A FRAMEWORK FOR CREATIVE VISUALIZATION-OPPORTUNITIES WORKSHOPS

by

Ethan Kerzner

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Computer Science

School of Computing

The University of Utah

May 2019

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Ethan Kerzner**

has been approved by the following supervisory committee members:

Miriah Meyer,	Chair(s)	**August 9, 2018** Date Approved
Erik Brunvand,	Member	**August 9, 2018** Date Approved
Jason Dykes,	Member	Date Approved
Charles Hansen,	Member	**September 13, 2018** Date Approved
Alexander Lex,	Member	**September 14, 2018** Date Approved

by **Ross Whitaker**, Chair/Dean of

the Department/College/School of **School of Computing**

and by **David B. Kieda**, Dean of The Graduate School.

ABSTRACT

Applied visualization researchers often work closely with domain collaborators to explore new, useful, and interesting visualization applications. 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 workshops to accelerate the early stages of applied work, eliciting a wealth of requirements in a few days of focused work. Yet, no guidance exists for how to use such workshops effectively. This dissertation's primary contribution is a framework — created through the meta-analysis of 17 workshops in 10 visualization contexts — that describes how and why to use workshops in the early, formative stages of applied work. The framework: 1) describes characteristics of effective workshops; 2) identifies a process model for using workshops; 3) describes a structure of what happens within effective workshops; 4) recommends 25 actionable guidelines for future workshops; and 5) proposes three example workshops as a starting point for researchers who are interested in using workshops in future projects. The creation of this framework exemplifies the use of critical reflection to learn about visualization in practice from diverse studies and experience.

The framework is grounded in two formative design studies that provide this dissertation's secondary contributions. In the first formative design study, we worked with defense analysts focused on improving the safety of military vehicles. From this design study, we contribute task analysis, data abstraction, and a validated visualization tool for the visual analysis of spatial and nonspatial ballistic vulnerability data. In the second formative design study, we worked with neuroscientists focused on retinal connectomics. From this design study, we contribute two new visualization techniques and a prototype system for visualizing connectivity in large graphs. Our experiences in these two design studies motivated and informed this dissertation's primary contribution.

To D, L, J, M, and A.

CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to extend my deepest gratitude to my advisor and supervisory committee: Miriah Meyer, Erik Brunvand, Jason Dykes, Charles Hansen, and Alexander Lex. Miriah welcomed me while I was an academic orphan and — with great patience — taught me how to be an effective visualization designer and researcher. Working with her and my committee has been a great honor.

I would like to thank the many people who contributed to this dissertation as coauthors, colleagues, and collaborators: Jefferson Amstutz, James Anderson, Rick zum Brunnen, Lee A. Butler, Mark Butkiewicz, Susan Coates, Jason Dykes, Christiaan Gribble, Sarah Goodwin, Cletus Hunt, James Hunt, Bryan Jones, Sara Jones, Alexander Lex, Robert Marc, Rebecca Pfiefer, Jodi Robertson, Matthew Rothwell, Nathaniel Nelson, Marco Salvi, Scott Shaw, Crystal Sigulinsky, Carolyn Stancoff, Timothy Urness, and Chris Wyman.

I am also grateful for the feedback and support of the members of the Vis Design Lab at the University of Utah: Alex Bigelow, Aspen Hopkins, Sean McKenna, Nina McCurdy, Jimmy Moore, Carolina Nobre, Jennifer Rogers, Sam Quinan, and Kris Zygmunt. Special thanks to Christine Pickett because her feedback helped me to write and to think clearly about this dissertation.

I am grateful to my parents and brother for their encouragement and support throughout my education. I would also like to thank my friends and extended family who — at the time of writing this — are in Boston, Chicago, Des Moines, Edina, Fremont, Grenoble, Iowa City, Lyon, Madison, Minneapolis, Portland, Salt Lake City, San Francisco, Santa Clara, St. Louis, St. Louis Park, Stillwater, and Redmond. You know who you are and how you have helped me. Thank you.

This dissertation extends published work for which we have permission to re-use. Reprinted [1], with permission, Kerzner, Ethan and Butler, Lee A. and Hansen, Charles and Meyer, Miriah; "A shot at visual vulnerability analysis;" *Computer Graphics Forum*; © 2015 John Wiley and Sons. Reprinted [2], with permission, Kerzner, Ethan and Lex, Alexander, and Sigulinsky, Crystal and Urness, Timothy and Jones, Bryan and Marc, Robert E. and Meyer, Miriah; "Graffinity: visualizing connectivity in large graphs;" *Computer Graphics Forum*; © 2017 John Wiley and Sons. Reprinted [3], with permission, from Kerzner, Ethan and Goodwin, Sarah and Jones, Sara and Dykes, Jason, and Meyer, Miriah; "A framework for creative visualization-opportunities workshops;" *Transactions on Visualization and Computer Graphics*, to be published; © 2019 IEEE.

This dissertation was supported in part by a cooperative agreement with the US Army Research Laboratory and NSF grant IIS-1350896. The work reprinted in this dissertation was also supported in part by NIH U01 CA198935, the DoD - Office of Economic Adjustment (OEA) ST1605-16-01, NIH EY015128, EY02576, and EY014800 Vision Core, an unrestricted grant from Research to Prevent Blindness to the Moran Eye Center.

CHAPTER 1

INTRODUCTION

This dissertation introduces a framework that describes how and why to use workshops in the early stages of applied visualization collaborations. Existing visualization process models recommend that researchers use interviews, observations, and similar user-centered design methods to explore visualization opportunities, but this process can take months [4]–[6]. In the formative work of this dissertation, we applied user-centered design methods in collaboration with defense analysts, but we struggled to piece together an understanding of visualization opportunities from the competing needs of project stakeholders [1]. After reflecting on this experience, in a subsequent design study we used a workshop to explore opportunities for visualizing neuroscience data [2]. Despite our successful use of a workshop and the reported success of workshops in the literature [7]–[9], there was no guidance about how to use workshops in the context of applied visualization. This dissertation contributes the first framework for using workshops in applied visualization [3]. It focuses on a specific type of workshop, a creative visualization-opportunities workshop, that can elicit a wealth of visualization opportunities in just a few days of focused work. The framework is grounded in our formative design studies and created from the meta-analysis of 17 workshops in 10 visualization contexts. Its creation demonstrates the use of critical reflection to learn about visualization in practice from diverse experiences.

1.1 Motivation and Overview

In applied visualization, researchers work closely with domain collaborators to create new and interesting visualization applications. Through a design process that is often messy and iterative [10], researchers and collaborators establish a deep and changing understanding of domain challenges and explore relevant visualization solutions [11]. Often, researchers work with a number of collaborators in parallel and must piece together a holistic understanding from disparate perspectives [12]. Existing visualization processes recommend interviews and observations to sample domain challenges, but as noted, these methods are time consuming [5]. In recent years, researchers have applied user-centered design methods — repeatable and modular processes that actively engage users in the design process — to complement traditional interviews and observations [6], [13].

In the formative work of this dissertation, we applied user-centered methods — visualization awareness [14], brainstorming [15], and others [16] — to discern how defense analysts think about ballistic simulation data [1]. In an ad hoc process of on-site meetings and video conferences, we pieced together ideas about how our collaborators produced and analyzed data. Although we developed a useful visualization tool based on this knowledge, reflecting on our experience revealed that we invested significant time in sampling the needs of various stakeholders and building consensus instead of building visualizations. Furthermore, we designed visualizations to aid in the day-to-day analysis instead of creating more audacious, potentially transformative, visualizations. Similar challenges plague applied collaborations in large organizations as researchers must piece together the perspectives of highly specialized analysts to understand organizational goals and data analysis pipelines [17]. Likewise, collaborators often may not know what is possible with visualization, reducing their ability to participate effectively in user-centered design methods [18]. Even with methods that show collaborators the possibilities of visualization, such as visualization awareness [14], exploring the opportunities for a project can still require a significant amount of time [1].

In a subsequent design study, we worked with collaborators who, again, had seemingly disparate analysis challenges [2]. In collaboration with neuroscientists, we considered applying user-centered design methods, but we decided that it would likely require a prohibitive amount of time because we would have to sample and integrate our collaborators' diverse perspectives to discover visualization opportunities. We therefore applied a workshop — based on workshops described in the visualization literature [7]–[9] — to expose shared analysis needs and explore ideas beyond the daily analysis challenges. Our workshop elicited a wealth of opportunities for our collaboration and provided a forum for diverse stakeholders to contribute to the direction of our project. Based on the opportunities that we discovered in the workshop, we created new and useful visualization tools. For example, we developed prototypes that enabled the discovery of new circuity in the retina [19] as well as tools and techniques for visualizing graph connectivity [2]. We speculate that the workshop replaced months of interviews with a few days of focused preparation, execution, and analysis. It also helped establish trust and rapport among stakeholders, as one collaborator told us: "The interpersonal leveling and intense revisiting of concepts made more progress in a day than we make in a year of lab meetings ... [the workshop] created consensus by exposing shared user needs."

Yet, our success was serendipitous. Reflecting on our experience revealed that we did not understand how or why the workshop was effective. We haphazardly adapted a workshop in the literature to our project, but we did not know how to run successful workshops in the future. Because of the usefulness of the workshop, we wanted to understand how to make them a repeatable and dependable method for applied visualization research. We therefore gathered researchers who had been involved with a number of workshops in applied visualization to reflect on our collective experience [7]–[9]. At first our discussions were focused on a seemingly simple question about two workshops [2], [8]: what could we do better next time?

To answer this question, we evaluated the methods used in our workshops and identified potential pitfalls and improvements for future workshops. Although our piecemeal analysis seemed useful for researchers who had experience with workshops, it lacked a conceptual framework that holistically described how and why to use them. Because the workshops in visualization have generally been based on *software requirements workshops* [20] and *creative problem-solving workshops* [15], we searched the literature from these fields as well additional workshop resources [21]–[23] for guidance that could be directly adopted in visualization research. Existing resources, however, did not appropriately emphasize three characteristics that are fundamental to applied visualization, which we term *visualization specifics*: the *visualization mindset* of researchers and collaborators characterized by a symbiotic collaboration [5] and a deep and changing understanding of the domain challenges and relevant visualizations [11]; the connection to *visualization methodologies*, including specialized process and design decision models [5], [24]; and the use of *visualization methods* that focus the workshop participants on data analysis challenges and visualization opportunities [8].

The core contribution of this dissertation is a framework that provides the first actionable guidance for workshops in light of the visualization specifics. We created the framework to describe workshops used in the early, formative stages of applied work. We term these workshops **creative visualization-opportunities (CVO) workshops**, because, when used effectively, they encourage creative thinking about visualization opportunities and support the cross-pollination of ideas among all project stakeholders.

The framework results from a 2-year international collaboration of five visualization and creativity experts. By applying a methodology of *critically reflective practice* [25]–[28], we intertwined analysis with action, reflecting on our experiences while continuing to use and analyze workshops [29]–[32]. In sum, we analyzed 17 workshops in 10 visualization contexts [2], [7]–[9], [29]–[34] and surveyed the existing literature of visualization [5],[6],[10],[14],[17],[24],[35], design [16],[36]–[40], software engineering [20],[41]–[46], and creative problem-solving [15], [22], [23], [47], [48]. Through our reflective analysis, we developed a framework that provides the first guidance for using CVO workshops in applied research. The framework articulates what we have learned from using workshops in applied collaborations over the past 10 years and provides a starting point for researchers who want to use workshops in their own projects.

Furthermore, the way in which we created the framework exemplifies how we can use critical reflection to learn about visualization in practice from diverse experiences. However, due to the nature of reflection, the framework does not provide predictive or causal knowledge [28]. Instead, it contributes an interpretive understanding and approach to practice [49]. We intend for researchers to carefully adapt the framework to local context, preference, and experience. In short, the framework is a thinking tool to help researchers plan, run, and analyze workshops as they work with collaborators to create new and useful visualizations.

1.2 Contributions

This dissertation's primary contribution is a framework for CVO workshops in applied visualization research. The framework consists of 1) common factors that influence workshop effectiveness; 2) a process model that identifies actions before, during, and after workshops; 3) a workshop structure that describes what happens in the beginning, in the middle, and at the end of effective workshops; 4) a set of 25 actionable guidelines for future workshops; and 5) three example CVO workshops and set of example methods for future workshops.

The secondary contributions of this dissertation arise from two formative design studies that grounded our creation of the CVO workshop framework. First, in a collaboration with defense analysts, we applied user-centered design methods to understand the needs of disparate stakeholders. This project contributes a task analysis, data abstraction, and a visualization tool for analyzing spatial and nonspatial ballistic simulation data [1]. Second, in a design study with neuroscientists, we used a CVO workshop to explore shared visualization needs of highly specialized researchers. This project contributes a set of software requirements for multivariate graph analysis, two techniques for visualizing graph connectivity, and an open-source implementation of those techniques [2], [19]. Our successful use of a CVO workshop with neuroscientists inspired us to create the framework for CVO workshops.

