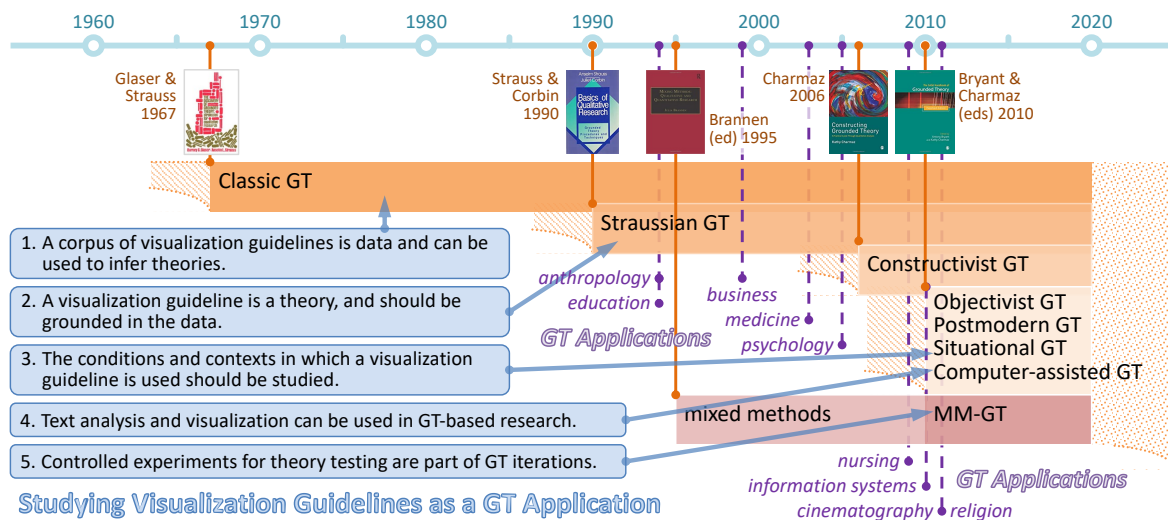# Characterizing Grounded Theory Approaches in Visualization

A. Diehl[1], A. Abdul-Rahman[2] , B. Bach[3], M. El-Assady[4], M. Kraus[5], R. S. Laramee[6], D. A. Keim[5], M. Chen[7]

[1]University of Zurich, Switzerland   [2]King's College London, UK   [3]Edinburgh University, UK
[4]ETH AI Center, Zurich, Switzerland   [5]University of Konstanz, Germany   [6]Nottingham University, UK   [7]University of Oxford, UK

Figure 1: *Over the last five decades, the scope of GT has been broadened gradually. Many methods used in VIS for studying discourse corpora fall within the scope of GT, and the combined uses of these methods can strengthen the GT applications in VIS.*

Abstract
Grounded theory (GT) is a research methodology that entails a systematic workflow for theory generation grounded on emergent data. In this paper, we juxtapose GT workflows with typical workflows in visualization and visual analytics, shortly VIS, revealing the characteristics shared by these workflows. We explore the research landscape of VIS to observe where GT has been applied to generate VIS theories, explicitly as well as implicitly. We discuss "why" GT can potentially play a significant role in VIS. We outline a "how" methodology for conducting GT research in VIS, which addresses the need for theoretical advancement in VIS while benefiting from other methods and techniques in VIS. We exemplify this "how" methodology by adopting GT approaches in studying the messages posted on VisGuides — an Open Discourse Forum for discussing visualization guidelines.

CCS Concepts
• *Human-centered computing* → *Visualization theory, concepts and paradigms;*

1. Introduction

Grounded theory (GT) is a well-established methodology in social science to scrutinize a postulated theory by *grounding* it in data. As depicted in Figure 1, GT was first proposed by Glaser and Strauss in 1967 for sociological research [GS67], and its scope has been further broadened since 1990. In addition to disciplines in social science, it has been applied to medicine, nursing, and computer science.

At a high level, GT workflows bear a strong resemblance to visualization and visual analytics (VIS) workflows. Both types of workflow aim to derive new knowledge from data (e.g., discovering theories in GT), and enable human users to make data-informed

decisions (e.g., determining categories for a concept). At a low level, most GT workflows are yet to benefit significantly from VIS techniques and tools. This motivates us with the first related question: *Should and can VIS techniques be deployed in GT workflows?*

Meanwhile, while GT methods have been applied to study VIS phenomena, the number of VIS publications that explicitly attributed the work to GT is relatively small. This motivates us with the second related question: *Should and can GT methods be more widely deployed in studying VIS phenomena?*

This discourse paper, which follows the tradition of many discourse papers in VIS (Section 2), reports a qualitative analytical investigation into the above two questions. In our investigation, we first research the GT literature and map out different schools of thought in GT. We conclude that classic GT would have welcomed VIS if such technology had been available in the 1960s when GT was first proposed. We find that modern GT encompasses computer-assisted GT methods as well as empirical studies (Section 3). We then compare a typical GT workflow with a typical VA workflow and a typical hypothesis-testing workflow. We examined the similarity and difference among them, and outlined a VIS4GT workflow that falls within the scope of modern GT and is consistent with the principles of classic GT (Section 4). We thus concluded that VIS techniques **should** be deployed in GT workflows.

Building on the above discourse, we considered several areas where VIS phenomena may be studied using GT methods, and discovered that although the number of GT papers in VIS is small, there have been many VIS research papers exhibiting GT-like research methods. We thus concluded that GT methods **should** and **can** be widely deployed in studying VIS phenomena (Section 5). Following a brief discourse on the benefit of using GT in VIS, we analysed our own experience of using GT in studying the discourse data on visualization guidelines, and enriching our categorization effort with VIS techniques and empirical studies (Section 6). We concluded that VIS techniques **can** be deployed in GT workflows.

2. Related Work

In this section, we provide a summary review of discussions and discourse papers in VIS. We then briefly review VIS publications where GT methods were mentioned.

2.1. Discourse-based Research in VIS

In VIS, a good number of discourse papers have had significant influence on the subject development of VIS, including, e.g., Van Wijk's discourse on the value of visualization [VW05] and Munzner on design workflows [Mun09]. The discourse on visual analytics (VA) by Thomas and Cook [TC05] and Keim et al. [KAF*08, KKEM10] shaped the subject of visual analytics today. In general, discourse-based research enables researchers to translate their experience and observations into insight on a specific topic, engage in scholarly discussions and critiques, and propose new methodologies and research directions. While it is not feasible to include many discourse papers in this brief review, here we highlight a number of important papers representing different research fora.

Discussions and Critiques. Tory and Möller discussed the subject division in VIS [TM04] and argued for the importance of the

human as a factor to consider during the visualization design process [Tor14]. Chen et al. questioned if the term *insight* answers the question "what is visualization really for?" [CFB14]. Streeb critiqued some existing arguments about the biases in VIS [SCK18]. Chen et al. provided analytical reasoning about the cost-benefit of deploying VIS applications in VR environments [CGJM18]. Dimara and Perin offered their answers to the question "what is interaction for VIS?" [DP19]. There are also discourse papers on discourse. For example, Streeb et al. examined the numerous arguments about the question "why we need visualization" or "visual analytics" [SEAKC21b, SEAKC21a]. Chen and Edwards examined a collection of schools of thought or "isms" in VIS [CE20].

Conceptual Models. Built on the VIS pipeline proposed by Van Wijk [VW05] and VA pipeline by Keim et al. [KAF*08], Sacha et al. proposed an extended VA pipeline to include analytical reasoning, knowledge acquisition and decision making processes in the mind.

Methodologies. Following Munzner's discourse on design workflows, there have been a series of discourse papers on design study methodologies [SMM12, BDFM14], using guidelines in design [MSQM15], criteria for rigorous design studies [MD19], problem-driven visualization [Mar17, HBH*19], discourse in design processes [BDRW20], and best design practices [PGBC20, CFGT20, Par21]. Lam et al., Isrnberg et al., Elmqvist, Saket et al. provided a series of discourse on evaluation methodologies in VIS [IIC*13, LBI*11, SES16, EY15]. Brehmer et al., Rind et al., and Kerracher et al. proposed methodologies for tasks abstraction in VIS [BM13, KK17, RAW*16].

Research Agenda. Johnson's discourse on scientific visualization research agenda outlined a number of important unsolved problems [Joh04]. Chen et al. mapped out the pathways for theoretical advances in VIS [CGJ*17]. Bradley et al. proposed interdisciplinary research agenda in conjunction with digital humanity (DH), including GT methods in VIS [BEC*18].

These discourse papers confirm the tradition and the need for discourse-based research in VIS. We can infer the need for using GT in studying discourse in VIS. This work provides a new discourse on an important interdisciplinary topic between VIS and social science. In particular, we discuss the need for increasing the uses of GT methods in VIS, and propose a VIS-assisted GT workflow that can benefit from VIS techniques and tools as well as empirical studies.

