
- results are valid in the real world

Trade-off: The more akin to real-world situations, the more the experiment is susceptible to uncontrolled sources of variation

Scope of Evaluation



Design



Threats & Evaluation

Scope of Evaluation

Pre-design

- e.g., interview your user base and understand their workflows

Design

- e.g., evaluate how your visual encodings and interactions are perceived

Prototype

- e.g., compare your vis prototype to other state-of-the-art systems

Deployment

- e.g., assess your vis effectiveness in the field

Re-design

- e.g., improve your design by identifying usability problems

FIN

Upcoming Dates

Apr 25: Homework 4 (Due at 11:59pm Eastern)

(Tentatively) Apr 20: Project Peer Feedback Groups will be assigned (Due Apr 25)

Apr 28: Final Quiz

Apr 29: Final Group Activity