# Exploring Important and Difficult Concepts to Learn in Information Visualization

Camilo Vieira*, Paul Parsons, Vetria Byrd

Purdue University

**ABSTRACT**

Information visualization (InfoVis) has acquired significant relevance across many disciplines, leading to the emergence of new courses and educational programs. As a consequence, visualization educators have started to explore pedagogical practices to support the student learning process. We present preliminary results of the first phase of data collection of a Delphi study aimed at identifying important and difficult concepts in information visualization education. We also identify some benefits of this work for InfoVis education, and describe future plans for the Delphi study.

**Keywords**: information visualization, important concepts, difficult concepts, pedagogy.

**Index Terms**: [Human-centered computing]: Visualization theory, concepts and paradigms; [Human-centered computing]: Empirical studies in visualization; [Social and professional topics]: Information technology education

## 1 INTRODUCTION

The vast amounts of data being produced and consumed over the past few decades have led to a growing interest in information visualization (InfoVis). Numerous job roles now require or prefer InfoVis knowledge and skills, particularly those related to data science and design. As a result, there is an increasing interest in the introduction of formal educational programs in InfoVis. InfoVis educators and practitioners have also recently developed interest in the exploration and documentation of InfoVis pedagogies. For instance, the 2016 IEEE VIS conference offered the first workshop in Pedagogy of Data Visualization. Also, various university departments have begun to offer majors or concentrations in InfoVis (e.g., [1, 2]), while some universities have offered specific courses in information visualization for almost 20 years now.

This paper contributes to the growing body of knowledge surrounding InfoVis education by exploring concepts and topics that are both important and difficult to learn. Identifying these topics and concepts will enable educators and researchers to inform the design of educational programs, as well as the design and implementation of pedagogies to support student learning.

The first of a three-phase Delphi study was conducted to identify these two elements. Educators, researchers, and practitioners in InfoVis were invited to list, describe, and justify the important and the difficult concepts to learn in this area. The research questions for this study are:

What are *important* concepts/topics for students to learn in InfoVis?

What are *difficult* concepts/topics for students to learn in InfoVis?

* cvieira@purdue.edu

This paper is organized as follows: section two describes the participants, data collection instruments, and data analysis procedures; section three presents the common important and difficult concepts described by the participants; and section four summarizes the findings and discusses next steps.

## 2 METHODS

The Delphi technique was employed to identify important and difficult concepts to learn in InfoVis. This technique has also been previously used for this purpose in introductory computing courses [3] and in engineering education courses [4]. The Delphi method aims at achieving consensus about a topic that cannot be reached in a traditional way among a group of experts [5]. To achieve this consensus, multiple rounds of data collection need to be conducted. Here we present the preliminary results for the first round of this process.

### 2.1 Participants and Data Collection

An online survey was sent to different groups of researchers, practitioners, and educators in InfoVis that included the following mailing lists: Women in Vis, Broadening Participation in Visualization (BPViz), AAPHDCS: African American PhDs in Computer Science, CHI-Announcements, Infovis-digest, and diagrams-list. Posts were also made on various social media outlets. Eight participants completed the survey. Six of the participants had a doctorate degree while two of them had a master's degree. Their current positions were distributed as follows: three faculty members, four research scientists, and one consultant. Regarding specific activities that they carry out as part of their jobs in InfoVis, five participants do research, while five participants teach InfoVis courses. Four of the participants do industry consultancy, three participants design and develop systems, while two participants do curriculum development.
The survey asked two questions in the form:

Please list and elaborate on what you consider to be the five most IMPORTANT concepts to learn in InfoVis.

Please list and elaborate on what you consider to be the five most DIFFICULT concepts to learn in InfoVis.

Besides listing the name of the concepts, each entry included a description, and a justification for why this was an important or a difficult concept to learn.

### 2.2 Data Analysis

In order to consolidate the important and the difficult concepts to learn, one researcher first assigned codes to each of the responses. This researcher created definitions for these codes based on the descriptions and justifications that respondents included into the survey. These codes represented concepts that were described by the participants.

Two other researchers went through the responses grouped by codes and validated this assignment. Whenever discrepancies among the codes were found, the three researchers negotiated them until reaching agreement.

## 3 RESULTS

This section describes the preliminary findings of the first round of data collection of this Delphi study. Although the survey asked

specifically about concepts, some of the participants listed broader topics instead. In order to summarize the important and difficult concepts and topics, we include only the codes that showed more than one instance.

### 3.1 Important Topics

Nine concepts or topics were identified as important from the participants' responses.

**Human Perception**: This concept was highlighted by seven out of the eight participants. This refers to how people see and interpret visualizations. One of the participants justified the relevance of this concept as:

> *"Without a sound foundation in perceptual psychology it is not possible to develop visualizations of high quality. Humans in general, and designers specifically, lack the insight into their own perceptual processes, therefore, the design of visualizations often does not support the users."*

**Interaction**: six participants identified interaction as an important concept to learn in InfoVis. These included existing interaction techniques (e.g. selecting, filtering, reconfiguring, etc.) as well as how to implement them. One of the participants explained that:

> *"Users are not passive recipients of visualizations, but active partners. As such they need to be able control what they're looking at to be able to ask multiple different questions from the same data, evolve their questions as they learn more, compare current observations to previous observations*"

**Data complexity**: Five participants mentioned that it is important for students to understand the implication of working with large, complex, temporal, heterogeneous data:

> *"Although our representations of data are most often static, the data we represent is constantly changing in response to yet other data, or due to alterations in the environments and contexts in which the data exists in as they change over time."*

**Computational and Statistical methods**: understanding analysis techniques that go from descriptive statistics to Bayesian analysis was an important concept described by five participants: *"Much InfoVis is concerned with communicating statistical concepts, and it isn't possible to communicate something you don't understand"*

**Design creativity**: This concept was also identified by five participants. One of them described it as: *"The process of coming up with a visualization / layout, generating ideas, choosing good ideas"*.