2.2. Applications of GT in VIS

GT has been accepted as a major qualitative research methodology by the field of human-computer interaction for at least two decades [Cre02]. Knigge and Cope [KC06] proposed an integrated analytical method that combines GT, geographic information systems, and ethnography in the area of environment and planning, and coined the term *Grounded Visualization*. In the context of visualization, Kandogan and Lee applied GT methods to study visualization guidelines [KL16]. Isenberg et al. used GT for evaluation in VIS and referred to it as *Grounded Evaluation* [IZCC08]. Seldmair et al. proposed a new design study methodology and compared it with other methodologies such as ethnography, action research, and GT [SMM12]. Lee et al. used GT to construct a grounded model of guidelines for making sense of unfamiliar visualizations [LKH*15].

Bigelow et al. used GT to investigate how VIS designers work with data [BWI20]. Lundgard and Satyanarayan used GT to understand the semantics of natural language descriptions and derived a four-level model of semantic content [LS21]. Diehl et al. collected discourse on different visualization guidelines and practices and applied GT to study the collected corpus [DKAR*20, DAREA*18].

In Section 5, we will discuss previous work in the VIS literature even further, which may feature GT-link workflows. In Section 6, we will use the work by Diehl et al. [DKAR*20] to inform the discourse on how VIS techniques and VIS empirical studies may be used in GT workflows.

3. Grounded Theory: Terminology and Concepts

Grounded Theory (GT) is a research methodology centered on the principle that theories must be grounded in the data[†]. GT was first proposed by Glaser and Strauss for sociological research [GS67] with the goal of constructing and scrutinizing theories from data. Applications of GT normally feature two major processes: (i) data collection, e.g., in the form of examples, counter examples, and case studies; and (ii) data analysis where data is examined through methods such as as *categorization*, *coding*, *constant comparative analysis*, *negative case analysis*, and *memoing*. The ultimate goal of GT is to derive a theory from raw data and/or scrutinize a proposed theory systematically by constantly and continuously sampling the data space, analyzing the data captured, and refining the theory until it reaches *theoretical saturation*. The first two columns of Table 1 summarize and compare these methods.

In addition, in the original proposal of the classic GT [GS67], Glaser and Strass articulated the need for computing statistical measures for the categorized data, and for placing the statistical measures in tables intelligently such that the relationships among different categories can be observed. These two methods remind us of the basic activities in "analytics" and "visualization" in VIS, while we can easily understand why these two words were absent in the original proposal of GT [GS67] due to the historical context. We therefore added two extra rows at the end of Table 1 to highlight the presence of "analytics" and "visualization" in GT. The formal descriptions of these methods and principles, which was compiled based on the work of Willig [Wil13], can be found in Appendix A.

As illustrated in Figure 1, the scope of GT has been broadened over the past five decades [KF14]. The original version of GT asserted that a theory must be "naturally" emerged from the data, and GT researchers must be open-minded and must abstain from existing theories in the literature. 1990 saw the first major extension of the scope. Strauss (one of the original proposers) and Corbin challenged this assertion, reinterpreting "open mind" as not "empty mind" [SC90]. This broadened the scope of GT beyond the restriction of studying "naturally-emerged" theories only. In 2006, Charmaz, a student of Strauss and Glaser, further broadened the scope of GT by allowing GT researchers to use their prior knowledge more actively, especially in coding, sampling, raising questions,

and designing further iterations [Cha06]. Meanwhile, Glaser argued to preserve the original scope of GT, which is now referred to as *Classic GT*, while the two extended versions are referred to as *Straussian GT* and *Constructivist GT* respectively. The insistence for maintaining the "pure GT" is often referred to as *Glaserian GT*.

Several other extensions were proposed before or around 2010. The SAGE Handbook of Grounded Theory edited by Bryant and Charmaz [BC10b] consists of 27 articles on GT, featuring different versions and applications of GT. In their introductory article of the handbook [BC10a], Bryant and Charmaz explicitly defined "grounded theory method as a family of methods," and marked the handbook as a celebration of difference. In another article, Denzin identified seven versions of GT, positivist, postpositivist, constructivist, objectivist, postmodern, situational, and computer assisted GT [Den10], with the first three roughly corresponding to the aforementioned classic GT and its two extensions.

In their introductory chapter [BC10a], Bryant and Charmaz also noted that the original GT did not in anyway reject quantitative methods. Glaser and Strauss "intended to show such [qualitative] research projects could produce outcomes of equal significance to those produced by the predominant statistical-quantitative, primarily mass survey methods of the day." "Glaser has always argued that the method applies equally to quantitative inquiry." In fact, Glaser and Strauss clearly stated in their 1967 book [GS67]:[‡] "Our position in this book is as follows: there is no fundamental clash between the purposes and capacities of qualitative and quantitative methods or data. ... We believe that *each form of data is useful for both verification and generation of theory,* whatever the primacy of emphasis. ... *In many instances, both forms of data are necessary*." Many scholars have adopted mixed-methods approaches featuring both qualitative and quantitative methods. For example, Brannen included GT as a major method in her 1995 book on mixed methods [Bra95]. Almost all books thereafter on mixed methods approaches have included GT as component method (e.g., [Cre02, TNMC04, Ber18]). In 2010, Johnson et al. proposed a mixed-methods version of GT (MM-GT) for "connecting theory generation with theory testing" [JMT10].

The contemporary, broadened scope of GT provides an ideal research framework for applying qualitative research methods to the discourse data in the field of VIS, while the GT-based studies can benefit from the research methods familiar to visualization researchers, such as digital data collection, text visualization, and empirical studies.

4. Comparing GT Workflows and VA Workflows

Grounded Theory (GT) advocates research workflows that are different from the traditional hypothesis-testing workflows. As illustrated in Figure 2, the two types of workflows differ in several ways:

- A hypothesis-testing workflow W_{HT} begins with reading the literature to formulate a hypothesis, while a GT workflow W_{GT} begins with defining phenomena to be studied without any hypothesis. Note that the notion "without any hypothesis" strictly belongs

---

[†] GT itself is not a theory. In this paper, the term "a theory" means a theoretical postulation that is to be formulated and evaluated using GT methods, unless it is stated explicitly as a confirmed theory.

[‡] The italicized excerpts were highlighted by Glaser and Strauss originally in their book [GS67].

Table 1: *A summary of the principles and methods of grounded theory, and an outlook on how computer-assisted implementation may complement the traditional implementation of the grounded theory (GT) methodology [Wil13]. In the last two rows, we use place "analytics" and "visualization" in quotation marks because they are not the original terms used in traditional implementation of GT methods.*

Methods and Principles	Traditional Implementation	Computer-assisted Implementation
Categorization	Close reading; category identification.	Distant reading with statistics and statistical graphics; cluster and similarity analysis; text analysis & visualization.
Theorization	Taxonomy & ontology construction; entity relationships (e.g., correlation, association, causality, etc.); guidelines; conceptual models.	Visual analytics; machine learning; algorithmic taxonomy & ontology learning; correlation, association analysis & causality analysis; dimensionality reduction; modelling & simulation.
Coding	Close reading and labelling; open coding; axial coding; selective coding.	Text and discourse analysis; text visualization; similarity & association analysis; network visualization; ontology mapping.
Comparative Analysis	Iterative close reading; comparing different options of category abstraction; continuous refinement of the categorization scheme; proposing new categories or categorization schemes.	Iterative and continuous effort for distant reading with statistics and statistical graphics; cluster & similarity analysis; topic modelling; text visualization; algorithmic taxonomy & ontology learning.
Negative Case Analysis	Close reading; negative case identification & analysis.	Algorithmic outlier analysis & anomaly detection.
Memoing	Memo-writing; sketch-drawing; systematically recording the ideas and actions related to the development of a theory.	Online discussion forum; crowd sourcing; provenance visualization; data collection from social media.
Theoretical Sensitivity	Testing a theory against old and new data (i.e., "Interact" with data); analysis of applicability, negative cases, and possible new categories.	Computer-assisted data collection; data analysis & visualization.
Theoretical Sampling	Data collection; empirical studies (e.g., controlled experiments, surveys, focus groups, field observation, etc.); interview transcription.	Online discussion forum, crowd sourcing; data collection from social media; computer-assisted experiments; computer-mediated group activities; observation of online activities.
Theoretical Saturation	Data collection and analysis until theoretical saturation has been achieved.	Computer-assisted platforms for facilitating the longevity and provenance of the GT processes.
"Analytics"	Computing statistical measures of categorized data.	Statistical inference; data mining; data modelling; visual analytics.
"Visualization"	Displaying statistical measures in tables with intelligent alignment.	General-purpose and special-purpose visual representations; interactive visualization; visual analytics.

to the Classic GT or Glaserian GT, as the broadened versions of GT, such as Straussian GT and Constructivist GT, considered that it would be natural to have thoughts on the phenomena to be studied.

- In W_{HT}, one collects data in order to test a hypothesis, while in W_{GT}, one collects all collectable data associated with the phenomena. With a known hypothesis, one can normally collect data by focusing on a few independent and dependent variables, where the notion of dependency is usually part of the hypothesis to be tested. Without any hypothesis, one normally has to collect less focused data featuring many variables.

- In W_{HT}, collected data represents the sampling of the independent variables that are predefined for a hypothesis, and typically statistical analysis are used to derive measures about predefined dependent variables. The statistical properties associated with both independent and dependent variables are used to reason about the hypothesized relationships among these variables, e.g., scatter plots may be more useful than parallel coordinates plots [KZZM12], visualization processes benefit from human knowledge [KARC17].