1.3 Organization

This dissertation summarizes related work to CVO workshops and then presents our work in the order that we completed the formative design studies and reflective analysis. In Chapter 2, we adopt definitions for workshops and creative workshops, connect CVO workshops to existing visualization process models, compare CVO workshops to existing methods for visualization design, and relate CVO workshops to those used in the fields of software engineering and design. In Chapter 3, we describe our design study with defense analysts, which we preface with a description of how it was formative to this dissertation. Similarly, in Chapter 4, we present our design study with neuroscientists, which, again, we preface with its formative aspects. In Chapter 5, we introduce the CVO workshop framework, including a workshop process model and structure, guidelines for effective workshops, and three example workshops. Then, in Chapter 6, we discuss implications and limitations of the CVO workshop framework. In Chapter 7, we conclude by summarizing this dissertation and identifying areas of future work.

CHAPTER 2

BACKGROUND AND RELATED WORK

In this chapter, we summarize the use of workshops in applied visualization and then characterize CVO workshops by their role in visualization design models. Next, we survey the use of visualization design methods that are similar to CVO workshops. After that, we frame CVO workshops more generally — as design methods — and relate them to work in human-computer interaction and design. Lastly, we delve into the origin of CVO workshops, which is grounded in software requirements engineering, creative problem-solving, and creativity support. We preface this chapter with terminology for describing workshops in applied visualization.

2.1 Terminology

Workshop is an overloaded term because workshops are used in practically every field, including technology, education, business, and the arts [21],[48],[50]–[53]. In education [21], a workshop can be defined as a "short-term learning experience that encourages active, experiential learning and uses a variety of learning activities to meet the needs of diverse learners." In a business setting [51], a workshop can be defined as "a collaborative working session in which a team achieves an agreed goal together." A dictionary characterizes workshops as structured meetings with activities focused on a specific subject [54]. In convergence with existing definitions, we define a **workshop** as a short-term event in which a group of people perform structured methods that are focused on a specific topic. The key differences between workshops and meetings are the use of structured **methods** — well-defined repeatable actions [55] — and the explicit focus on a topic or goal.

Creativity is also an overloaded term, with many historical and cultural implications [56]. A full review of creativity research is outside the scope of this dissertation, but we refer the reader to summaries by Sawyer [57] and Mayer [56]. Because of the complexities associated with creativity, we avoid proposing a complete definition. Instead, we describe characteristics of creativity that are particularly relevant to workshops in applied visualization. Creativity involves the generation of novel and appropriate ideas [56], which often results from series of interconnect mini-insights [57]. Creativity can be fostered by exploring a broad space of ideas [58], providing time to rest between periods of focused work [59], and encouraging open communication and cross-pollination of ideas among groups [52]. These characteristics provide a partial definition of creativity, but they are not necessarily actionable. Nevertheless, they provide a foundation for practitioners who want to teach or encourage creative thinking [60].

We use the term **creative workshop** (or creativity workshop) to refer to workshops that deliberately and explicitly encourage creative thinking [15]. The key distinguishing feature of creative workshops is the use of methods that support the aforementioned characteristics of creativity — generating novel and appropriate ideas, exploring a broad space of ideas, balancing activity with rest, and supporting open communication. Arguably, all workshops are creative workshops to some extent [21]. Regardless, we focus on creative workshops because evidence suggests that the use of methods to explicitly foster creativity can have positive impacts on the interpersonal relationships and ideas generated by visualization collaborations [8].

However, because creativity is a complex sociocultural phenomenon, determining if creative workshops can actually enhance creativity is an open question [60]. Yet, this dissertation is *not* about whether workshops can actually enhance creativity. We analyze creative workshops as reflective practitioners [25],[26],[28] who want to better understand how to use them in the context of applied visualization.

Granted, creative workshops have been extensively studied, and a plethora of resources describe how to use them effectively outside of visualization [15],[21],[48],[50],[51]. Yet, CVO workshops differ from those used in other fields because they explicitly focus on visualization, which implies three **visualization specifics** for effective workshops and workshop guidance:

- Workshops should promote a **visualization mindset** — the set of beliefs and attitudes held by project stakeholders, including an evolving understanding about domain challenges and visualization [5],[11] — that fosters and benefits an exploratory and visual approach to dealing with data while promoting trust and rapport among these stakeholders [61].

- Workshops should contribute to **visualization methodologies** — the research practices of visualization, including process and decision models [6], [24] — by creating artifacts and knowledge useful in the visualization design process.

- Workshops should use **visualization methods** that explicitly focus on data visualization and analysis by exploring visualization opportunities with the appropriate *information location* and *task clarity* [5].

We refer to the three visualization specifics throughout this dissertation. They are based on the careful analysis of workshops used in a number of applied visualization collaborations.

2.2 Workshops in Applied Visualization

To our knowledge, visualization researchers have used and reported on workshops in six applied collaborations. In three of these collaborations, researchers reported on a series of workshops. Dykes et al. [7] described three imagination exercises to explore opportunities for enhancing map legends with visualization. Goodwin et al. [8] built on these experiences, reporting on their collaboration with energy analysts in which they used a series of workshops to discover opportunities for visualization, to develop and iterate on prototypes, and to evaluate the resulting visualizations. Walker et al. [9] also applied three workshops in a collaboration with defense analysts to understand needs, create designs, and evaluate prototypes. Reports of these three projects contain detailed descriptions of their workshop experiences, but do not provide instructions about how others could use workshops in the future.

In three other projects, researchers reported on a single workshop used to jump-start applied collaborations. First, Kerzner et al. [2] used a full-day workshop to understand the analysis needs of neuroscientists. Second, Goodwin et al. [29] applied a full-day workshop to explore visualization opportunities in the field of constraint programming. Third, Nobre et al. [30] used a half-day workshop to elicit requirements from analysts working with psychiatric data. These three projects showed that a single workshop can help researchers rapidly explore visualization opportunities, but they provided no guidance for future workshops.

This dissertation is the first meta-analysis of workshops used in applied visualization. We are interested in learning from workshops used in the context of real visualization projects, working with real collaborators to create visualizations that are useful for real data analysis. We focus on workshops used in the early, formative stages of applied work or as the *first* in a series of workshops. We propose the term CVO workshop to describe these events because they encourage creative thinking and promote exploration of visualization opportunities and constraints.

In general, CVO workshops are a method for pre-design empiricism [62] or an evaluation of work practices [4]. More specifically, workshops can be used to fulfill the *winnow*, *discover*, and *design* stages of the design study methodology's nine-stage process model [5]. Alternatively, they correspond to the *problem domain analysis* in the user-centered visualization design framework described by Koh et at. [14] and the *task analysis* and *design* phases of the human-centered design cycle [10].

To graphically represent where CVO workshops fit in the visualization design process, we use the design activity framework because it is, to our knowledge, the only process model that explicitly characterizes design methods by their motivation and intended outcomes [63]. Fig. 2.1 represents CVO workshops in the design activity framework. CVO workshops fulfill the *understand* and *ideate* design activities, which focus on discovering visualization opportunities and constraints as well as proposing visualization ideas at a high level of abstraction [6]. Because design occurs in a cycle, the *understand* activity can also be framed as an evaluation of deployed systems. Therefore, CVO workshops can also fulfill the *deploy* activity because they evaluate existing tools. Nevertheless, we focus on workshops in the early formative stages of applied work.

CVO workshops can also be described by how they influence design decision models. These workshops create artifacts and knowledge useful for problem characterization, data/task analysis, and initial exploration of the encoding in the nested model for visualization design [24]. Their output can help define blocks, providing a foundation for subsequent design work [64]. For example, Goodwin et al. [8] identified key themes in their workshop output that were used as a starting point for visualization designs.

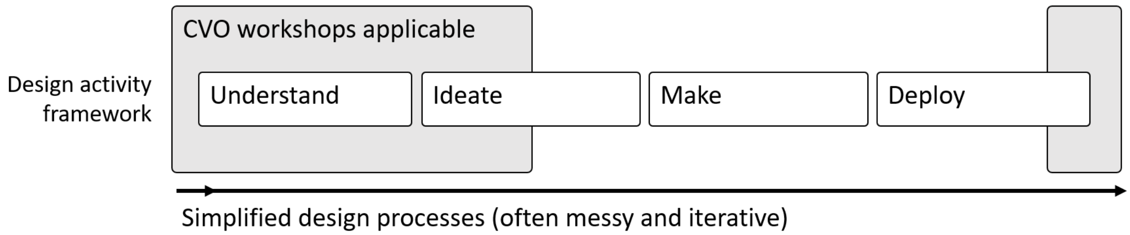

Figure 2.1. CVO workshops in the visualization design process. CVO workshops fulfill the *understand* and *ideate* activities of the design activity framework [6]. Because design often occurs in a cycle, CVO workshops can also be framed as methods for testing and evaluating deployed visualization systems, which, alternatively, is also part of understanding opportunities and analyzing domain tasks. Of course, this description is an abstraction and simplification of the often messy and iterative design processes.

Workshop output can also be used in requirements specification processes, such as the activity-centered framework for problem characterization [65].

CVO workshops have impacts beyond the initial problem characterization. They can help overcome the challenges of working in large organizations by eliciting ideas from diverse collaborators and providing time for stakeholders to step away from daily challenges [17]. Their output can be revisited throughout collaborations, providing criteria to evaluate and validate design decisions. For example, Goodwin et al. [29] refer to the workshop results to evaluate designs for visual analysis of constraint programs. If appropriately documented and preserved, the output can also be used for meta-analysis of workshops — as in this dissertation.

2.3 Visualization Design Methods

CVO workshops are one of many methods that can be used to understand the analysis needs of collaborators in the early stages of applied work. Interviews and observations are the workhorses of applied researchers [5]. Interviews are conversations with the purpose of gathering information [66]. Because of this broad definition, interviews are common across practically every documented design study. Yet, interviews often fail to capture important details because humans struggle to communicate tacit knowledge — we know more than we can say [66], [67].

Accordingly, researchers supplement interviews with observations to examine work in the context in which it occurs [17]. Common forms of observation are contextual inquiry [68] — or contextual interviews [66] — that allow researchers to gather information from collaborators as they perform real work. Contextual inquiry differs from mere observations because it encourages researchers to engage with collaborators, interpreting their reactions and asking clarifying questions as needed [69]. In our experience, contextual inquiry and interviews are invaluable design methods for the early stages of applied work [1].

However, interviews and contextual inquiry are typically used with one researcher and one collaborator [66], [68]. When used in a large organization, they require researchers piece together diverse perspectives to understand analysis needs. They also tend to focus collaborators on ideas to benefit day-to-day work [17]. CVO workshops, in contrast, elicit ideas from a group of collaborators and are designed in a way to encourage thinking about opportunities that are potentially beyond the daily workflow and analysis. Often, CVO workshops complement interviews and observations. The former can be used to explore a breadth of opportunities, and the latter can be used to elicit a depth of information about specific opportunities.

Beyond interview and observations, recent work has reported on user-centered design methods in visualization. Sakai and Aert [70], for example, describe the use of card sorting for domain problem characterization. Hogan et al. [71] propose a form of interview for eliciting subjective perceptions of visualization. Lloyd and Dykes [13], in collaboration with cartographers, evaluate their experiences using a number of methods including lectures, domain scenarios, volare templates, and paper prototypes. Roberts et al. [72] describe a method for exploring and developing visualization ideas through structured sketching. McKenna et al. [73] summarize the use of qualitative coding, personas, and data sketches in collaboration with security analysts. This dissertation provides a framework for thinking about how user-centered design methods can be integrated into CVO workshops.

Additional types of workshops are used in visualization education. Huron et al. [74] describe data physicalization workshops for constructive visualization with novices. He et al. [75] describe workshops for students to think about the relationships between domain problems and visualization designs. Although we frame CVO workshops as a method for experienced researchers to pursue domain problem characterization, methods for visualization education — e.g., constructive visualization [76] and sketching [77] — can be integrated into CVO workshops.

2.4 Participatory (and Similar) Design Methods

More generally, CVO workshops can be framed as a method for user-centered design [78], participatory design [79], or co-design [80] because they involve users directly in the design process. A full survey of these fields is beyond the scope of any one dissertation because their definitions are ill-defined and change over time [81]. We refer the reader to recent summaries of design methods [16], [82] and practices [81], [83], [84] for more complete descriptions of these fields. This section focuses on design methods and frameworks closely related to CVO workshops.

The thinking tools to analyze design methods apply to CVO workshops. Muller [79], for example, describes two orthogonal axes to characterize participatory methods by who participates with whom, and where the methods are used in the design process. Similarly, Sanders et al. [39] characterize design methods by their role in the design process and if they work with local or remote participants. Biskjaer et al. [38] analyze methods based on concrete, conceptual, and design space aspects. Couger et al. [85] differentiate methods by the type of thinking that they are intended to stimulate and evaluate the use of six methods to create business software. These thinking tools provide the terminology to describe the methods of workshops but they do not account for visualization specifics, such as the explicit focus on visualization and data.

Yet, there is one form of participatory design that focuses on data: the CoDesign with Data Framework describes how to incorporate data visualization into participatory workshops to create new products and services [40]. Developed through reflection on experiments that evaluated the use of data visualization in design workshops [86], [87], it recommends principles about using data visualizations in participatory design, such as creating visualizations to display relevant data before the workshop and avoiding ambiguity in any visualizations presented. This framework, however, focuses on using data in the design of products, services, and other initiatives — e.g., case studies illustrating the CoDesign with Data Framework describe workshops used to invent campaigns that encourage recycling or reduce waste on a university campus. Thus, it does not appropriately emphasize the use of workshops in visualization methodologies. Nevertheless, it is grounded in workshops used for creative problem-solving and software engineering, and workshops from these fields also provide a foundation for the workshops used in visualization.