**Considering the Context**: Four participants highlighted the importance of learning to consider the context of the data, the audience, and the task. Their justification was aligned as:

> *"Like any sociotechnical artifact, infovis software exists within larger cultural contexts. It is important to understanding how design and modeling in tools affects what kind of analysis can be done, what biases might be introduced, and what problems are ignored"*.

**Design considerations**: These considerations were highlighted by three participants and include: *"how different display projections work, how raw data is transformed and positioned within the projection, how styling is applied to data, how annotations, labels, are overlaid over the data, and interactive elements attached to provide interactive control."*

Finally, two participants mentioned evaluation of visualization artefacts supported by elements of human computer interaction (HCI) as important concepts, and two participants talked about the relevance of learning the visualization process—i.e., understanding visualization pipelines.

### 3.2 Difficult Topics

Several of the concepts included in the list of important concepts were also included in the list of difficult ones. Seven participants mentioned computational and statistical methods; six participants talked about evaluation of visualization artifacts supported by elements of HCI; five mentioned design creativity; and four participants included data complexity, the design considerations, and visualization pipelines as difficult concepts or topics.

Three of the participants considered that learning about human perception, and learning to consider context of the data, the user, and the task were all difficult for students. Likewise, interaction design for visualizations was also described as difficult to learn by three of the participants.

Here we cite some of the justifications that participants included in their responses. For instance, when explaining why evaluations and HCI were included, one of the participants said:

> *"Humility is always a hard thing to learn. Get out of the way of the actual task at hand. Support that work by doing good work. ...Also, most user interaction frameworks/concepts are a real pain to use. Computer science is still working on this one…"*

Regarding interaction design, one of the participants explained that: *"It is often difficult to understand which interaction technique is useful for which purpose."*

Two new categories emerged from this question: (1) Simplicity in design and (2) Communicating uncertainty. Simplicity in design included concept descriptions such as avoiding perfectionism and "*less is often more*", and was mentioned by three participants.

The two participants that included communicating uncertainty among the difficult concepts mentioned that *"there are rich taxonomies of uncertainty that can be applied to inform visualization design, but designers are not trained on them"* and *"Classical mathematics education emphasizes proof and correctness, meaning students are poorly prepared for numerical uncertainty"*

## 4 Conclusions and Next Steps

This study presented the preliminary findings of the first phase of Delphi study aimed at identifying the important and difficult topics and concepts to learn in InfoVis. The important topics included: human perception; interaction design; data complexity; computational and statistical methods; design creativity; considering the context; design considerations; evaluation of visualization artifacts and HCI; and visualization pipelines. This set of important concepts and topics can be used to inform InfoVis courses and programs.

The most difficult topics to learn in InfoVis had a large overlap with the important concepts. This list included: computational and statistical methods; evaluation and HCI; design creativity; data complexity; design considerations; visualization pipelines; human perception; considering the context of the data, the user, and the task; designing interactive visualizations; simplicity in design; and communicating uncertainty. This list of difficult concepts and topics should be considered by educators and educational researchers when designing and evaluating effective pedagogical strategies in InfoVis.

Understanding the complexity of the data, and learning computational tools and methods require special consideration since both of them are near the top in both lists. These results suggest that these two concepts are not only important but difficult to learn, and therefore, educational research should make an extra effort to investigate instructional strategies to support students' learning.

Several of the elements identified as important and difficult to learn fall into the category of topics rather than specific concepts. Additional research is required to decompose these topics into

specific concepts. Understanding what the difficult concepts to learn are, will enable researchers to identify why these are difficult, and how educators can better support student learning of these concepts.

The small number of participants may bring bias to our current study. Thus, next steps for this project include inviting more educators, researchers, and practitioners to contribute with their expertise to complete this initial list. The research team will take the opportunity of presenting at this workshop to invite educators and researchers to participate in this study. Next steps for this project include inviting more educators, researchers, and practitioners to contribute with their expertise to complete this initial list. The consolidated list will be shared during a second round of data collection asking participants to rank them based on their importance and their level of difficulty. A third and final round will be carried out to reach a consensus among experts regarding the most important and difficult concepts to learn in InfoVis.

## REFERENCES


[1] Purdue Polytechnic. Data Visualization - Purdue Polytechnic Institute. [online] Available at: https://polytechnic.purdue.edu/degrees/data-visualization [Accessed 20 Jul. 2017].

[2] Berkeley. Master of Information and Data Science. [online] Available at: https://datascience.berkeley.edu/academics/ [Accessed 20 Jul. 2017].

[3] K. Goldman, P. Gross, C. Heeren, G. Herman, L. Kaczmarczyk, M. C. Loui, and C. Zilles. Identifying important and difficult concepts in introductory computing courses using a delphi process. *ACM SIGCSE Bulletin*, *40*(1), 256-260. 2008

[4] R. Streveler, B. M. Olds, R. L. Miller, and M. A. Nelson. Using a Delphi study to identify the most difficult concepts for students to master in thermal and transport science. In American Society of Engineering Education, Annual Conference, June 2003.

[5] J. Pill. The delphi method: substance, context, a critique and an annotated bibliography. Socio-Economic Planning Sciences, 5(1):57–71, 1971