In W_{GT}, as independent and dependent variables are not predefined, it is necessary to analyze the collected data to identify such variables by "coding", which attempts to identify from data a number of *concepts* (or *attributes*) and the categories (or types) of each concept (or attribute). Mathematically, a concept (or an attribute) is variable, and a category (or a type) is a valid value of a variable. As shown in the lower half of Figure 2, coding and data analysis are two highly integrated processes.

- In W_{HT}, data is often collected in an organized tabular form, while in W_{GT}, data is often collected as free-form textual data or less-organized tabular data (e.g., tabular cells may have inconsistent or missing data values). Hence, GT methods are often used to study discourse data.

- In Figure 2, W_{GT} appears to be more iterative than W_{HT}. Later we will point out that W_{HT} is also iterative but in a different way.

Although the terms "hypothesis" and "hypothesis testing" do not appear in the GT workflow illustrated in Figure 2, it would be overly-strict to assume the absence of hypotheses. For example, to derive a categorization scheme using coding, the researcher has
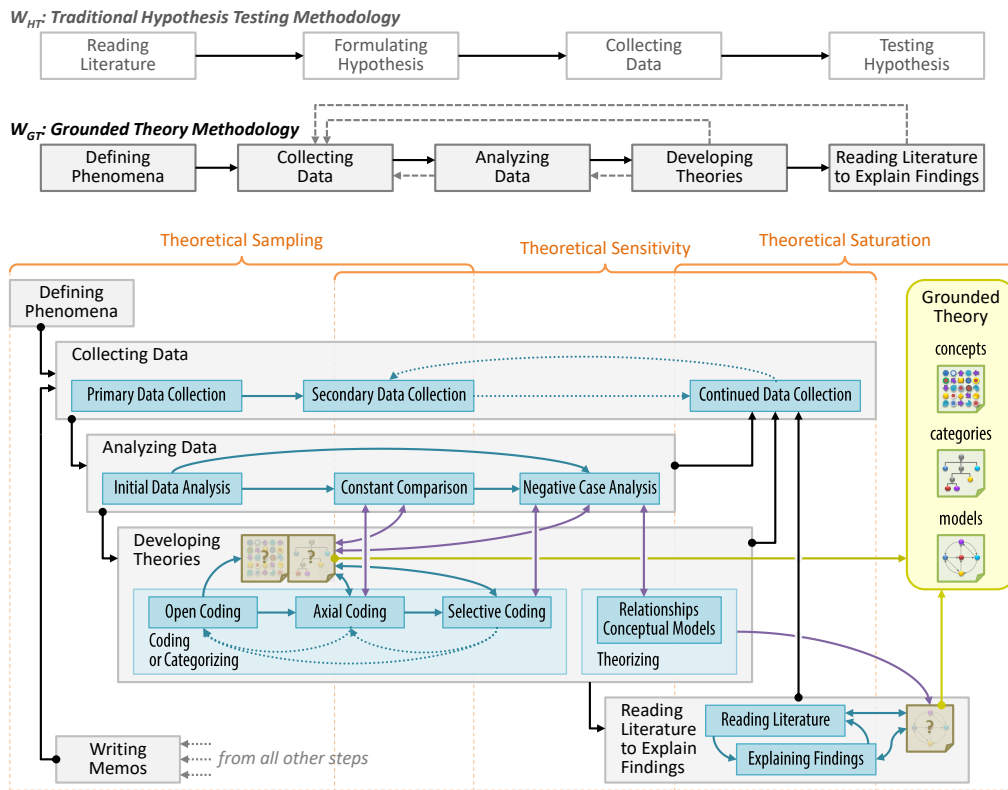
Figure 2: *A traditional hypothesis testing workflow is juxtaposed with a GT workflow, which is detailed in the lower half of the figure.*

to contemplate different optional schemes and hypothesize their meaningfulness and usefulness. The judgement about theoretical sensitivity and saturation involves hypotheses as to whether new effort for data collection, coding, or data analysis would change the theoretical findings obtained so far.

Although the terms "statistics" and "visualization" do not appear in the GT workflow illustrated in Figure 2, they are very much part of the data analysis process in GT since it was proposed [GS67]. As mentioned in Section 3 and Table 1, after coding, the researcher typically derives statistical measures for different categories (or types) for each concept (or attribute). Glass and Strauss particularly articulated the need to design appropriate data tables where statistical measures can be aligned intelligently to reveal relations among different categories (or types) and among different concepts (or attributes). Although such visual observation can clearly benefit from visualization techniques, likely and understandably, Glass and Strauss had limited access to visualization techniques at that time (more than five decades ago).
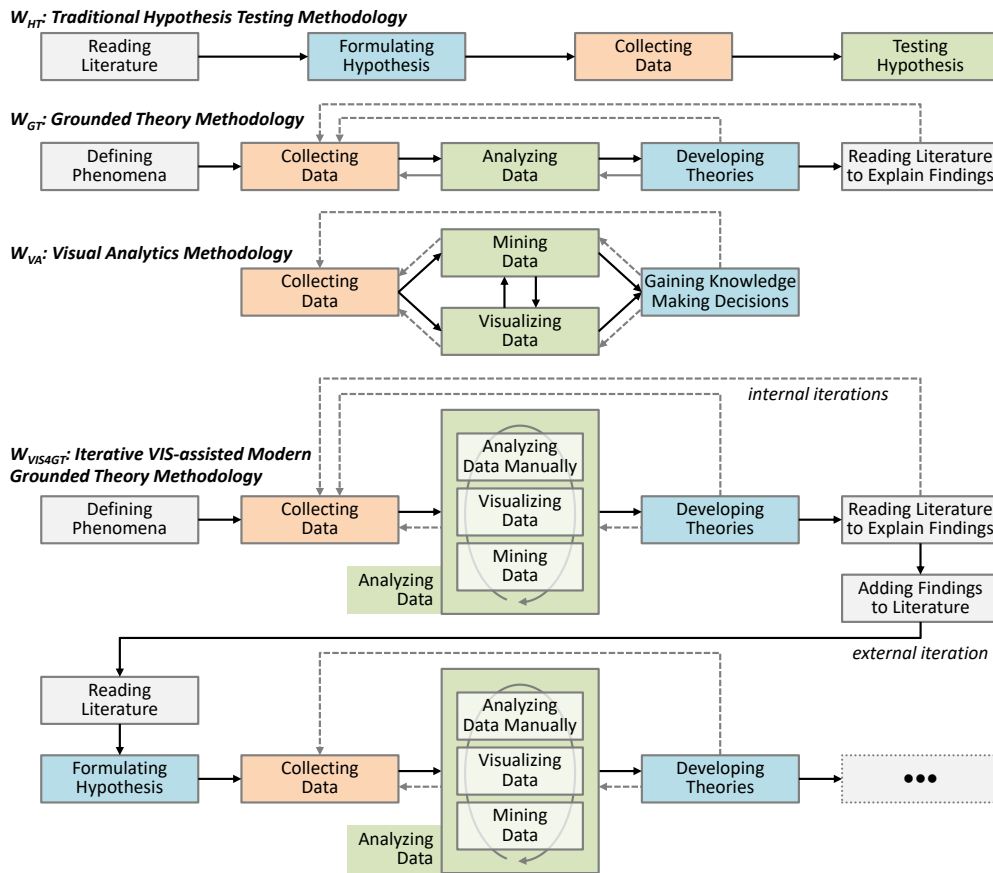
This reminds us of visual analytics (VA) workflows outlined by Keim et al. [KAF*08]. The upper half of Figure 3 juxtaposes W_{HT} and W_{GT} with a VA workflow W_{VA}. We encode process-blocks of a similar function using the same color. In general, "Testing Hypothesis" in W_{HT} is a data analysis process, while "Analyzing Data" in W_{HT} can be seen as a technically simpler version of "Mining Data" and "Visualizing Data" in W_{VA}.

In the lower half of Figure 3, we add the "Mining Data" and "Vi-

sualizing Data" processes in W_{VA} into the "Analyzing Data" process in W_{GT}. Such an integration is an implementation of Denzin's proposal of computer assisted GT [Den10]. We refer to such a workflow as VIS-assisted GT methodology, and denote it as W_{VIS4GT}. In the right column of Table 1, we elaborate various VA techniques that may be used for assisting different GT methods.

Although the W_{HT} workflows in Figures 2 and 3 do not feature any feedback loop, traditional hypothesis-testing workflows are in fact fairly iterative. In particular, in psychology and VIS, many hypotheses have been evaluated by multiple empirical studies over a long period, which often feature different experiment designs and different analytical methods, and lead to different conclusions. For example, Kuang et al. conducted a study to test a hypothesis that scatter plots (SPs) are more effective than parallel coordinates plots (PCPs), and obtained a positive conclusion [KZZM12]. Kanjanabose et al. noticed the "theoretical sensitivity" of the conclusion, and designed a study to collect new data for comparing SPs, PCPs and data tables [KARC15]. The data analysis showed that data tables are more effective than SPs and PCPs for value retrieval tasks, PCPs are more effective for some tasks such as clustering, outlier detection, and change detection. On the topic of perceiving correlation through SPs, there are over 20 hypothesis-testing studies in the literature (see literature review by Sher et al. [SBLC17]).