2.5 Creativity Workshops and Support Tools

CVO workshops are based on principles and practices of creativity workshops for software requirements engineering and creative problem-solving [8]. The CVO workshop framework adapts and adopts existing practices from these fields for use in visualization.

Creative requirements-workshops elicit software specifications for large-scale systems from diverse stakeholders [20]. For example, Maiden et al. [43] reported on a workshop to elicit requirements for air traffic control software. Similarly, Jones et al. reported on a workshop to create requirements for e-learning software [42]. A number of other workshops were developed and reported for software requirements engineering [20], [41], [45], [46], [88]. With one exception, these workshops were part of a requirements engineering process that included human activity modeling, system goal modeling, use case modeling, and requirements management [41]. One exception is Hollis and Maiden [89], who reported on the use of creativity methods in an agile process to invent more creative software requirements. We ground the CVO workshop framework in the reported use of software requirements workshops. Software requirements workshops, however, are not connected to visualization methodologies, and their methods mention data only implicitly.

Creative problem-solving workshops are workshops that deliberately and explicitly foster creative thinking for a specific goal [23]. A number of other frameworks describe how to foster creative problem-solving (e.g., Creative Problem Solving [23], Lateral Thinking [48], and Synectics [47]). Surveys of these frameworks reveal that they share the same underlying principles: encouraging open communication, promoting trust and risk taking, providing time for focused work, fostering divergent and convergent thinking, supporting iteration of ideas, emphasizing problem finding and solving, and eliciting group creativity [60]. Existing workshop guidance, however, does not completely describe how to use CVO workshops, because it lacks a focus on data and does not account for the visualization mindset.

No conclusive evidence, however, indicates that any framework can actually enhance creative thinking [60]. Laboratory experiments show that using methods to enhance group creativity — such as brainstorming [15] — often result in fewer ideas than individuals working alone [90]. Yet, critics argue that laboratory experiments rely on contrived metrics and lack ecological validity [91]. Experimentally determining the relationship between workshops and creativity is beyond the scope of this dissertation, as we focus on encouraging creative thinking about applied visualization research with real collaborators in uncontrolled environments.

Just as we aim to encourage creative thinking, creativity support tools aim to encourage creative workflows and processes [37]. Many of the guidelines for creativity support tools apply to CVO workshops. The guidelines include encouraging exploration, providing a low barrier to entry, and promoting open communication for cross-pollination of ideas [36]. Similar to CVO workshops, creativity support tools are often evaluated in real environments where controlled experiments are not possible [91]. Furthermore, frameworks for creating creativity support tools encourage the reflection on and analysis of new ideas [37], which is the goal of this dissertation.

2.6 Conclusion

This dissertation is grounded in the rich history of workshops and related methods used in design, software engineering, and creative problem-solving. It is also based on meta-analysis of workshops used in real applied visualization collaborations conducted on 3 continents over the past 10 years. The CVO workshop framework is the first explicit guidance for how and why to use workshops in the early, formative stages of applied visualization research. It is motivated by and grounded in our experiences in two formative design studies.

CHAPTER 3

FORMATIVE DESIGN STUDY — SHOTVIEWER: VISUAL ANALYSIS OF BALLISTIC VULNERABILITY DATA

This chapter is about a design study in which we applied user-centered design methods to understand how defense analysts reason about ballistic simulation data and ultimately created a new visualization tool for analyzing the simulation results. More specifically, it contributes data and task analysis for the domain of ballistic vulnerability analysis, as well as ShotViewer, a validated visualization tool that uses three linked views for visual analysis of combined spatial and nonspatial ballistic simulation data. Before detailing the design study, we describe how this work influenced the CVO workshop framework.

3.1 Formative Aspects

In this project we worked with a large organization that consisted of specialized analysts working together to achieve organizational goals [17]. We therefore needed to understand how the organization as a whole produces and analyzes data — the perspective of any one analyst was incomplete. Accordingly, we worked with analysts, fellow tool builders, and specialists who produced data as well as those who consumed analysis results. To establish buy-in from all of the stakeholders, we experimented with user-centered design methods, including visualization awareness [14] and brainstorming [15]. While these methods were useful, we applied them in an ad hoc process that required significant energy and time commitment. For example, while visiting our collaborators' organization, we used visualization awareness three times — each with a different group of stakeholders.

Nevertheless, the user-centered design methods proved valuable — visualization awareness helped to engage and excite collaborators with the possibilities of our work and brainstorming provided a forum for stakeholders to contribute ideas to the project. Yet, since we applied the methods in a piecemeal fashion, we also spent significant time piecing together disparate perspectives to understand how the organization uses ballistic simulation data. Furthermore, throughout the design process, we had to manage expectations of diverse stakeholders. To maintain high levels of engagement, we focused on visualizations that could fulfill immediate analysis needs and avoided more audacious, potentially transformative, ideas.

This chapter is formative to this dissertation in that it demonstrates the utility of user-centered design methods even though they may require significant time commitment. In subsequent work (Chapter 4), when we were faced with similar challenges — specialized analysts in a large organization — we applied user-centered design methods as a workshop that ultimately inspired the CVO workshop framework.

3.2 Motivation and Overview

Simulations enable engineers to test designs in ways that may be too expensive or time consuming for the real world. For example, vehicle manufacturers use simulations to understand how a vehicle may perform in extreme conditions, such as during a collision [92], providing insights that can ultimately lead to safer vehicles as engineers revise designs based on the simulation results. Similarly, military organizations use ballistic simulations to understand the vulnerability of their vehicles, allowing them to improve safety by modifying vehicle designs.

In military applications, analyzing the vulnerability of vehicles relies heavily on ballistic simulations [93]. Consuming these simulation results, however, necessitates human insight as analysts must find trends, patterns, and outliers in vehicle vulnerability that require experiential knowledge to spot. Furthermore, the analysts must reason about *multityped data*: output attributes from physics-based simulations, computer-aided design (CAD) models, and hierarchical relationships between the model's components. Analysts use their findings to debug simulation inputs, identify components likely to be damaged, plan live-fire testing, and ultimately make recommendations about vehicle design. The complete analysis process is labor intensive and relies heavily on the expert knowledge of analysts. As one senior vulnerability analyst told us, consuming ballistic simulation results is ”more of an art than a science.”

In this design study, we worked closely with three vulnerability analysts to demystify the process of consuming ballistic simulation results. The collaboration also enabled us to design and deploy a prototype visualization system that leverages multiple linked views for supporting reasoning about spatial and nonspatial vulnerability data. As a result, several novel contributions arise, including a problem characterization, data abstraction, and task analysis for the vulnerability analysis domain, as well as Shotviewer, a carefully justified and validated software prototype for visual vulnerability analysis. Furthermore, reflection on our design process illuminates a strategy for exploiting view-design parallelism while creating multiview visualizations and a list of four recommendations for conducting design studies in large organizations with sensitive data.

3.3 Methods

We describe our methods using the design activity framework [6], a process model for visualization design, as it provides the vocabulary necessary to describe the parallel nature of our design process and maps to the well-known nested model for visualization design [24]. The framework identifies four discrete design activities: **understand** the users' needs; **ideate**, or generate ideas, to support those needs; **make** tangible prototypes of the ideas; and **deploy** visualization prototypes to users. We discuss our methods in the context of these activities after describing our project constraints; namely, developing remotely, designing for sensitive data, and working within a large organization.

The design study was conducted over the course of 15 months. We worked remotely throughout the project as we were two time zones away from our collaborators. We met weekly via video conferences, and spent 3 weeks onsite with analysts. Unless otherwise specified, all the methods were conducted in video conferences, making it challenging to develop rapport as video conferences can be impersonal.

The challenges of working remotely were exacerbated by our collaborators' sensitive data that could not be moved from their secure machines, machines that were not connected to outside networks. We had access to their real data during our onsite visits, but we otherwise used simplified test data, which are shown in all of the figures in this paper.

Our collaborators are employees of a large organization, which is characterized by specialized individuals relying on cooperation to accomplish tasks coupled with decentralized authority over workflows [94], [95]. Thus, we sought input from various individuals who fulfill different roles in the organization. In particular, we spoke with engineers who create input data for the analysts and the analysts' customers who communicate results to the vehicle manufacturers.

For the understand activity, we initially conducted unstructured interviews with one analyst and one fellow tool builder [5] to learn domain vocabulary and to create an initial data abstraction. We refined the abstractions and performed a task analysis through onsite contextual inquiry [68] with three analysts. To engage the analysts and convince them of the potential for visualization to improve their workflow, we conducted three visualization awareness workshops as described by Koh et al. [14].

For both the ideate and make activities, we used parallel prototyping [96] and participatory design [79], but with prototypes of varying fidelity for the two activities. For instance, we used low-fidelity paper prototypes during the ideate activity compared to higher fidelity data sketches [13] in the make activity. To elicit feedback about our prototypes, we used the *rose-bud-thorn method* [97], where individuals are prompted to give three comments on an idea: one positive (rose), one negative (thorn), and one identifying an opportunity for future work (bud). This allowed us to elicit both positive and negative feedback even though our collaborators were initially reluctant to criticize our ideas.

We deployed our final visualization system, Shotviewer, to three analysts who have used it in their daily work. Wide-scale deployment, however, requires integrating it with existing vulnerability analysis software. We discuss the proposed integration in Sec. 3.6.

Although we discuss the design activities linearly here, we performed them in a parallel and staggered fashion as shown by our project's timeline in Fig. 3.1. Our software design consists of three linked views to display spatial and nonspatial data, and each row in the timeline corresponds to per-view activities. For simplicity, we do not show the iterative nature of the design process, which often involves backward movement through activities. Sec. 3.8 contains our strategy and justification for designing multiple linked views in parallel.

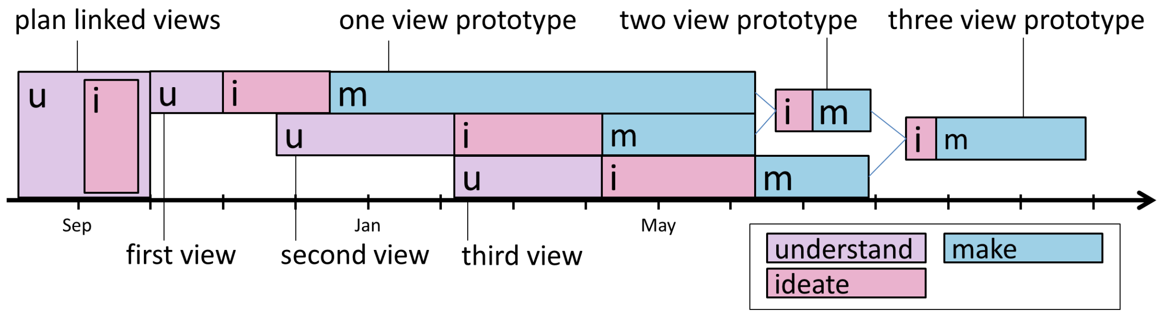

Figure 3.1. Timeline of the Shotviewer design study. After conducting initial understand and ideate activities, we decided on a multiple linked view system. We designed each of these views in parallel and ultimately combined them into a full system.

3.4 Problem, Data, and Tasks

In this section we discuss our problem characterization, data abstraction, and task analysis contributions for the domain of ballistic vulnerability analysis. We characterize the domain problem by providing a high-level overview of ballistic simulations and describing the broad goals of vulnerability analysts. Next, we propose a data abstraction for the simulation, followed by the specific tasks the analysts need to conduct in order to understand the simulation results.

3.4.1 Problem Characterization

Our collaborators focus on understanding ballistic simulation outputs. These outputs are best explained by examining the origin of ballistic simulation software: an optical ray tracer [98],[99]. Ray tracers compute photon paths through an environment, shade surfaces using physically based lighting models, and output pixel color based on primary visibility rays. Ballistic simulations replace photons with **shots**, a projectile being simulated, and compute energy transfer using physically based penetration models. Historically, the simulations output statistical summaries such as the total vulnerable area of a vehicle, but these summaries have a relatively nebulous definition that make it almost impossible to reason about why a vehicle may be vulnerable.

To understand simulation outputs in greater detail, analysts rely on images output from the ballistic simulations, called **cell plots** as shown in Fig. 3.2, where cell color represents vehicle capability damage from a shot for a given trajectory. The number of cell plots and cell size varies between vehicles and analyses. Typically, analysts run simulations for between 3 and 42 industry-standard trajectories around a vehicle, with a cell size ranging from 10-100mm [93]. Unfortunately, consuming cell plots is labor intensive as they are ambiguous: they contain no information about *why* a cell is a certain color. These colors are derived from a variety of information, both spatial and nonspatial, related to a vehicle's three-dimensional (3D) geometry and capabilities. Vulnerability analysts rely on their extensive domain knowledge to intuit about the cause of a cell's color, which they use to identify patterns and outliers in vehicle vulnerability.

3.4.2 Data Abstraction

Fig. 3.3 contains an overview of the simulation data relevant to understanding cell plots. The vehicle inputs are collections of 3D meshes, called **components**, shown in Fig. 3.3(a). A **dependency graph**, shown in Fig. 3.3(b), describes the functionality of the vehicle in terms of its components [100]. The graph's leaves are **critical components**, a subset of the vehicle's components that contributes to its capabilities. The graph's internal nodes are aggregations of components called **systems**, and its roots are aggregations of systems called **capabilities**.

Fig. 3.3(c) shows the simulation launching a shot at the vehicle. For a given trajectory, the simulations compute **shotlines**, the path of a shot through the vehicle. These shotlines contain the physical properties of the shot — quantitative attributes such as mass and velocity defined along the one-dimmensional (1D) line. Shotlines also contain per-component quantitative values that represent the amount of damage to intersected components [93]. Simulations aggregate per-component damage up the dependency graph from components to capabilities in order to compute per-shotline capability damage values. Each cell in a cell plot, which represents a single shotline, is then colormapped to encode the damage value.

These damage values, and resulting cell plots, are an ineffective representation of the simulation output as they aggregate complex spatial and abstract data into a single value. We focus on unpacking and visualizing the information behind cell plots in order to un-

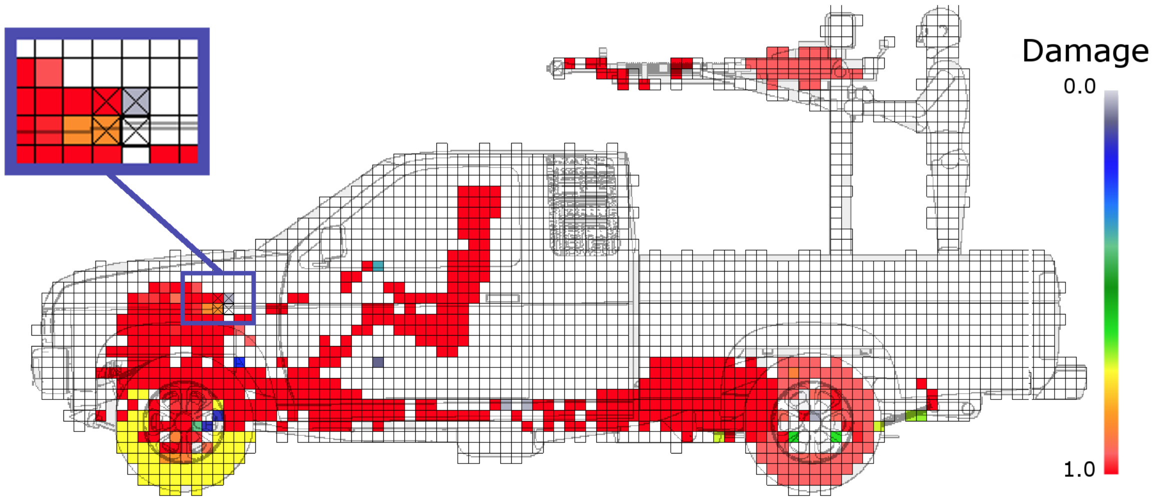

Figure 3.2. Cell plot output by ballistic simulations. Cell color encodes the quantitative damage that each shot inflicts on this vehicle's mobility or firepower capabilities. Analysts understand cell plots by comparing cells to their neighbors of different colors, such as the highlighted cells marked here.

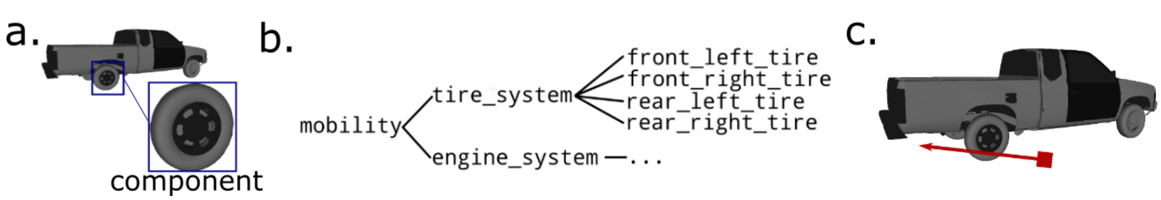

Figure 3.3. Overview of ballistic vulnerability analysis data. It consists of three datatypes: a) vehicles are composed of various 3D meshes called components; b) a subset of these components appears in a dependency graph, which describes the vehicle's capabilities; and c) the simulations trace shots through the vehicle and aggregate per-shot damage using the dependency graph.

derstand the rich and descriptive simulation data.

3.4.3 Task Analysis

The high-level goal of this work is to enable a deeper understanding of the data behind cell plots. Analysts use cell plots to debug simulation inputs and to understand vehicle vulnerability. In both cases, they must explain the color behind an individual cell or a group of cells to check the validity of their inputs or to identify trends in vulnerability. Our task analysis revealed that analysts understand cells in the context of their spatially adjacent neighbors. For instance, the differences between red and white cells highlighted in Fig. 3.2 can be used to explain what components are being damaged by the central group of red cells. To understand differences between shots, analysts perform the following tasks per-shot and compare their outcomes:

T1 Understand a shot's degradation and component damage.

T2 Understand a shot's spatial context.

T3 Understand a shot's systemic impact.

These three tasks require understanding the simulation output, i.e., shots and component damage, in the context of the simulation input, i.e., the vehicle's geometry and dependency graph — we describe these tasks in greater detail through our design requirements in Sec. 3.6. Currently, analysts understand a shot's degradation and component damage (T1) using textual debugging logs. These logs are massive and not particularly human readable — in our test data one of the debugging logs is 2.7Gb in size. Furthermore, analysts trace a shot's systemic impact (T3) by viewing yet another textual description of the dependency graph. Finally, to see a shot's spatial context (T2), the analysts rely on offline rendering software or their own mental models. We postulate that effectively making sense of this multityped simulation data requires a multiview visualization that combines both the abstract two-dimensional (2D) and the spatial 3D data.

3.5 Related Work — Spatial and Nonspatial Data

To our knowledge, no previous research has focused on visualizing ballistic simulation results. Recent work within the vulnerability analysis domain involves accelerating simulations with modern ray tracers [99] and designing software for preparing simulation inputs [101]. Ours is the first project to analyze the consumption of simulation results. Due to the lack of existing visualization software for vulnerability analysis, in this section we discuss the research related to the challenges of our design study — creating 2D representations of inherently spatial data and designing multiview visualizations for multityped data.

Visual comparison of inherently 3D data is made challenging by occlusion, clutter, and difficult navigation. To avoid these issues, we project 3D geometry and shot data into 2D while preserving spatial relationships. This is inspired by Keefe et al. [102], who derive 2D representations of 4D animations with geometry tracers. Similarly, Landge et al. [103] project 3D network structures to a 2D view in order to avoid occlusion while preserving spatial relationships. Also, Weber et al. [104] and Meyer et al. [105] represent embryo cell position using a lossless 2D parameterization of its structure. Recently, Al-Awami et al. [106] have proposed a 2D representation of brain connections that preserves spatial relationships. Similar to these designs, we use a 2D projection of 3D geometry and shot data.

We also use linked multiform 2D and 3D views. Closely related to our work is the SimVis application [107] as it combines spatial 3D and abstract 2D views, although the 2D representations are limited to scatterplots and histograms. Similarly, the GRACE application [108] combines 3D spatial brain data with abstract functionality data, but it does not consider 2D projections of the 3D data. Chang and Collins [109] augment 3D views with 2D views containing summaries of vehicle damage from highway reports, but they do not consider the resulting vehicle functionality from that damage. Design studies on in-car communication networks also combine 2D abstract views with 3D spatial views [110], [111], although they do not apply to the problem of vulnerability analysis.

3.6 Shotviewer

In this section we present Shotviewer, a prototype visualization system that combines spatial and nonspatial data for ballistic vulnerability analysis. Its three linked views, inspired by the multiform encodings of Maries et al. [108], correspond to the tasks of understanding cell plots: the **Shotline View** enables comparison of shot degradation and component damage (T1); the **Geometry View** provides information about a shot's spatial context (T2); and the **System View** displays the shot's systemic impact (T3). We discuss Shotviewer by identifying per-view requirements and using them to justify our design decisions. We preface this with two application-wide requirements:

- *Support current workflow and offer new capabilities.* Analysts read text files to make sense of cell plots, allowing them to compare two shotlines at a time. We created Shotviewer to visualize from one to four shotlines simultaneously. Although it will not scale to more than four shotlines, Shotviewer supports analysts' current workflow and offers new capabilities that we discuss for each view. Shotviewer also incorporates textual data representations where possible to help analysts build trust in the visualizations as they can verify that the encodings match the data.

- *Interface with existing tools.* We designed Shotviewer so that analysts may launch it from their existing software. Existing software [101] displays cell plots to which Shotviewer then provides details-on-demand. We propose that cell plots act as a legend by assigning a categorical color to each cell, as shown in Fig. 3.4. We use these colors to identify shots within our application.

3.6.1 Shotline View

The Shotline View in Fig. 3.5 shows information about the shots' degradation and the damage to intersected components. This information is shown in three subviews: a Table View (top); a Compare View (center); and Line Plot Views (bottom). While the Table View supports existing workflows with textual data representations, we based the other subviews on the following requirements:

- *Show shotlines as linear events.* The Compare View uses a lossless 1D parameterization of shotlines. The shot's vehicle entry point is at the far left and the horizontal position encodes distance from it. We represent shots with straight lines as penetration models do not allow for refraction. This linear display corresponds to reading textual shot descriptions from the entry point through the vehicle.

- *Differentiate between air and components, while preserving thickness of arbitrarily sized components.* We represent air and components along a shotline using a horizontal set

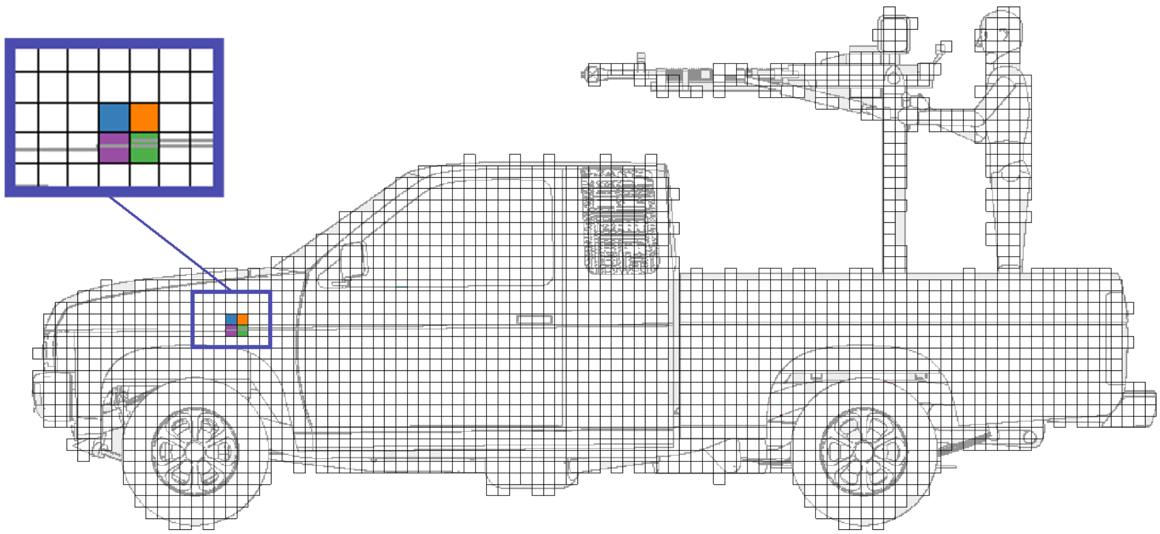

Figure 3.4. Example cell plot as a legend for Shotviewer. We propose that cell plots serve as an interactive legend for Shotviewer. Here, the user selected four shotlines that are assigned categorical colors used within Shotviewer.

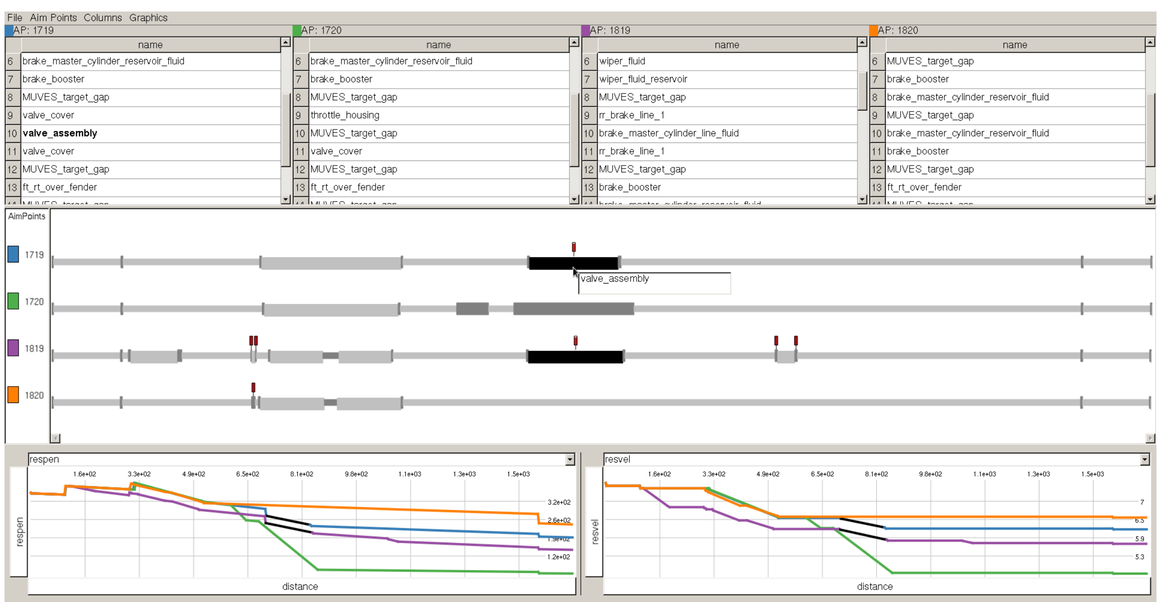

Figure 3.5. Shotline View of Shotviewer. It consists of three subviews. The Table View (top) provides human-readable details about the shot. The Compare View (center) uses a 2D projection of shotlines to enable visual comparison through juxtaposition [112]. The two Line Plots Views (bottom) show trends in shot degradation.

of rectangles. Since analysts focus on component damage, rectangles representing components have larger height than those representing air. Rectangle width encodes the component thickness, which is particularly useful as component thickness is important in physically based ballistic models. Rectangle color has no inherent meaning, but rather alternates to distinguish between components. This is necessitated by the varying width of components, for instance, ranging from an engine block to an electrical wire. Our parameterization and damage representations were inspired by the linear gene representation in Variant View [113].