For hypothesis testing, a hypothesis is usually iteratively tested by different research teams in different W_{HT} workflows. We refer to such iterative research workflows as *external* iterations. For a GT study, it is often the same researcher that conducts multiple iterations

W_{HT}: Traditional Hypothesis Testing Methodology

Reading Literature → Formulating Hypothesis → Collecting Data → Testing Hypothesis

W_{GT}: Grounded Theory Methodology

Defining Phenomena → Collecting Data → Analyzing Data → Developing Theories → Reading Literature to Explain Findings

W_{VA}: Visual Analytics Methodology

Collecting Data → Mining Data / Visualizing Data → Gaining Knowledge Making Decisions

W_{VIS4GT}: Iterative VIS-assisted Modern Grounded Theory Methodology

internal iterations

Defining Phenomena → Collecting Data → Analyzing Data (Analyzing Data Manually / Visualizing Data / Mining Data) → Developing Theories → Reading Literature to Explain Findings → Adding Findings to Literature

external iteration

Reading Literature → Formulating Hypothesis → Collecting Data → Analyzing Data (Analyzing Data Manually / Visualizing Data / Mining Data) → Developing Theories → ●●●

Figure 3: *Juxtaposing a traditional hypothesis testing workflow, a GT workflow, and a VA workflow, we can observe that their key processes blocks share similar functionality, e.g., orange blocks for* data collection*, green blocks for* data analysis*, and cyan blocks for data-informed decision making, knowledge acquisition, and hypothesis/theory generation. As the modern GT scope encompasses computer-assisted techniques and hypothesis testing, we can embrace an iterative VIS-assisted GT workflow (VIS4GT) as illustrated in the lower half of this figure.*

as shown in Figure 2. We refer to such iterative research workflows as *internal* iterations. With the increasing adoption of GT methods in different applications (Figure 1), it becomes more common for different research teams to study the same set of phenomena through external iterations. The lower half of Figure 3 illustrates both internal and external iterations in such a VIS-assisted GT workflow W_{VIS4GT}. Based on the above reasoning, we can conclude:

1. A corpus of discourse collected in the field of VIS can be studied using GT in order to infer theories as GT workflows are particularly suitable the free-form textual data in such a corpus;
2. A proposed visualization theory (e.g., a guideline) should be grounded in the data, and thus GT methods can be used to study such a theory;
3. When a visualization theory (e.g., a guideline) is considered as a phenomenon to be studied using GT methods, the data collected will likely feature the conditions and contexts in which the theory is applicable or not;
4. Many VA techniques, including text analysis and visualization, can be used to assist GT research;
5. Controlled experiments for theory testing are part of GT iteration,

if the hypothesis to be tested is a proposed theory originally derived from previous iterations.

As illustrated in Figure 1, these conclusions are consistent with the broadened scope of GT.

5. Where are GT Approaches in VIS?

From the previous two sections, we can observe that the Classic GT does include statistics and table-based "visualization" as part of GT workflows. As depicted in Figure 1, the contemporary schools of thought embrace external iterations, computer-assisted GT and empirical studies, all of which hinge on the interpretation that "open mind" is "not empty mind" [SC90]. Nevertheless, these schools of thought all agree that theories to be postulated and evaluated in GT workflows must be grounded in real-world data. In other words, there must be a provenance for a sequence of GT workflows, regardless internal and/or external iterations, with which such a theory can be traced back to some real-world data. As most GT applications feature qualitative data and free-form textual data, we consider here such discourse, text, and imagery corpora in the following visualization contexts:

- *Context A*: Research papers and discourse about them;
- *Context B*: Visualization guidelines and discourse about them;
- *Context C*: Visualization images and discourse about them;
- *Context D*: Discourse about subject developments;
- *Context E*: Discourse about schools of thought;
- *Context F*: Discourse about design practices;
- *Context G*: Discourse about visualization literacy.

5.1. Explicitly-Labelled GT Approaches

Visualization researchers have been studying such textual corpora for decades. In some cases, the terms of "grounded theory" or "open coding" were mentioned in the published work. These explicitly-labelled GT work in VIS include:

Context A. Isenberg et al. [IIS*16] conducted a retrospective analysis of keywords in VIS papers. The use of GT was mentioned. Kim et al. [KJRK21] applied GT to study papers published for the last 20 years on visualization accessibility. Using coding and thematic analysis, they derived a design space for accessible visualization.

Context B. Kandogan and Lee [KL16] conducted an extensive GT study on visualization guidelines. They extract some 550 guidelines from books, papers, and blogs, and formulated five high-level concepts (i.e., data, visualization, user, insight, device) as well as qualifiers and sub-concepts. They examined patterns, relations, and co-occurrences among them. Diehl et al. [DKAR*20] used GT to study the posts on an open discourse forum, VisGuides, specially for discussing visualization guidelines. They followed the contemporary schools of thought in GT, and used text analytics and visualization and empirical studies in their GT workflow. Bigelow et al. [BWI20] applied GT to analyze the interviews with data workers and identified concepts and categories, from which they derived guidelines.

Context C. Lundgard and Satyanarayan [LS21] used GT to study the semantic content of natural language descriptions that accompany visualization images. Lee et. al [LKH*15] used GT to analyze different high-dimensional multivariate visualizations that are considered unfamiliar to novice users in order to derive a model of novice users' sense making processes.

Context D. Chandrasegaran et al. [CBK*17] proposed to integrate VA techniques into GT practices in qualitative text analysis. Sperrle et al. [SEAG*21] used GT to analyze qualitative feedback and findings on human-related factors that influence human-centered machine learning.

Context E. *We have not found explicitly-labelled GT work on schools of thought in VIS.*

Context F. Rajabiyazdi et al. [RPOC20] used GT to study focus group discussions in relation to patient-generated data visualization. Leon et al. [LLB20] used GT to study survey data on designing global choropleth maps. The work by Kim et al. [KJRK21], mentioned in the *context A* also falls into this context.

Context G. Adar and Lee [AL20] applied GT to a corpus of communicative visualization and interview data. They built their GT workflows on Bloom's taxonomy for teaching, learning, and assessing [BAC*01]. The work demonstrates "open mind is not empty mind" – a contemporary school of thought in GT.

5.2. Implicitly-labelled or Unlabeled GT Approaches

Many research papers in VIS involved extensive study of research papers, a collection of which is a text corpus. Hence such work would have featured GT-like workflows, though GT was rarely explicitly mentioned. Consider the aforementioned seven contexts:

Context A. Many researchers proposed categorization schemes, typologies, taxonomies and ontologies based on extensive literature research. The effort for identifying taxa in a taxonomy or an entity in an ontology is very similar to coding in GT, while determining hierarchical and ontological relations is a form of theorization. Examples of such papers include [BCD*12, BKT*21, LBI*11, ARCL20].

Many survey papers, which also fall into Context A, produced categorization schemes, taxonomies, ontologies, guidelines, and models that also feature other contexts. For example,

- **Context B.** Borgo et al. [BKC*13] collected a list of guidelines on glyph-based visualization.
- **Context D.** Rogowitz et al., Tovanich et al., Crisan et al. examined the issues of subject development of VIS (e.g., history, diversity, polarization, and biases) through the lens of research papers and conferences [RDI*19, TDI21, CFGT20].
- **Context E.** A few papers extracted the categories of different schools of thought or arguments in VIS, e.g., [SEAKC21b, CE20].
- **Context F.** Many papers proposed design methods, design spaces, and design guidelines based on literature research, e.g., [HBH*19, HQC*18, CJS*21].
- **Context G.** Liu et al. [LAB*21] developed a taxonomy of visualization resources for improving visualization literacy.

Some research papers report categorization schemes, typologies, taxonomies, guidelines, methodologies, and conceptual models directly from focus group discussions and interviews, e.g., [MD19, Par21] in Context F.

6. Why and How to Adopt GT Approaches in VIS?

Visualization and visual analytics (VIS) has always and will continue to provide other disciplines with effective and efficient techniques and tools for supporting data intelligence workflows that humans need to gain information and knowledge from data and to make data-informed decisions. As a technology, VIS offers a invaluable bridge between machine-centric processes (e.g., digital data collection, statistics, algorithms, computational models, etc.) and human-centric processes (e.g., perception, cognition, knowledge acquisition, decision making, etc.). As a scientific subject, VIS is an interdisciplinary study of both machine- and human-centric processes in the context of visualization, and in particular, the interactions between these two types of processes. As Chen and Jänicke observed through the lens of information theory, when a visualization process is considered as a data communication process, we normally know the encoder well, but not the decoder [CJ10]. In comparison with the algorithms and software for the "encoders", we have rather vague understanding about the "decoders".