- *Display damaged components.* The Compare View shows component damage using the height and color of a fixed width glyph above each component. We set the width to a fixed size to show damage of small and large components with equal salience. This encoding is particularly useful in vehicles with complex electrical wiring, which is common in our collaborators' sensitive data.

- *Enable comparison of shot degradation between and within shotlines.* Shot degradation is the change in mass, velocity, and other quantitative values defined on the shotline. We initially considered encoding these in the Compare View, for instance, by using rectangle height, but with such encodings it is difficult to compare values between shotlines. We instead use the two Line Plot Views to display the shots' physical degradation: the x-axis uses the same lossless parameterization as the Compare View and the y-axis encodes the physical properties. These views enable comparison both along and between shotlines, allowing analysts to verify that the physically based simulations work as intended.

- *Identify components hit by multiple shots.* Interaction allows analysts to highlight components that have been shot, which is useful for identifying components hit by more than one shot, such as the valve assembly in Fig. 3.5, which has been hit by two shots. Moving the mouse over a component in any of these subviews highlights it throughout the application in order to support reasoning about the component's spatial location and systemic impact.

3.6.2 Geometry View

The Geometry View in Fig. 3.6 displays shots' 3D position in the context of the vehicle geometry. We designed this view from the following requirements:

- *Represent shots and shot components in 3D.* The Geometry View displays shotlines as 3D cylinders. Similar to the Shotline View, it uses colors to identify shots and cylinder radius to differentiate between air and components. We also render each of the components intersected by the shot.

- *Show spatial context with respect to vehicle geometry.* Rendering just the shotlines and intersected components provides insufficient spatial context about the entire vehicle. Rendering the entire vehicle, however, is impractical due to slow rendering performance and visual clutter. Although we initially considered exploded geometry views [114], they failed to provide information about the true spatial relationships of components needed by our analysts. Instead, we observed the analysts using offline renderers to draw intersected components along with the geometry of known locations such as crew or wheels in order to place the components in their mental model of the vehicle. Based on this observation, we developed the concept of **landmark geometry**, a user-defined set of recognizable components. In Shotviewer we allow analysts to define landmark geometry with regular expressions.

- *Display spatial context with respect to specific systems.* Aliasing leads to unexpected simulation results as shotlines simulated by infinitely thin rays fail to intersect high-frequency geometry, such as wires. Often, analysts will verify that aliasing occurred by looking at a shot with respect to an individual system they expect to be damaged. We support this reasoning by letting users select systems rendered with the shotlines, shot components, and landmark geometry.

- *Customize colors used in 3D rendering.* In their existing tool chains, analysts have little control over color schemes used in 3D rendering. As one analyst complained, the vehicles "look like clown cars." The analysts would often manually color geometry images when presenting results to their peers or customers. Thus, in Shotviewer we allow users to control the geometry color and opacity. Although this feature may seem insignificant, improving the visualization aesthetics significantly improved our rapport with analysts.

3.6.3 System View

The System View in Fig. 3.7 visualizes systemic impact of component damage using a node-link diagram. Although the simulations output a list of damaged components, these lists fail to sufficiently describe the impact of that damage. For example, in Fig. 3.7 the valve assembly has been damaged, and it would be difficult to intuit that this is a part of the engine system without the dependency graph for context. The idea of showing the dependency graph structure appears in our design requirements for the System View:

- *Show only relevant components, systems, and capabilities.* Dependency graphs often contain up to thousands of nodes. Although we initially considered space-efficient representations of these graphs, such as treemaps [115], analysts found these hard to understand. We instead filter the dependency graph to only nodes impacted by selected shots.

- *Clearly encode parent-child relationships and damage.* Analysts must understand the damage propagation through a dependency graph, from the components (leaves) through the systems (internal nodes), and to the capabilities (roots). We use a hierarchical node-link diagram layout (computed with graphviz [116]) as it encodes parent-child relationships both with connections and spatial position. We display damage with the height of a vertical bar to the right of each node, and we identify the shot causing damage with the bar's color. Links represent damage propagating up the tree. Moving the mouse over a node highlights the path to its parents.

- *Enable top-down analysis (from capabilities to components).* Before our design study, analysts could understand systemic impact by aggregating component damage up to the capabilities—in essence, manually aggregating the details to get an overview of the data. In Shotviewer, we additionally allow analysts to select a node of interest and filter the graph from the top-down, showing only damaged nodes that contribute to their selection. This top-down workflow is novel in the vulnerability analysis domain.

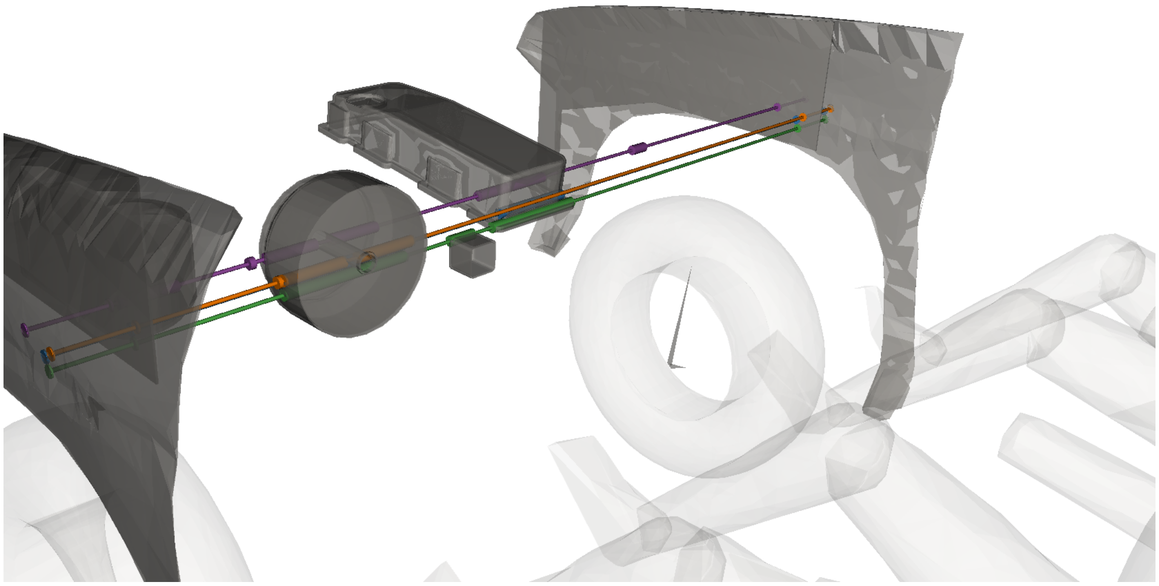

Figure 3.6. Geometry View of Shotviewer. It renders 3D representations of the shotlines along with user-defined landmark geometry to provide spatial context while avoiding clutter and occlusion. Here, the crew and wheels serve as landmark geometry.

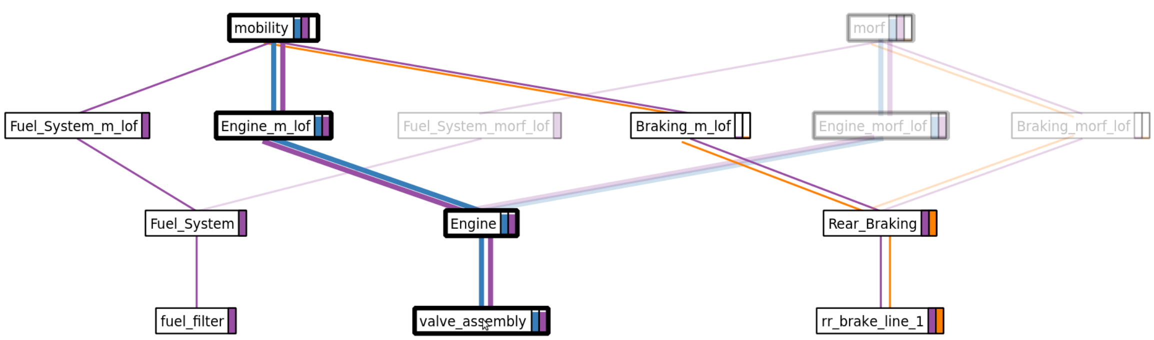

Figure 3.7. System View of Shotviewer. It uses a node-link diagram to show the systemic impact of damaged components. The full dependency graph is filtered to show only damaged components (leaves) and their parents (systems and capabilities). The vertical bars on the right of each node encode damage, and color identifies the shots causing the damage. Using this view, analysts can filter the damaged capabilities to focus on one of interest. Here, highlighting enables the user to trace the valve assembly damage up to the vehicle's mobility.

3.6.4 Implementation

We implemented Shotviewer using C++ and Qt. We focused on rapid prototyping in order to quickly get user feedback and iterate on designs. We also considered the following requirement in our implementation:

- *Develop where we deploy.* We treated the hardware and software available to analysts as constraints. For instance, analysts did not have graphics-processing units, so we used a CPU ray tracing library [117] to render the 3D geometry. We found that virtual machines did not reproduce the behavior of analysts' machines (for example, with windowing libraries that use hardware-accelerated rendering), so we instead developed on real machines that matched their configurations. This approach enabled fast installation and feedback of our prototypes.

3.7 Validation

We validate both our task analysis and Shotviewer's design. Early in the project, we interviewed our analysts' customers to find immediate validation of our choice to support the analysis of cell plots. After completing the design study, we sought informal user-feedback from three vulnerability analysts and two case studies that describe Shotviewer's usefulness.

Talking to our analysts' customers provided initial validation of our problem characterization and task analysis. Through early interviews we began to understand the importance of cell plots as well as how difficult they are to interpret, leading to our task analysis, which focuses on understanding these plots. One of the analysts' customers who is responsible for communicating vulnerability analysis results to vehicle manufacturers validated this analysis. In particular, he said, "We don't actually do a good job of turning cell plots into actionable knowledge" and that "the data behind cell plots are intelligible to only a few people." Creating software to understand cell plots, he confirmed, would be beneficial to the entire vulnerability analysis community.

We received overwhelmingly positive user feedback on Shotviewer. One analyst stated that the tool reduced his time-to-understanding the data behind cell plots from minutes to seconds. We have also seen a growing demand for Shotviewer as our analysts' coworkers have seen them working with it and asked to use it for their own work.

Our first case study occurred while demonstrating a final version of Shotviewer to an analyst. She was examining a dataset from a project that she already completed and shipped to a customer. Although she believed the data to be bug free, Shotviewer enabled her to discover an error in it. Particularly, while looking at a shot with a low capability damage, she noticed a component in the Geometry View that was hit by the shot. Although she expected that component to cause high-capability damage, it did not appear in the System View. After checking the simulation inputs, she discovered an error that caused damage to that component to be incorrectly aggregated.

We present a second case study from an analyst who used Shotviewer to understand simulations of a new shot type against an existing vehicle. After preparing inputs and running the simulation, she inspected a cell plot. The cell plot contained some groups of red cells (high-capability damage) surrounding a green cell (low-capability damage). She launched Shotviewer to see the data behind these cells so that she could explain their differences. Using the Compare View's damage glyphs and interactive highlighting, she identified a component that was damaged by the three red shotlines but not by the green one. She hypothesized that this component caused the high-capability damage and used the System View to confirm this by tracing the path from the leaf up through its root. Using the Geometry View, she was also able to see that the green shotline had passed very close to the component and that it would have likely been damaged in a live-fire test. She concluded that the outlier was a result of aliasing, and the component in question was vulnerable to this particular shot. In her old workflow, a similar analysis would have involved looking up information from three different text files and running an offline rendering application to see the shots' spatial context.

3.8 Reflections

In this section we present the two methodological contributions that result from reflection on our design process. The first is a strategy that exploits **view-design parallelism** for creating multiview visualizations of multityped data. This strategy proposes designing views in a parallel and staggered fashion in order to effectively use time and resources while keeping collaborators engaged in the design process. The second contribution is a list of four recommendations to increase the efficiency of designing visualizations in a large organization with sensitive data.

3.8.1 View-Design Parallelism

View-design parallelism complements the growing area of design study process models [5], [6], [8], [12]. These models focus on the iterative and multilinear nature of design, including the execution of parallel design activities, but they do not directly address the challenges that we as designers face in problem-driven work: potentially overwhelming amounts of data and collaborators who expect tangible prototypes early on in the design process. View-design parallelism is a divide-and-conquer strategy for effectively designing individual views of a multiview visualization system. We present this strategy in the context of our work with Shotviewer, where it enabled us to avoid the paralysis caused by staggering amounts of data and uncountably many design possibilities. It allowed us to quickly deliver tangible prototypes, which refined our understanding of the problem and further engaged our collaborators.

The need for multiview visualizations arises when a single view is not optimal for all necessary tasks [118], such as with ballistic simulation data. In the case of Shotviewer, we knew early on that we would need multiple views to support the three types of data. At first we were overwhelmed while trying to understand all of the data and associated tasks at once. We noticed that by trying to stick to the understand activity for all three data types, our efforts turned counterproductive as we, and our collaborators, became mentally exhausted. To overcome this, we pushed on to the ideate and make activities for one of the data types, while still understanding the other two. This solution injected new energy into the project, gave us tangible results to pass on to our collaborators, and also helped us to better understand the nuances of the data and tasks overall.

We continued to design the three views in a staggered and parallel fashion, a strategy we call view-design parallelism. The key idea of view-design parallelism is to begin with the understand activity for one data type, and then when moving on to the ideate activity, beginning the understand activity for a second data type, and so on — this continues for each data type and individual view. This strategy is shown in our timeline in Fig. 3.1. However, this strategy leaves open the question of how to decide which data type and view to focus on first, second, etc. To answer this, we advocate the use of the following considerations to rank each data type:

1. Availability: do we have the necessary data and support?
2. Usefulness: will our collaborators use the visualization?
3. Impact: will the visualization impact our collaborators?
4. Time: how quickly can we deliver the visualization?

The first item is a constraint: it is impossible to design views without data and support such as parsers [5]. The other items are considerations that may vary between projects. In essence, when we have the opportunity to begin designing a view, we select one that is useful, that will have a large impact, and that can be designed quickly.

While building Shotviewer, we exploited view-design parallelism. We started with the Shotline View as we had access to the shotline data while waiting on access to parsers for the other data types. Next, we built the System View as we believed it would have a significant impact by enabling a new way of looking at the dependency graph through top-down analysis. We concluded by building the Geometry View as it was the most labor intensive in terms of implementation. Throughout our design process, we were able to transfer insights between the design of each view. For instance, in designing the System View we used the encoding of vertical height to show per-component damage values that we had already designed and tested for the Shotline View.

After designing the individual views, we began to think about how to combine them into a multiview system. At this point, we benefited from getting user feedback on the views individually, as well as connected to each other. Through informal experiments, we observed that presenting just single views to users elicited feedback on the individual encodings, while presenting views linked together elicited feedback on the choice of interaction and linking. This observation helped us decide which views to show users in order to focus feedback appropriately.

In summary, view-design parallelism is a strategy for designing visualizations that overcomes the challenges of understanding overwhelming data, of quickly creating tangible prototypes, and of diminishing returns for time spent on individual design activities. We believe it is particularly useful when designing multiview visualizations. A useful metaphor to describe the strategy is to compare visualization design to a reduced instruction set CPU architecture that exploits instruction-level parallelism [119]. The design of each individual view is like a CPU instruction — just as CPU pipelines overlap and stagger instruction execution to efficiently use limited hardware resources (e.g., for instruction fetch, decode, arithmetic, and memory access), view-design parallelism overcomes our limited mental resources (e.g., time and energy dedicated to a single activity). Instruction-level parallelism increases CPU throughput, and similarly we believe view-design parallelism increases the productivity of visualization design.

3.8.2 Recommendations

Data constraints have become increasingly common with problem-driven research — in our case, these constraints arose from working on sensitive data within a large organization. While researchers are sometimes able to obtain surrogate data with similar characteristics to the real data, such as Walker et al. [9] who used online business reviews in place of human-terrain reports, there are domains where no such surrogate data exist, such as in vulnerability analysis. In other cases, metadata can be used to create visualizations without direct data access [120], but metadata do not exist for our domain. Thus, we were forced to work with simplified test data, and we identify the following recommendations based on our experience. These recommendations are an extension of previously identified design study pitfalls [5] and recommendations for evaluating visualizations in large organizations [17]:

- *Sample the relevant data pipeline.* By talking with the *producers* and *consumers* of our analysts' data, we built a more accurate data abstraction. These conversations also allowed us to validate our initial choice of explaining data behind cell plots early in the design process.

- *Recognize test data are not real data.* By talking with our collaborators about how our test data differed from the real data we were able to develop visualizations that were more likely to work on the real data, for instance, by handling high-frequency geometry in the Shotline View.

- *Budget time for transitioning to real data.* During our onsite visits, we realized that the real data did not always match the format of our test data. By budgeting time at the beginning of our onsite visits specifically for debugging our application with real data, we were able to ensure we could demo a working system to our collaborators with their real data, maximizing our productivity during the limited time we could spend with them.

- *Automate everything.* By automating the installation process of Shotviewer, we were able to more quickly get feedback from analysts on new designs. This was particularly useful when we could not remotely access our analysts machines due to their data's sensitive nature.

Although these four recommendations appear obvious in hindsight, they are solutions to pitfalls that we encountered throughout our project. We believe they are worthwhile considerations for future design studies conducted within large organizations and with sensitive data.

3.9 Conclusion

In this paper we present the results of a 15-month design study in the domain of vulnerability analysis. The contributions of this work include a problem characterization, data abstraction, and task analysis for this domain, as well as Shotviewer, a prototype vulnerability visualization tool that uses multiple linked views to display spatial and non-spatial data. We validate Shotviewer with user feedback and two case studies. Reflections on our design process also present two methodological contributions: view-design parallelism, a strategy for designing multiview visualizations; and four recommendations for conducting design studies in large organizations with sensitive data.

This design study focuses on effective encodings for multiview visualizations of multityped data. Future work could formally describe and evaluate user-interactions with multiview systems. We hope to refine the process of problem-driven research in the context of large organizations, in particular by identifying a methodology for balancing research with product development. We also see opportunities for workshop-based methods to elicit requirements from diverse stakeholders in large organizations.

CHAPTER 4

FORMATIVE DESIGN STUDY — GRAFFINITY: VISUAL ANALYSIS OF CONNECTIVITY IN LARGE GRAPHS

This chapter describes a formative design study in which we used a CVO workshop to expose shared needs of neuroscientists who were working on seemingly disparate analysis problems. The workshop revealed that many of our collaborators had questions about graph connectivity, which we explored through iterative prototyping. One of the prototypes was used to discover new circuity in the mammalian retina [19]. Our subsequent design and development created the contributions described in this chapter: 1) two novel visualization techniques that work in concert for summarizing graph connectivity; and 2) Graffinity, an open-source implementation of these visualizations supplemented by detail views to enable a complete analysis workflow.

4.1 Formative Aspects

We started this design study through interviews and observations with neuroscience collaborators. Because each neuroscientist was interested in specific research questions, interviews and observations revealed seemingly disparate opportunities for visualization. Furthermore, in the early stages of this project, senior analysts were less willing to meet with us directly, preferring to delegate meetings to junior analysts — i.e., the professors told us to talk to their graduate students.

The challenges of this project resembled our formative experiences with defense analysts. We were faced with piecing together specialized needs to find common visualization opportunities, and we had to engage diverse members of the organization — from undergraduates to professors. Instead of applying standard user-centered design methods, we decided to use a workshop that, we hoped, would provide a forum to find shared ideas. We planned a workshop by haphazardly adapting workshop descriptions from the visualization literature [7]–[9].

The workshop was tremendously successful. It fostered trust and engagement with collaborators as senior analysts were willing to meet with us regularly in subsequent design efforts. It also generated a number of visualization opportunities and revealed the a consensus of what our collaborators thought were the most pressing problems. As one participant reported, "The structured meeting created consensus by exposing shared user needs."

This work is formative to this dissertation in that it was our first experience with a workshop in applied visualization. Through further thinking on this experience, we discovered that we did not entirely understand how or why the workshop was successful. To make workshops a repeatable and dependable method for applied visualization research, we started discussions with fellow visualization researchers who had used workshops in three of their projects [7]–[9]. At first, we discussed ways to more effectively run workshops. Ultimately, these discussions expanded in scope, and we created the framework described in Chapter 5.

4.2 Motivation and Overview

Graphs are an important datatype across many domains, from transportation to neuroscience. Graph nodes represent entities and edges represent connections or relationships between entities. For instance, graphs can model the flights (edges) between airports (nodes) or synapses (edges) between neurons (nodes). In multivariate graphs, both nodes and edges can be associated with categorical attributes, such as the city of an airport, and quantitative attributes, such as the size of a synapse. Analyzing multivariate graphs often involves understanding some combination of the graph's topology and attributes [121], [122].

One important area of graph analysis is concerned with examining the direct and indirect connections between entities, their connectivity, for understanding structures implied by the graph's topology, such as airline routes that directly or indirectly connect cities. Understanding the direct and indirect connections involves analyzing a combination of the graph's adjacency (direct connections), connectivity (presence of paths connecting entities), and accessibility (entities reachable from a certain one) [121],[123]. We use the term **connectivity** to refer to the direct and indirect connections between entities based on paths, potentially considering node and edge attributes.

Understanding the connectivity of a graph is challenging because the number of possible paths connecting two entities increases exponentially with graph size [124]. This scalability problem is exacerbated by standard graph visualizations such as node-link diagrams and adjacency matrices, which have their own limitations when used for connectivity analysis. Node-link diagrams excel at topology-based tasks for small graphs but degenerate to hairballs for larger graphs [125]. Adjacency matrices are slightly more scalable for tasks related to adjacency in large graphs, but are ill-suited for tasks involving indirect connectivity because they require tracing across rows and columns to follow paths [126]. As the size of a graph increases, specialized techniques are needed to make sense of its connectivity.

Query-based approaches (e.g., [127]–[131]) are helpful when dealing with large graphs in general, and for understanding graph connectivity in particular. These systems allow analysts to query the graph for the connections between a set of nodes and return a subset of the entire graph. These subsets are often displayed as lists [131] and subgraphs [128], or use a specialized representation [130]. As the size of query results increases, however, analyzing connectivity again becomes challenging due to the large number of potential paths.

In this chapter, we propose a new technique for making sense of connectivity in large graphs. Our technique provides a flexible overview of path-based connectivity, enabling a user to explore interesting subsets of paths in a highly scalable way. The design of the technique was motivated by a collaboration with neuroscientists which, along with a review of visualization literature, allowed us to identify a set of design requirements for summarizing graph connectivity in a query-based system.

Based on these requirements, we present two contributions: 1) two novel and complementary visualization techniques for summarizing the connectivity in a subset of a graph selected by queries, the connectivity matrix, and the intermediate node table; and 2) Graffinity, an open-source implementation of these techniques. Although this work is motivated by our collaboration with neuroscientists, our visualization techniques and prototype generalize to graph analysis in other domains. We validate this work through illustrative examples and case studies with flight and neuroscience data.

4.3 Requirements

We introduce a set of requirements (R1-R5) for visualizations designed to summarize graph connectivity. We identified these requirements in a user-centered design process involving a group of up to 8 neuroscientists over a period of 18 months. We used methods including contextual inquiry [68], a CVO workshop, and informal interviews to elicit requirements and receive feedback on prototypes. The requirements were also influenced by prior visualization research, discussed in Sec. 4.4.

These requirements were informed by a domain collaboration, but we argue that they apply broadly. They are, however, not meant to be exhaustive for general graph analysis, but are targeted at a use case of analyzing connectivity between node sets. This so-called many-to-many analysis is useful for understanding relationships in a graph at a higher level of abstraction than individual nodes. For instance, an airline analyst may be interested in how two states, both with many airports, are connected by air travel. Another example is the analysis of trade or migration between geographic regions that are represented as sets of nodes [132]. In neuroscience, researchers examine the flow of signals between different types of neurons [19].

Many-to-many analysis is often performed on graphs that are too large to be drawn directly. In these cases, analysts often use queries to identify interesting subgraphs [122]. Hence, our requirements focus on query-based connectivity analysis between node sets.

We assume that all measures of connectivity are based on short paths connecting the nodes. Hence, our requirements address both abstract measures of connectivity as well as specific paths connecting the nodes.

R1 *Query many-to-many paths.* Analysts should be able to specify path queries based on node lists, shared attributes of starting or ending nodes, or the types of nodes and edges involved in the paths.

R2 *Visualize an overview of connectivity.* Analysts should be presented with a visual summary of the relationships between the nodes that they queried for. It is important that this representation appropriately scales to handle large numbers of paths. Connectivity can also be defined in various ways; hence a system targeted at analyzing connectivity should allow analysts to specify different metrics to represent connectivity.

R3 *Support dynamic aggregation of nodes and paths.* In order to understand higher level structures in a network, analysts may be interested in relationships between node sets. To support this type of analysis, dynamic aggregation of nodes, and consequently of the paths connecting these nodes, should be supported.

R4 *Visualize path details.* The details of paths, including the individual nodes and edges that make up paths, as well as the node and edge attributes, should be accessible on demand.

R5 *Visualize path context.* The context of a path describes how it is embedded within the topology of the graph. Also, when appropriate, a meaningful spatial representation of the nodes and edges should be available.

Finally, underlying our requirements is the assumption that analysts have already identified interesting queries about the connectivity. These queries may be based on existing domain knowledge and bottom-up analysis such as tracing paths in node-link diagrams or other visualization techniques discussed in the next section.