Based on above reasoning, the field of VIS should feature a fair amount of social science research. Perhaps it is self-evident that the amount of social science research in VIS has reached the level that is necessary for understanding the "decoders". Most social

science research in VIS has been conducted in the form of empirical studies, largely related to psychology, one of the many subjects in social science. As shown in the survey by Abdul-Rahman et al. [ARBC*19], the number of VIS publications on empirical studies has been increasing in recent years. There is no doubt a need for significantly more social science research in VIS, because:

- *Understanding VIS needs discourse.* — As reviewed in Section 2, many discourse papers have played a significant role in VIS, e.g., Van Wijk's values of visualization [VW05] and Munzner's nested model [Mun09]. Many ideas raised in these discourse papers can benefit from community participation in discussions to enrich and improve these ideas over a period. In many cases, new ideas can emerge or be explored before they become mature enough for a publication. In VIS, ideally, there are more opportunities for researchers to participate in discussing various discourse topics, through discussion fora as well as formal publications. As a scientific discipline, VIS can prosper from discourses, discussions, or debates, which in GT terms, are "theoretical sampling", facilitate "comparative analysis" and "negative case analysis", and can reveal "theoretical sensitivity" (Table 1).
- *Discourse needs abstraction.* — When a scientific or scholarly subject reaches a certain level of maturity, the scientists or scholars will naturally attempt to make abstraction and generalization from empirical evidence and practical experience [CE20]. The researchers in VIS have always pursued and striven for this. GT was developed to assist in abstraction and generalization processes. For example, the most commonly-adopted GT method, coding, is essentially an abstraction method for reducing diverse and complex textual descriptions to a number of categories. The intelligent table-based display of statistics proposed by Glaser and Strauss [GS67] was intended to discover abstract and generalizable relationships among different categories. Meanwhile, iterative applications of GT workflows with continuous sampling, comparative analysis, and negative case analysis are quality assurance processes for abstraction and generation.
- *Abstraction needs VIS assistance.* — GT workflows have been primarily human-centric processes, and will continue to be in the foreseeable future. Using VIS to assist GT is not in any way intended to replace human-centric processes with machine-centric processes, since visualization and interaction are human-centric processes [SEAG*21]. VIS can significantly improve the efficiency and effectiveness of GT workflows, which will be discussed in detail in the following four subsections. On the one hand, when GT workflows are employed to study VIS discourse, VIS researchers are the domain experts as well as technology-enablers. They will naturally employ different VIS techniques to enrich and improve their GT4VIS workflows. Hopefully, the success of employing VIS techniques can be translated to wide deployment of VIS4GT in many other GT workflows.

The increasing use of VIS techniques will no doubt increase the know-how about GT4VIS and VIS4GT in the field of VIS. In the following four subsections, we describe our experience of GT4VIS and VIS4GT when we used GT methods for analyzing discourse on visualization guidelines. While such experience is not a comprehensive instruction about how to deliver GT4VIS and VIS4GT solutions, we hope that it will prompt VIS colleagues to develop additional and improved GT4VIS and VIS4GT solutions.

6.1. Open Discourse

In general, online discourse fora have been more effective in gathering discourse that can be studied using GT methods. VisGuides (`https://visguides.org/`) is an online discussion forum dedicated to visualization guidelines [DAREA*18]. It allows registered users to pose questions and offer advice on visualization, propose and recommend guidelines, share positive and negative experience about certain guidelines, report and reason about successes, failures, and conflicts of guidelines, and suggest ways to refine guidelines.

A guideline essentially defines a causal relation that an action A will cause a consequence B under condition C. In many cases, guidelines are expressed as "one should do A", implying a negative consequence. In most cases, the condition C is yet to be precisely defined. Such a guideline would be considered as a postulated theory in social science and many other fields. With hundreds of visualization guidelines available in the literature [KL16], it would take a very long time for all proposed guidelines to be investigated thoroughly.

VisGuides serves as a digital platform to supply both the *postulated theories*, i.e., the visualization guidelines to be developed and investigated, and the *data*, i.e., the discourses that can be used to generate, evaluate, and refine the theories.

From the GT perspective, VisGuides provides a mechanism for community-based "theoretical sampling", facilitating the evaluation of "theoretical sensitivity". For example, when the failure of a guideline under a specific condition is reported, it is a "negative case" to be analysed, indicating that the guideline is still theoretically sensitive.

VisGuides was launched on September 27, 2017, and according to the latest count, the forum received 577 posts on December 1, 2021, which include contributions from users with different visualization backgrounds, ranging from undergraduate students to highly-respected visualization experts (e.g., in alphabetic order, Niklas Elmqvist, Stephen Few, Eduard Gröller, Robert Kosara, Liz Marai, Theresa-Marie Rhyne, and Bernice Rogowitz, among others).

VisGuides has been also used for visualization education, providing a platform for students to discuss their visual designs, seek advice on visualization guidelines, and offer their discourse on individual guidelines [DFTW*21].

6.2. Support Coding

As mentioned in Section 5, coding has been used by many VIS researchers in categorizing discourse data and other complex data. In the case of discourse data from VisGuides, we have applied open coding, axial coding, and selective coding to the data. Figure 4 illustrates the main steps for transforming raw discourse data (a) to coded concepts and categories in spreadsheets (b), which is then transformed to statistical graphics (c). The detailed descriptions of these steps can be found in an arXiv report by Diehl et al. [DKAR*20].

As mentioned in earlier sections, Glaser and Strauss considered statistics as a post-coding step and used data tables to display statistical measures [GS67]. Today, most users can easily create statistical graphics for such data tables. While tables can display data more precisely, statistical graphics can enable users to overview and compare

| (a) corpus of posts | (b) coding | (c) statistics | (d) interactive visualization |
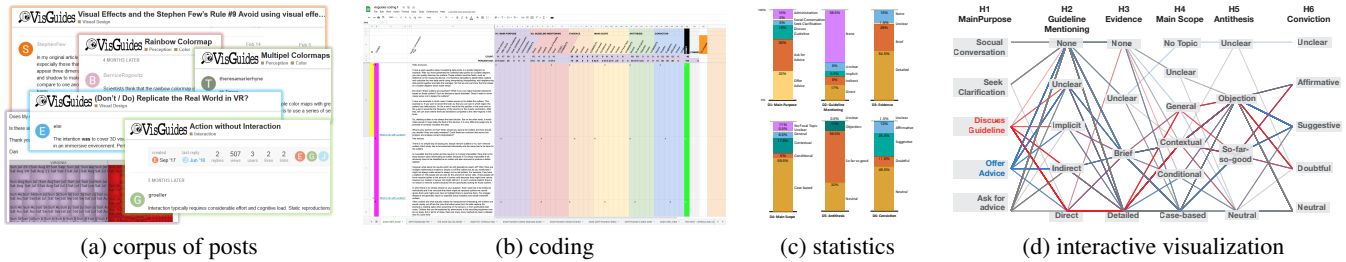
Figure 4: *Using statistical graphics and interactive visualization to support coding and categorisation of posts on VisGuides. The parallel coordinates plot in (d) is used for further analysis and theorization. For detailed descriptions about these images, please see [DKAR*20].*

many statistical measures quickly. For example, during coding, GT researchers frequently encounter a dilemma whether to merge two or a few proposed categories. The relative sizes of different categories (i.e., the numbers of data samples in different categories) are usually important factors that influence the decision. Statistical graphics can enable data comparison more efficiently than data tables.

6.3. Enable Analysis and Theorization

For reasoning about the relationships among coded categories, the GT proposers, Glaser and Strauss, recommend aligning the numbers in the data table intelligently. This is indeed a multivariate data visualization problem. For example, Figure 4(d) shows a parallel coordinates plot for investigating the relationships among different categories under concepts (i.e., variables) H1, H2, ..., H6. The thickness or opacity and the lines encodes the percentage/proportions of a given category and dimension. A GT researcher can interactively brush categories (e.g., the categories of "Discuss Guideline" (red) and "Offer Advice" (blue), and observe the strength of their relationships with other categories. There is little doubt that interactive visualization can significantly improve the process of aligning table items intelligently.

There are many other multivariate analytical and visualization techniques that can be used for supporting GT, especially, the processes of comparative analysis, negative case analysis, and theorization. For example, one can have tree visualization for proposed taxonomy; graph visualization for proposed ontology, association, and causality; VA-based clustering and anomaly detection for comparative analysis and negative case analysis; and coordinated multiple views for complex analysis and theorization in GT workflows.

6.4. Facilitate Theoretical Sampling

Because the GT processes in earlier parts of a GT workflows typically focus on coding free-form textual data for which categorization schemes are not yet defined, close reading of raw data has always been a dominant mechanism. We do not suggest replacing automated text analysis with close reading in GT workflows.

However, all those who have experienced a coding exercise appreciate the time-consuming nature of close reading. Although the coding activities are required to be "open-minded", in practice, contemplating and exploring many optional categorization schemes is usually hindered by lack of human and time resources. The difficulties in contemplating and exploring options fundamentally hamper

the GT principle of "Theoretical Sampling", i.e., one cannot afford to sample many options.

Text analysis and visualization is a family of VA techniques that can significantly alleviate the difficulties in contemplating and exploring optional categorization schemes. With such techniques, a GT researcher can contemplate a categorization scheme using a collection of keywords that are mapped to different concepts and categories. The text algorithms can scan the raw text and attempt to determine the category (or categories) of the text concerned. Text visualization can quickly display the results, enabling the GT researcher to judge if the proposed categorization scheme is promising or not, and if further close reading should be performed. In this way, the GT researchers can explore more optional categorization schemes, as required by the GT principle of "Theoretical Sampling".