4.4 Related Work — Graph Connectivity

We focus our discussion of related work on techniques that support path-based connectivity analysis in large, multivariate graphs. Summaries of the extensive research on graph visualization beyond path analysis are available for various areas, including visualization of large graphs [133], dynamic graphs [134], and multivariate graphs [135].

Representations for paths in graphs include traditional node-link layouts, adjacency matrices, and path-listing techniques [131] as shown in Fig. 4.1. Each technique can be combined with an initial query step to reduce a larger graph into a smaller subgraph (R1) to enhance scalability. Path queries are supported by general purpose graph software

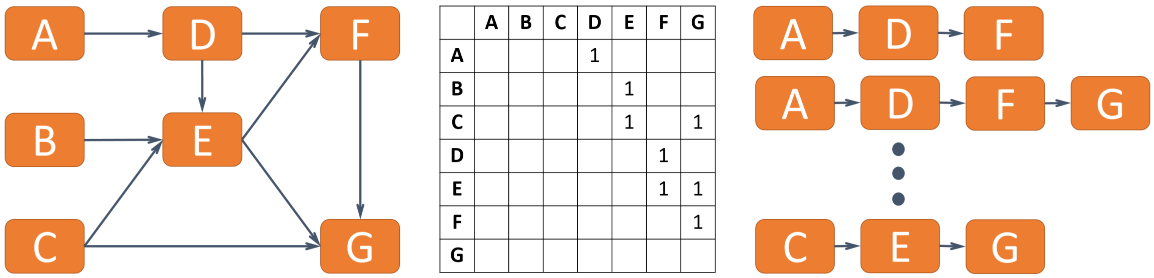

Figure 4.1. Standard encodings for displaying paths in graphs. Analyzing node connectivity is challenging with traditional graph encodings and path listing techniques. Suppose a query for paths connecting nodes A, B, C with nodes F, G returned the subgraph shown here. Node-link diagrams (left) give an overview of graph topology but require manual tracing to analyze relationships between the start and end nodes. Adjacency matrices (middle) are ill-suited for connectivity analysis as tracing paths necessitates indirection between rows and columns. Path lists (right) do not provide a connectivity overview.

packages, such as Tulip [136] and Gephi [137], and databases such as Neo4j [138].

Traditional node-link diagrams support the exploration of connectivity by enabling analysts to trace paths to identify the relationships between nodes (R4) within the graph's topological context (R5) [126]. They fail to scale to many nodes and paths (R2), however, as they require manual tracing of paths, and they quickly degenerate to hairballs when they exceed about 50 nodes and 200 links [125]. RelFinder [128] and the path topology view in Pathfinder [131] are examples of node-link diagrams being used to display the results of path queries.

Adjacency matrices are generally considered ill-suited for path-related tasks because they require tedious manual tracing between rows and columns to follow the paths [126]. Augmented matrices exist to support browsing paths and accessing details of those paths (R4). In MatLink, Henry et al. [139] augmented adjacency matrices with additional edge representations. This approach has been expanded by Shen and Ma [140], who draw links directly on top of matrices. Recently, the EgoLines tool [141] has used a similar approach for representing paths in egocentric adjacency matrices. These augmented matrix approaches are appropriate for following a relatively small number of paths, but do not provide a scalable overview of connectivity (R2).

Matrices can also be augmented to display aggregations of relationships (R3). Aggregated matrices such as those used in Honeycomb [142] and MapTrix [132] are suitable for analyzing adjacency between sets of nodes. Yet, these approaches suffer from the same problem as nonaggregated matrices when considering connectivity and hence do not provide an adequate overview of connectivity (R2).

There are also aggregation approaches for node-link diagrams. PivotGraphs [143] create aggregate representations of graphs based on node attributes (R3), but this representation hides the paths that connect individual nodes and hence does not meet the requirements related to individual paths (R4, R5). GraphCharter [130] is a pivot graph implementation modified to support iterative query-based path browsing, but it does not adequately provide information about connections between many-to-many nodes. Details-to-overview-via-selection-and-aggregation [122] enables users to transform node-link representations of a graph into aggregated summaries (R3). These summaries can explain connections between sets of nodes, but they do not necessarily support analysis of connections within those sets or between individual nodes (R4).

Statistical summaries can be used to give an overview of connectivity. B-Matrices [144], [145] and graph prisms [146] offer such summaries of nodes and edges in a graph, such as the number of reachable nodes, but these approaches contain little information about relationships between specific nodes (R2).

Specialized techniques are particularly suitable for query-based path analysis and have focused extensively on querying paths between a small number of start and end nodes. RelClus [129] clusters paths hierarchically according to length and co-occurring nodes (R3) and displays these clusters in a tree view. This technique, however, does not provide an overview of many-to-many relationships without manual aggregation (R2). Aleman-Meza et al. [127] support browsing paths (R4) to identify interesting regions of a graph, but do not provide explicit summaries of the resulting connectivity (R2). PathFinder [131] supports querying for paths between sets of nodes (R1) and interactive browsing and ranking of those paths (R4), but it does not provide an adequate overview of the connectivity (R2).

This work aims to support queries between large sets of start and end nodes, visualize an overview of their connectivity, and then support analysis of the paths in detail.

4.5 Connectivity Overviews

In this section, we describe two visualization techniques for providing an overview of graph connectivity. These two complementary techniques are designed to give an overview of paths between nodes (R2) and support dynamic aggregation of those paths (R3). The first technique is the connectivity matrix, which provides an overview of paths as relationships between start and end nodes. The second, complementary, technique is the intermediate node table, which provides additional details about the role of intermediate nodes in these paths. These two techniques are implemented in a prototype, called Graffinity, that addresses the other requirements of querying for paths (R1), accessing path details (R4), and providing context (R5). We discuss these features of Graffinity in Sec. 4.6. Here, we describe the connectivity matrix and the intermediate node table assuming that a user has provided a query for paths connecting sets of nodes.

4.5.1 Connectivity Matrix

We designed the connectivity matrix to provide users with an overview of path-based connectivity when they query for paths between sets of nodes. The connectivity matrix visualizes sets of paths connecting start and end nodes. We apply metrics to these path sets, such as the count of paths, and display the results of these metrics in a matrix. The matrix rows correspond to the start nodes and the columns correspond to the end nodes. This matrix representation is a generalization of the adjacency matrix for showing path relationships. In the remainder of this subsection, we provide a definition of path sets and example metrics to analyze those sets, and discuss the aggregation of paths.

A query returns a subgraph $G = (N, E)$ that contains paths between the user-specified start nodes, $N_{start} = \{start_0,\ start_1,\ \dots\}$ and end nodes, $N_{end} = \{end_0,\ end_1,\ \dots\}$. The paths are $P = \{p_0,\ p_1,\ \dots\ , p_k\}$. We define *connectivity sets*, C, for all pairs of the start and end nodes as the set of paths that connect those nodes.

Formally,

$$C(start, end) = \{p \mid p \in\ P\ \wedge\ Start(p) = start\ \wedge\ End(p) = end\}$$
$$\forall\ start \in N_{start}, \forall\ end \in N_{end}.$$

Each set contains the paths matching the query criteria that connect a pair of start and end nodes. An example derivation of the path sets is shown in Fig. 4.2.

Each row in the connectivity matrix corresponds to a start node, $start \in N_{start}$, and each column corresponds to an end node, $end \in N_{end}$. Each matrix cell represents the path set, $C(start, end)$.

We use the cells of the matrix to visualize a metric derived from its path sets. A metric is a function that operates on a path set and returns one or more values representing those paths. The path count is an intuitive metric for summarizing query results. Two additional domain-agnostic metrics are the count and minimum length of paths in a set. Fig. 4.2 shows the path count and minimum length metrics, yet there are many possibilities for other metrics that account for node and edge attributes, e.g., taking edge weights into account.

The result of the metrics can be displayed using various visual encodings. Color coding the cells (i.e., creating a heat map) provides a visual summary of connectivity when using metrics that return a single value per set. More complex metrics that return an array of values could make use of a small multiples display of the table or a glyph representing multiple values in a cell [147]. These are described in more detail in Sec. 4.6.

Aggregating nodes in N_{start} and/or N_{end} can help to further simplify a connectivity matrix. For example, we could group nodes and their associated paths by node attributes to capture higher level phenomena in the network, to, e.g., group all airports in the New York City area. Aggregation is realized by taking the union of path sets. For instance, if two nodes $(start_a,\ start_b) \in N_{start}$ are to be aggregated, then a new aggregated connectivity set is computed by taking the union of both existing sets,

$$C(start_a \cup start_b,\ end) = C(start_a,\ end) \cup C(start_b,\ end) \quad \forall\ end \in N_{end}.$$

These aggregated sets can be displayed using the aforementioned metrics and encodings as in Fig. 4.2. Note, however, that the scales of aggregated and nonaggregated values can be quite different, which potentially requires dedicated visual encodings when showing both aggregated and nonaggregated connectivity in the same matrix.

The connectivity matrix intentionally hides information about the intermediate nodes of paths to support analysis of the general connectivity between start and end nodes. However, understanding the role of intermediate nodes can be important for certain analysis tasks, such as identifying major hubs in a flight network. Thus, we introduce an additional, complementary visualization that focuses on the intermediate node information, which is described next.

4.5.2 Intermediate Node Table

The intermediate node table, illustrated in Fig. 4.3, visualizes the properties of a path set defined by an intermediate node at a specific position in a path. For instance, in the the flight graph, queries for paths of length three identify paths of three flights between four airports. The intermediate node table defines path sets based on the airports used for layovers and whether those airports are the first or second stop in the journey. Again, we provide a formal definition of path sets and describe considerations for visualizing these sets.

Formally, the intermediate node table defines path sets based on the intermediate node, the path length, and node position. Let L be the maximum length of all paths in the query result P. Let (j, l) represent position j in paths of length l. Also, let $node(p, j)$ return the node at position j in path p. The intermediate node sets, I, are defined, formally,

$$I(n_i, (j, l)) = \{p \mid p \in P \wedge Node(p, j) = n_i \wedge Len(p) = l\}$$
$$\forall\, n_i \in N, \forall\, j \in [1, \ldots, l], \forall\, l \in [1, \ldots, L].$$

These sets are represented in a table where the rows correspond to nodes, and the columns correspond to the position of the node in a path of a given length.

The number of columns in the intermediate node table depends on the length of paths returned by a query. In queries for paths of length two, the table contains only one column representing the middle node position in all of the paths. In queries for paths of length three, the table contains three columns representing the possible positions for nodes inside the paths, as shown in Fig. 4.3.

Various metrics can be used for summarizing the path sets in the intermediate node table. In addition to the count metric used in Fig. 4.3, other metrics could include the weight of paths passing through an intermediate node, or the number of unique start and end nodes that an intermediate node connects.

Just as the connectivity matrix supports various visual encodings to represent metric results, the intermediate node table supports similar encodings. Likewise, dynamic aggregation of the intermediate node table based on node attributes is possible.

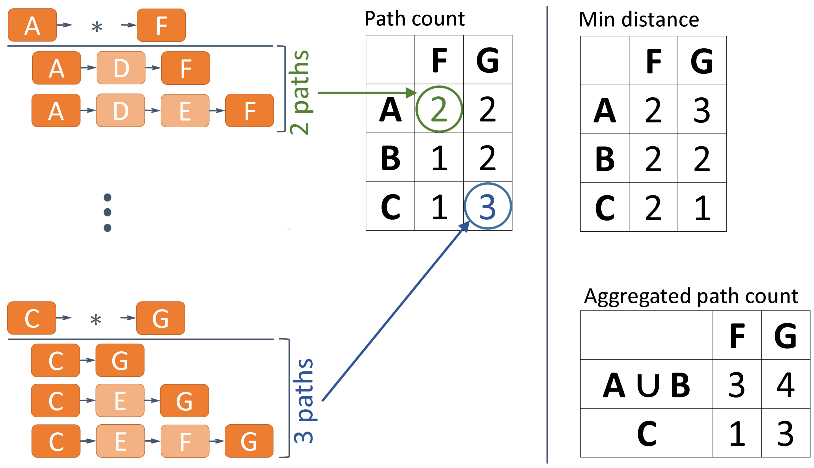

Figure 4.2. Example derivation of the connectivity matrix. Here, we show the construction of a connectivity matrix using the subgraph introduced in Fig. 4.1. We create path sets based on common start and end nodes, and then represent those sets in a matrix where each cell shows a metric applied to paths connecting a pair of nodes. Examples of path-based metrics shown here are the count of paths connecting two nodes and the minimum distance between two nodes. Additionally, the matrix rows and columns can be aggregated by computing the union of the corresponding path sets.

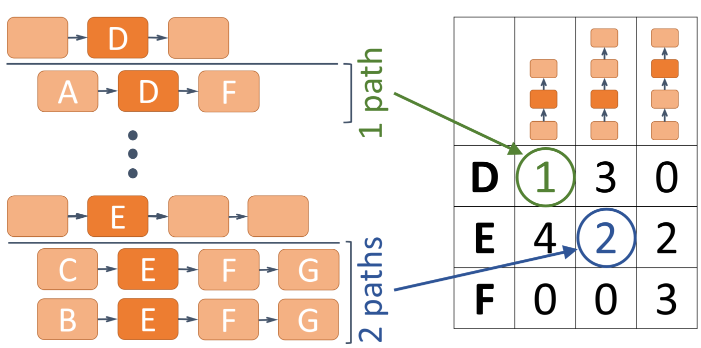

Figure 4.3. Example derivation of the intermediate node table. In this intermediate node table for the connectivity matrix shown in Fig. 4.2, rows correspond to nodes and columns correspond to a node's position in a path of a certain length. Here, node D appears once as the middle node in paths of length two. Node E is included twice as the second node in paths of length three.