In addition to statistical graphics for displaying the measures of the proposed categories, graph visualization can be used to depict the relationships between the categories. For example, Figures 5(a,b) show two visual representations used in conjunction with category analysis (i.e., referred to as entity analysis in text mining).
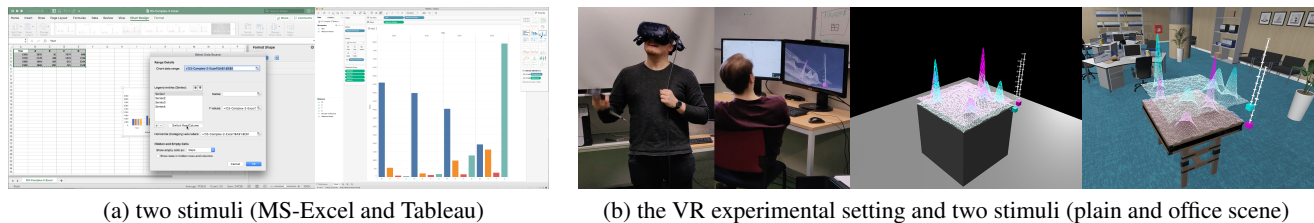
Text visualization can also support close reading. Figure 5(c) shows a lexical episode plot [GREA15] used to depict the evolution of the discourse over time. The analytical algorithm automatically detects compact chains of *n*-grams and highlights them beside an abstract overview of the complete text. We can interactively zoom-in to the interesting text areas for close reading. Large versions of Figure 5, plus further details about using text VA to analyze VisGuides discourse, can be found in an arXiv report by Diehl et al. [DKAR*20].

6.5. Address Theoretical Sensitivity

Theoretical Sensitivity is another important principle of GT. It obligates GT researchers to address uncertainty in a postulated theory generated in the previous iterations by continuing the iterative workflow, including collecting more data if necessary. One contention in the early schools of thought of GT was whether empirical studies would be considered as GT methods. On the one hand, empirical studies are typically designed to test certain hypotheses, which appeared not to be open-minded. On the other hand, apart from the first iteration, all other iterations in a GT workflow must be influenced by the findings in the previous iterations. Such findings are hypotheses until theoretical saturation. Many researchers argued for the extension of the traditional GT to include controlled empirical studies in order to evaluate and validate hypotheses [HP92, HP03, BL12].

(a) a named entity graph [EASG*17]   (b) a topic tree [EASD*19]   (c) a Lexical Episode Plot [GREA15]
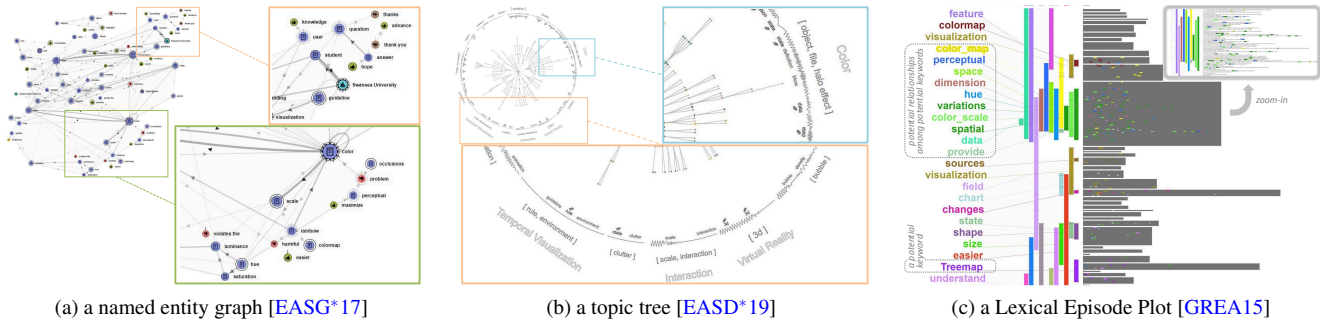
Figure 5: *Three different visual text analytics techniques were used in supporting theoretical sampling during the workflow for studying the discourse on VisGuides. For a high resolution version of the images, see [DKAR*20]; and for more detailed explanations, see [EAJS*19].*

(a) two stimuli (MS-Excel and Tableau)   (b) the VR experimental setting and two stimuli (plain and office scene)

Figure 6: *The data collection settings for the two empirical studies (Section 6.5), which were designed to address the theoretical sensitivity of two visualization guidelines: (a) "Use interaction in visualization sparsely and cautiously", and (b) "Don't replicate the real world in VR".*

As mentioned in Section 3, Strauss and Corbin reinterpreted "open mind" as not "empty mind" [SC98], opening the door for the inclusion of empirical studies in GT as mixed method GT.

During the close reading of VisGuides discourse, we noticed two threads of posts suggesting some VIS guidelines are still theoretically sensitive (i.e., not yet conclusive). One thread [eas17] was initiated by VisGuides user `east101` to discuss a guideline *"Use interaction in visualization sparsely and cautiously ..."* [Grö08]. In the forum, user `groeller` explained that *"interaction typically requires considerable effort and cognitive load."*. User `jamesscottbrown` later presented a contrasting argument: *"... there is generally little cost (to the user) associated with adding such interactivity to a visualisation ..."*. In order to address the theoretical sensitivity shown in this thread of discussion, we conducted an empirical study to collect more targeted data. Another thread was initiated by VisGuides user `matthias.kraus` on the topic *"(Don't / Do) Replicate the Real World in VR?"* [mat18]. In this thread, the user asked whether the guideline *"Rule #7: Don't Replicate the Real World"* [Elm] would still hold in a reality (VR) context. The thread received a fair amount of discussion, among others, a comment by Nilkas Elmqvist (user `elm`), confirming that his guideline was meant to include VR. Similarly we noticed the theoretical sensitivity, and conducted an empirical study to collect more targeted data. Figure 6 shows some stimuli used in the study.

Both empirical studies are reported in the arXiv report by Diehl et al. [DKAR*20]. In the context of this paper, these two studies show that the existing knowledge and experience of conducting empirical studies in VIS can enrich GT workflows for studying VIS phenomena. In these two cases, the phenomena are about two visualization guidelines, and the discourse data collected initially through VisGuides exhibits a high-level of theoretical sensitivity.

Further data collection and analysis is necessary, and empirical studies enable us to collect data on a narrowly-focused topic more quickly than a discourse forum.

7. Discussions and Conclusions

In this paper, we present our qualitative analysis of the role of visualization and visual analytics (VIS) in grounded theory (GT) workflows. Our study confirms the following findings:

1. GT methods are particularly useful for studying discourse data. VIS is an interdisciplinary field involving the study of humans, and can benefit from adopting social science research methods.
2. Discourse research is not uncommon in VIS and has significant impact on VIS development. More discourse, including open discourse fora, should be encouraged and facilitated.
3. Since the birth of GT, statistics have been part of GT workflows. Contemporary schools of thought in GT embrace computer-assisted GT. GT workflows will welcome visual analytics.
4. The early idea of arranging data tables intelligently was an attempt of analyzing multivariate data using a crude form of visualization. Interactive VIS can significantly enhance such analytical processes.
5. The requirement for iterative theoretical sampling in GT logically implies hypothesis-based iterations, since it is not feasible to ensure totally open-mindedness in succeeding iterations. Contemporary schools of thought in GT embrace empirical studies as part of mixed method GT (MMGT).

Our recommendation. The field of VIS is naturally suitable for developing research agenda of GT4VIS as well as VIS4GT. We can start by incorporating more VIS techniques and empirical studies in GT4VIS workflows, and aspire to provide a family of VIS4GT techniques to the broad GT communities in other disciplines.

References

[AL20] ADAR E., LEE E.: Communicative visualizations as a learning problem. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2020), 946–956. 7

[ARBC*19] ABDUL-RAHMAN A., BORGO R., CHEN M., EDWARDS D. J., FISHER B.: Juxtaposing controlled empirical studies in visualization with topic developments in psychology. arXiv 1909.03786, 2019. 8

[ARCL20] ABDUL-RAHMAN A., CHEN M., LAIDLAW D. H.: A survey of variables used in empirical studies for visualization. In *Foundations of Data Visualization*. Springer, 2020, pp. 161–179. 7

[BAC*01] BLOOM B. S., AIRASIAN P., CRUIKSHANK K., MAYER R., PINTRICH P., RATHS J., WITTROCK M.: *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Prentice Hall, 2001. 7

[BC10a] BRYANT A., CHARMAZ K.: Grounded theory in historical perspective: An epistemological account. In *The SAGE Handbook of Grounded Theory*. SAGE Publications Ltd., 2010, p. 31=57. 3

[BC10b] BRYANT A., CHARMAZ K. (Eds.): *The SAGE Handbook of Grounded Theory*. SAGE Publications Ltd., 2010. 3

[BCD*12] BORGO R., CHEN M., DAUBNEY B., GRUNDY E., HEIDE-MANN G., HÖFERLIN B., HÖFERLIN M., LEITTE H., WEISKOPF D., XIE X.: State of the art report on video-based graphics and video visualization. In *Computer Graphics Forum* (2012), vol. 31, Wiley Online Library, pp. 2450–2477. 7

[BDFM14] BIGELOW A., DRUCKER S., FISHER D., MEYER M.: Reflections on how designers design with data. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces* (2014), pp. 17–24. 2

[BDRW20] BEECHAM R., DYKES J., ROONEY C., WONG W.: Design exposition discussion documents for rich design discourse in applied visualization. *IEEE transactions on visualization and computer graphics* (2020). 2

[BEC*18] BRADLEY A. J., EL-ASSADY M., COLES K., ALEXANDER E., CHEN M., COLLINS C., JÄNICKE S., WRISLEY D. J.: Visualization and the digital humanities: Moving towards stronger collaborations. *IEEE Computer Graphics and Applications 38*, 6 (2018), 26–38. 2

[Ber18] BERGIN T.: *An Introduction to Data Analysis: Quantitative, Qualitative and Mixed Methods*. SAGE Publications Ltd., 2018. 3

[BKC*13] BORGO R., KEHRER J., CHUNG D. H., MAGUIRE E., LARAMEE R. S., HAUSER H., WARD M., CHEN M.: Glyph-based visualization: Foundations, design guidelines, techniques and applications. In *Eurographics (State of the Art Reports)* (2013), pp. 39–63. 7

[BKT*21] BRESSA N., KORSGAARD H., TABARD A., HOUBEN S., VERMEULEN J.: What's the situation with situated visualization? a survey and perspectives on situatedness. *IEEE Transactions on Visualization and Computer Graphics* (2021). 7

[BL12] BADREDDIN O., LETHBRIDGE T.: Combining experiments and grounded theory to evaluate a research prototype: Lessons from the umple model-oriented programming technology. In *Proceedings International Workshop on User Evaluation for Software Engineering Researchers* (2012), pp. 1–4. doi:10.1109/USER.2012.6226575. 9

[BM13] BREHMER M., MUNZNER T.: A multi-level typology of abstract visualization tasks. *IEEE transactions on visualization and computer graphics 19*, 12 (2013), 2376–2385. 2

[Bra95] BRANNEN J.: *Mixing Methods: Qualitative and Quantitative Research*. Routledge, 1995. 3

[BWI20] BIGELOW A., WILLIAMS K., ISAACS K. E.: Guidelines for pursuing and revealing data abstractions. *IEEE Transactions on Visualization and Computer Graphics 27*, 2 (2020), 1503–1513. 3, 7

[CBK*17] CHANDRASEGARAN S., BADAM S. K., KISSELBURGH L., RAMANI K., ELMQVIST N.: Integrating visual analytics support for grounded theory practice in qualitative text analysis. *Computer Graphics Forum 36*, 3 (2017), 201–212. 7

[CE20] CHEN M., EDWARDS D. J.: "isms" in visualization. In *Foundations of Data Visualization*. Springer, 2020, pp. 225–241. 2, 7, 8

[CFB14] CHEN M., FLORIDI L., BORGO R.: What is visualization really for? In *The Philosophy of Information Quality*. Springer, 2014, pp. 75–93. 2

[CFGT20] CRISAN A., FIORE-GARTLAND B., TORY M.: Passing the data baton: A retrospective analysis on data science work and workers. *IEEE Transactions on Visualization and Computer Graphics 27*, 2 (2020), 1860–1870. 2, 7

[CGJ*17] CHEN M., GRINSTEIN G., JOHNSON C. R., KENNEDY J., TORY M.: Pathways for theoretical advances in visualization. *IEEE computer graphics and applications 37*, 4 (2017), 103–112. 2

[CGJM18] CHEN M., GAITHER K., JOHN N. W., MCCANN B.: An information-theoretic approach to the cost-benefit analysis of visualization in virtual environments. *IEEE transactions on visualization and computer graphics 25*, 1 (2018), 32–42. 2

[Cha06] CHARMAZ K.: *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*. SAGE Publications, London, 2006. 3

[CJ10] CHEN M., JÄNICKE H.: An information-theoretic framework for visualization. *IEEE Trans. on Visualization & Computer Graphics 16*, 6 (2010), 1206–1215. 7

[CJS*21] CAKMAK E., JACKLE D., SCHRECK T., KEIM D., FUCHS J.: Multiscale visualization: A structured literature analysis. *IEEE Transactions on Visualization and Computer Graphics* (2021). 7

[Cre02] CRESWELL J. W.: *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 2nd ed. SAGE Publications, Inc., 2002. 2, 3

[DAREA*18] DIEHL A., ABDUL-RAHMAN A., EL-ASSADY M., BACH B., KEIM D., CHEN M.: VisGuides: A forum for discussing visualization guidelines. In *Proc. EuroVis Short Papers* (2018). 3, 8

[Den10] DENZIN N. K.: Grounded theory and the politics of interpretation. In *The SAGE Handbook of Grounded Theory*. SAGE Publications Ltd., 2010, pp. 454–471. 3, 5

[DFTW*21] DIEHL A., FIRAT E. E., TORSNEY-WEIR T., ABDUL-RAHMAN A., BACH B., LARAMEE R., PAJAROLA R., CHEN M.: VisGuided: A Community-driven Approach for Education in Visualization. In *Eurographics 2021 - Education Papers* (2021), Sousa Santos B., Domik G., (Eds.), The Eurographics Association. doi:10.2312/eged.20211003. 8

[DKAR*20] DIEHL A., KRAUS M., ABDUL-RAHMAN A., EL-ASSADY M., BACH B., LARAMEE R. S., KEIM D., CHEN M.: Studying visualization guidelines according to grounded theory. arXiv 2010.09040, 2020. 3, 7, 8, 9, 10

[DP19] DIMARA E., PERIN C.: What is interaction for data visualization? *IEEE transactions on visualization and computer graphics 26*, 1 (2019), 119–129. 2

[EAJS*19] EL-ASSADY M., JENTNER W., SPERRLE F., SEVASTJANOVA R., HAUTLI-JANISZ A., BUTT M., KEIM D.: lingvis.io - a linguistic visual analytics framework. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (Florence, Italy, July 2019), Association for Computational Linguistics, pp. 13–18. URL: https://aclanthology.org/P19-3003, doi:10.18653/v1/P19-3003. 10

[eas17] EAST101: Action without interaction. http://visguides.org/t/action-without-interaction/70, 2017. VisGuides. 10

[EASD*19] EL-ASSADY M., SPERRLE F., DEUSSEN O., KEIM D., COLLINS C.: Visual Analytics for Topic Model Optimization based on User-Steerable Speculative Execution. *IEEE Trans. on Visualization and Computer Graphics 25*, 1 (2019), 374–384. 10

[EASG*17] EL-ASSADY M., SEVASTJANOVA R., GIPP B., KEIM D., COLLINS C.: NEREx: Named-Entity Relationship Exploration in Multi-Party Conversations. *Computer Graphics Forum 36*, 3 (2017), 213–225. 10

[Elm] ELMQVIST N.: 3d visualization for nonspatial data: Guidelines and challenges. https://sites.umiacs.umd.edu/elm/2017/10/03/. accessed in March 2019. 10

[EY15] ELMQVIST N., YI J. S.: Patterns for visualization evaluation. *Information Visualization 14*, 3 (2015), 250–269. 2

[GREA15] GOLD V., ROHRDANTZ C., EL-ASSADY M.: Exploratory Text Analysis using Lexical Episode Plots. In *Proc. of Eurographics Conf. on Visualization, EuroVis Short Papers* (2015), pp. 85–89. 9, 10

[Grö08] GRÖLLER E.: Visualization - I see it my way (Action without Interaction). In *Dagstuhl Seminar 07291 Abstracts Collection Scientic Visualization*, Ebert D. S., Hagen H., Joy K. I., Keim D. A., (Eds.). 2008. URL: http://drops.dagstuhl.de/opus/volltexte/2008/1414/pdf/07291_abstracts_collection.1414.pdf. 10

[GS67] GLASER B. G., STRAUSS A. L.: *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Aldine, Chicago, 1967. 1, 3, 5, 8

[HBH*19] HALL K. W., BRADLEY A. J., HINRICHS U., HURON S., WOOD J., COLLINS C., CARPENDALE S.: Design by immersion: A transdisciplinary approach to problem-driven visualizations. *IEEE transactions on visualization and computer graphics 26*, 1 (2019), 109–118. 2, 7

[HP92] HENWOOD K. L., PIDGEON N. F.: Qualitative research and psychology theorizing. *British Journal of Psychology 83* (1992), 97–111. 9

[HP03] HENWOOD K. L., PIDGEON N. F.: Grounded theory in psychological research. In *Qualitative Research in Psychology: Expanding Perspectives in Methodology and Design*, Camic P. M., Rhodes J. E., Yardley L., (Eds.). APA Publications, 2003, pp. 131–155. 9

[HQC*18] HULLMAN J., QIAO X., CORRELL M., KALE A., KAY M.: In pursuit of error: A survey of uncertainty visualization evaluation. *IEEE transactions on visualization and computer graphics 25*, 1 (2018), 903–913. 7

[IIC*13] ISENBERG T., ISENBERG P., CHEN J., SEDLMAIR M., MÖLLER T.: A systematic review on the practice of evaluating visualization. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (2013), 2818–2827. 2

[IIS*16] ISENBERG P., ISENBERG T., SEDLMAIR M., CHEN J., MÖLLER T.: Visualization as seen through its research paper keywords. *IEEE Transactions on Visualization and Computer Graphics 23*, 1 (2016), 771–780. 7

[IZCC08] ISENBERG P., ZUK T., COLLINS C., CARPENDALE S.: Grounded evaluation of information visualizations. In *Proceedings of the 2008 Workshop on BEyond time and errors: novel evaLuation methods for Information Visualization* (2008), pp. 1–8. 2

[JMT10] JOHNSON R. B., MCGOWAN M. W., TURNER L. A.: Grounded theory in practice: Is it inherently a mixed method? *Research in the School 17*, 2 (2010), 65–78. 3

[Joh04] JOHNSON C. R.: Top scientific visualization research problems. *IEEE Computer Graphics and Applications 24*, 4 (2004), 13–17. 2

[KAF*08] KEIM D. A., ANDRIENKO G. L., FEKETE J., GÖRG C., KOHLHAMMER J., MELANÇON G.: Visual analytics: Definition, process, and challenges. In *Information Visualization - Human-Centered Issues and Perspectives*. 2008, p. 154–175. 2, 5

[KARC15] KANJANABOSE R., ABDUL-RAHMAN A., CHEN M.: A multi-task comparative study on scatter plots and parallel coordinates plots. *Computer Graphics Forum 34*, 3 (2015), 261–270. 5

[KARC17] KIJMONGKOLCHAI N., ABDUL-RAHMAN A., CHEN M.: Empirically measuring soft knowledge in visualization. *Computer Graphics Forum 36*, 3 (2017), 73–85. 4

[KC06] KNIGGE L., COPE M.: Grounded visualization: integrating the analysis of qualitative and quantitative data through grounded theory and visualization. *Environment and planning A 38*, 11 (2006), 2021–2037. 2

[KF14] KENNY M., FOURIE R.: Tracing the history of grounded theory methodology: From formation to fragmentation. *The Qualitative Report 19*, 52 (2014), Article 103. 3

[KJRK21] KIM N., JOYNER S., RIEGELHUTH A., KIM Y.: Accessible visualization: Design space, opportunities, and challenges. In *Computer Graphics Forum* (2021), vol. 40, Wiley Online Library, pp. 173–188. 7

[KK17] KERRACHER N., KENNEDY J.: Constructing and evaluating visualisation task classifications: Process and considerations. In *Computer Graphics Forum* (2017), vol. 36, Wiley Online Library, pp. 47–59. 2

[KKEM10] KEIM D., KOHLHAMMER J., ELLIS G., MANSMANN F.: Mastering the information age: solving problems with visual analytics. Goslar: Eurographics Association, 2010. 2

[KL16] KANDOGAN E., LEE H.: A grounded theory study on the language of data visualization principles and guidelines. *Electronic Imaging 2016*, 16 (2016), 1–9. 2, 7, 8

[KZZM12] KUANG X., ZHANG H., ZHAO S., MCGUFFIN M. J.: Tracing tuples across dimensions: A comparison of scatterplots and parallel coordinate plots. *Computer Graphics Forum 31*, 3 (2012), 1365–1374. 4, 5

[LAB*21] LIU X., ALHARBI M., BEST J., CHEN J., DIEHL A., FIRAT E., REES D., WANG Q., LARAMEE R. S.: Visualization resources: A starting point. In *2021 25th International Conference Information Visualisation (IV)* (2021), IEEE, pp. 160–169. 7

[LBI*11] LAM H., BERTINI E., ISENBERG P., PLAISANT C., CARPENDALE S.: Empirical studies in information visualization: Seven scenarios. *IEEE transactions on visualization and computer graphics 18*, 9 (2011), 1520–1536. 2, 7

[LKH*15] LEE S., KIM S.-H., HUNG Y.-H., LAM H., KANG Y.-A., YI J. S.: How do people make sense of unfamiliar visualizations?: A grounded model of novice's information visualization sensemaking. *IEEE transactions on visualization and computer graphics 22*, 1 (2015), 499–508. 2, 7

[LLB20] LEÓN G. M., LISCHKA M., BREITER A.: Mapping the global south: Equal-area projections for choropleth maps. In *2020 IEEE Visualization Conference (VIS)* (2020), IEEE, pp. 91–95. 7

[LS21] LUNDGARD A., SATYANARAYAN A.: Accessible visualization via natural language descriptions: A four-level model of semantic content. *IEEE transactions on visualization and computer graphics* (2021). 3, 7

[Mar17] MARAI G. E.: Activity-centered domain characterization for problem-driven scientific visualization. *IEEE transactions on visualization and computer graphics 24*, 1 (2017), 913–922. 2

[mat18] MATTHIAS.KRAUS: Don't replicate the real world. http://visguides.org/t/dont-do-replicate-the-real-world-in-vr/125, 2018. VisGuides. 10

[MD19] MEYER M., DYKES J.: Criteria for rigor in visualization design study. *IEEE transactions on visualization and computer graphics 26*, 1 (2019), 87–97. 2, 7

[MSQM15] MEYER M., SEDLMAIR M., QUINAN P. S., MUNZNER T.: The nested blocks and guidelines model. *Information Visualization 14*, 3 (2015), 234–249. 2

[Mun09] MUNZNER T.: A nested model for visualization design and validation. *IEEE transactions on visualization and computer graphics 15*, 6 (2009), 921–928. 2, 8

[Par21] PARSONS P.: Understanding data visualization design practice. *IEEE Transactions on Visualization and Computer Graphics* (2021). 2, 7

[PGBC20] PARSONS P., GRAY C. M., BAIGELENOV A., CARR I.: Design judgment in data visualization practice. In *2020 IEEE Visualization Conference (VIS)* (2020), IEEE, pp. 176–180. 2

[RAW*16] RIND A., AIGNER W., WAGNER M., MIKSCH S., LAM-MARSCH T.: Task cube: A three-dimensional conceptual space of user tasks in visualization design and evaluation. *Information Visualization 15*, 4 (2016), 288–300. 2

[RDI*19] ROGOWITZ B., DIEHL A., ISENBERG P., BORGO R., ABDUL-RAHMAN A.: Marshalling the many facets of diversity. In *Diversity in Visualization*. Morgan Claypool, 2019. 7

[RPOC20] RAJABIYAZDI F., PERIN C., OEHLBERG L., CARPENDALE S.: Exploring the design of patient-generated data visualizations. In *Graphics Interface Conference (GI)* (2020), Canadian Information Processing Society. 7

[SBLC17] SHER V., BEMIS K. G., LICCARDI I., CHEN M.: An empirical study on the reliability of perceiving correlation indices using scatterplots. *Computer Graphics Forum 36*, 3 (2017), 61–72. 5

[SC90] STRAUSS A. L., CORBIN J.: *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. SAGE Publications, Newbury Park, CA, 1990. 3, 6

[SC98] STRAUSS A. L., CORBIN J.: *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, 2nd ed. Sage, 1998. 10

[SCK18] STREEB D., CHEN M., KEIM D. A.: The biases of thinking fast and thinking slow. In *Cognitive Biases in Visualizations*. Springer, 2018, pp. 97–107. 2

[SEAG*21] SPERRLE F., EL-ASSADY M., GUO G., BORGO R., CHAU D. H., ENDERT A., KEIM D.: A survey of human-centered evaluations in human-centered machine learning. *Computer Graphics Forum* (2021). 7, 8

[SEAKC21a] STREEB D., EL-ASSADY M., KEIM D. A., CHEN M.: Why visualize? arguments for visual support in decision making. *IEEE Computer Graphics and Applications 41*, 2 (2021), 17–22. 2

[SEAKC21b] STREEB D., EL-ASSADY M., KEIM D. A., CHEN M.: Why visualize? untangling a large network of arguments. *IEEE Transac-tions on Visualization and Computer Graphics 27*, 3 (2021), 2220–2236. doi:10.1109/TVCG.2019.2940026. 2, 7

[SES16] SAKET B., ENDERT A., STASKO J.: Beyond usability and performance: A review of user experience-focused evaluations in visualization. In *Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization* (2016), pp. 133–142. 2

[SMM12] SEDLMAIR M., MEYER M., MUNZNER T.: Design study methodology: Reflections from the trenches and the stacks. *IEEE transactions on visualization and computer graphics 18*, 12 (2012), 2431–2440. 2

[TC05] THOMAS J. J., COOK K. A. (Eds.): *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. NVAC, 2005. 2

[TDI21] TOVANICH N., DRAGICEVIC P., ISENBERG P.: Gender in 30 years of ieee visualization. *IEEE Transactions on Visualization and Computer Graphics* (2021), 1–1. doi:10.1109/TVCG.2021.3114787. 7

[TM04] TORY M., MOLLER T.: Rethinking visualization: A high-level taxonomy. In *IEEE symposium on information visualization* (2004), IEEE, pp. 151–158. 2

[TNMC04] TODD Z., NERLICH B., MCKEOWN S., CLARKE D. D. (Eds.): *Mixing Methods in Psychology: The Integration of Qualitative and Quantitative Methods in Theory and Practice*. Routledge, 2004. 3

[Tor14] TORY M.: User studies in visualization: A reflection on methods. In *Handbook of Human Centric Visualization*. Springer, 2014, pp. 411–426. 2

[VW05] VAN WIJK J. J.: The value of visualization. In *VIS 05. IEEE Visualization, 2005.* (2005), IEEE, pp. 79–86. 2, 8

[Wil13] WILLIG C.: *Introducing Qualitative Research in Psychology*, 3rd ed. Open University Press, 2013. 3, 4