The intermediate node table, hence, displays a summary of the intermediate nodes returned by a path query. When paired with the connectivity matrix, the two techniques display an overview of paths connecting start/end nodes as well as of the importance of the intermediate nodes that those paths pass through. Interactive highlighting and selections can be used to access the relationships of paths between the two views. These interactions, along with a detailed discussion of metrics, encodings, and aggregation, are described in the following section.

4.6 Graffinity

We have implemented the connectivity matrix and the intermediate node table in a prototype system called Graffinity, shown in Fig. 4.4. Graffinity includes three additional components: a query interface and two supplemental views.

While our system was designed with neuroscience data in mind, we introduce its functionality with a flight dataset. This dataset is a graph of flights in the US over three days in 2015. It consists of 308 airports (nodes) and 13K flights (edges) connecting the airports. Nodes have categorical attributes, including a unique three-letter airport code, a city name, and a state. The categorical elements of this dataset have a hierarchical structure: one or multiple airports are associated with one city, one or multiple cities are associated with one state. Nodes also have quantitative attributes, such as their degree, as well as geographic locations. Edges have categorical attributes, such as an identifier for the airline, and quantitative attributes, including arrival time, departure time, and length of any delays.

4.6.1 Queries

The query interface supports visually defining queries for many-to-many paths either by specifying lists of start and end nodes or defining node sets based on shared categorical attributes (satisfying R1). In addition, a maximum path length must be provided. Fig. 4.5 shows an example where the start nodes are airports in any of four states, and the end nodes can be defined using any of the options shown.

Graffinity also supports defining advanced queries in the graphical interface, including restrictions on the edge types and intermediate nodes. Additionally, queries can be

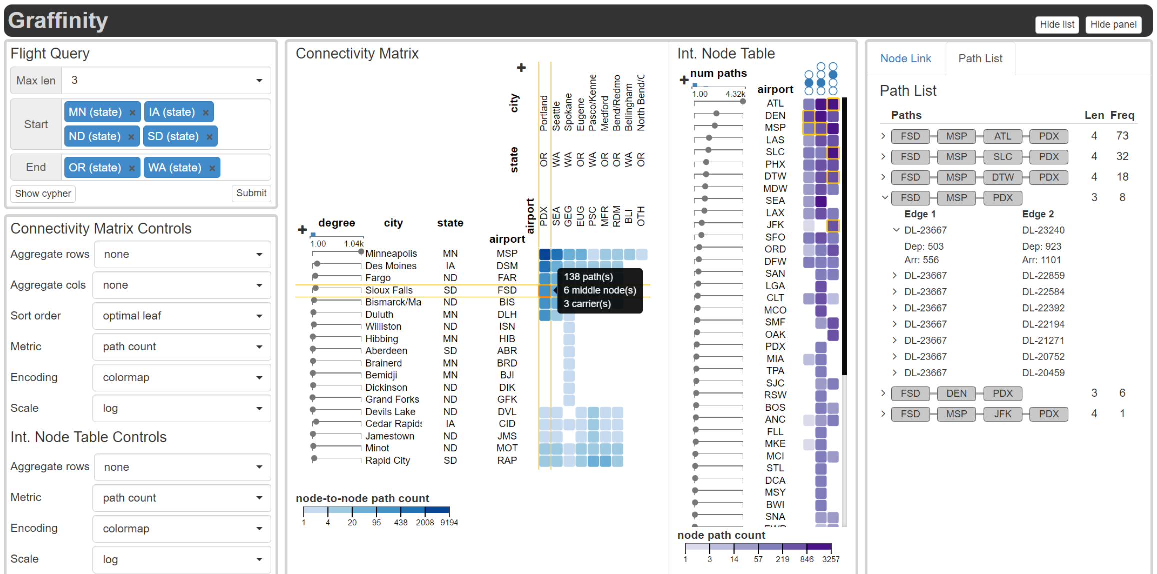

Figure 4.4. Graffinity visualizing 11727 flight paths. The paths have length $\leq$ three, and they connect states in the mid-western USA (Minnesota, Iowa, North Dakota, and South Dakota) to states in the Pacific Northwest (Oregon and Washington). Graffinity consists of five views: the query interface, the connectivity matrix, the intermediate node table, and two views showing details about selected paths: the path list and the node-link view. The 138 paths connecting the airport FSD (Sioux Falls, SD) to PDX (Portland, OR) are selected and displayed in the path list view.

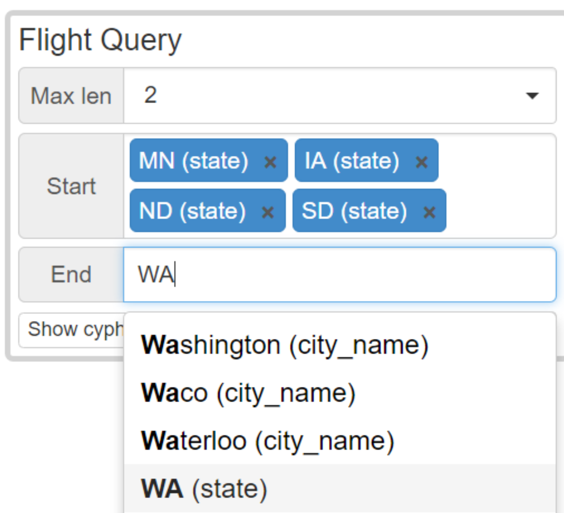

Figure 4.5. Example flight query in Graffinity. The flight query interface defines paths by a maximum length as well as attributes of the start and end nodes. Here, the user can select any attributes matching the input string “WA.”

specified in the cypher language [138], which enables queries of arbitrary complexity that are not easily specified using a graphical user interface.

In addition to queries, paths can also be filtered by quantitative or categorical node attributes. By filtering out nodes with a high degree, for example, we can reveal connections that do not go through the major hubs of a network.

4.6.2 Connectivity Overview

The connectivity overview consists of the connectivity matrix and intermediate node table as described in Sec. 4.5. Here, we describe the details of their implementation, including the display of path metrics and visual encodings, dynamic aggregation, node attributes, reordering, highlighting, and selection.

The cells in the connectivity matrix and the intermediate node table display the result of metrics applied to path sets. The default metric for both views is a path count displayed with a quantitative color map as in Fig. 4.4. There are many other possible metrics beyond a path count, such as the percentage of delayed flights connecting two airports, which is shown in Fig. 4.6.

In addition to dynamic metrics, Graffinity supports interactively changing the visual encodings. Fig. 4.7, for example, shows two encodings that use bar charts. The left example uses a bar to encode the total number of paths. The right example contains two bars, where the first bar visualizes the number of paths of length one, and the second bar visualizes the number of paths of length two.

Graffinity supports dynamic aggregation of nodes based on their attributes. This is important for analyzing higher level relationships in the graph, for instance, to understand connections between states instead of individual airports. This aggregation is demonstrated in Fig. 4.8, where the starting nodes are aggregated by state. Aggregated sets can be expanded to show the nested rows or columns. We use different color scales for aggregated values to make it obvious that a row or column is aggregated and to account for the often significantly different data ranges between aggregates and individual nodes.

Graffinity also displays node attributes. Node attributes are visualized adjacent to the rows and columns of the connectivity matrix and intermediate node table. Categorical attributes are visualized as strings. Quantitative attributes are shown using dotplots as

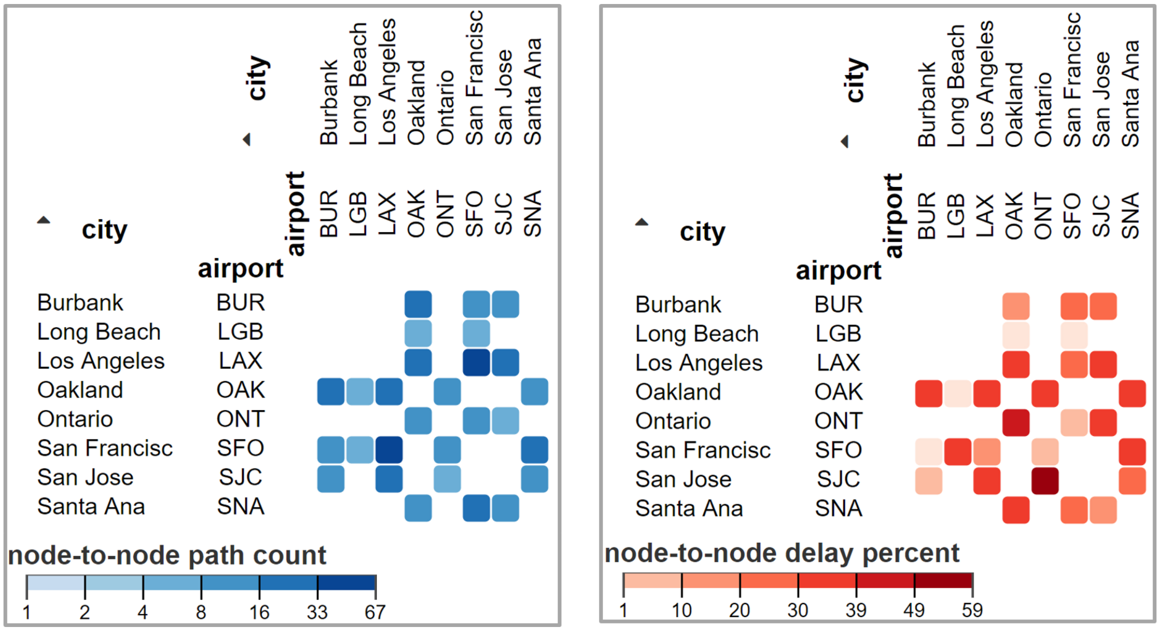

Figure 4.6. Example metrics visualized in Graffinity. Different metrics can be applied to path sets, such as the flights between Los Angeles and San Francisco area airports. Here, the count of flights is shown on the left, and the percentage of flights with more than a 15-minute delay is shown on the right.

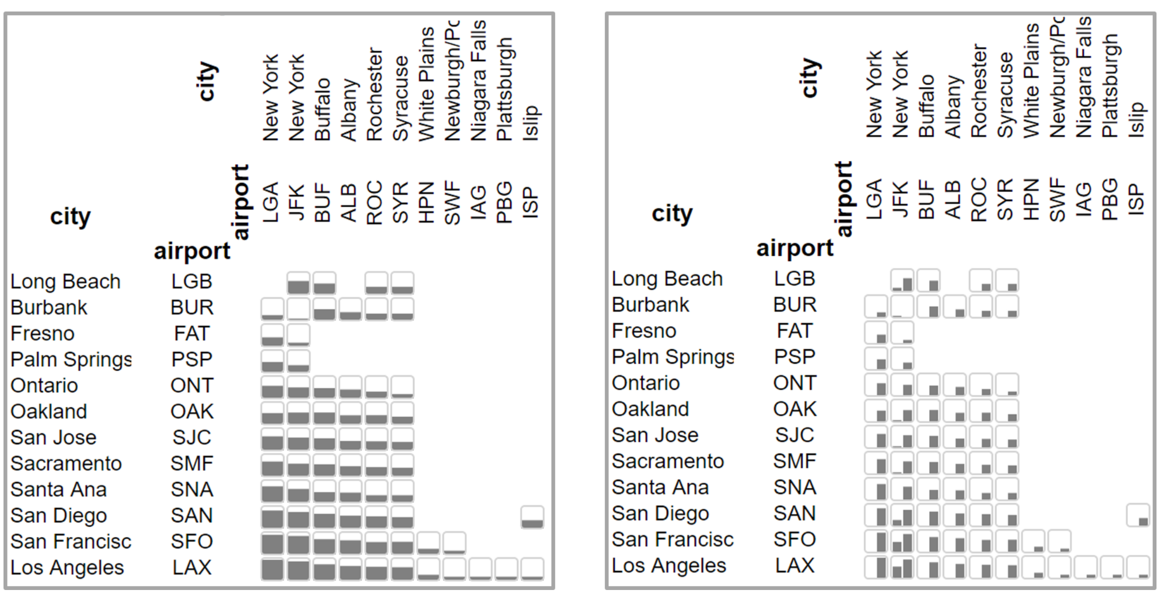

Figure 4.7. Example encodings used in Graffinity. Here, we show two encodings for the number of paths connecting California to New York. Left is a bar chart where height encodes the number of paths. Right is a bar chart where the left bar encodes paths of length one and the right bar encodes paths of length two.

in Fig. 4.8. The dotplots are well suited to display multiple entries, which is particularly important for representing aggregated sets of nodes.

The features that can be discovered in matrices are strongly influenced by the matrix ordering [148]. Consequently, Graffinity supports dynamic reordering based on either node attributes or using matrix reordering algorithms [149]. An example of the optimal leaf ordering applied to a matrix is shown in Fig. 4.9.

Linked highlighting reveals relationships between the connectivity matrix and intermediate node table. For example, hovering over a node or path set in the intermediate node table reveals the flights and paths that pass through that node in the connectivity matrix. Similarly, hovering over a node or path set in the connectivity matrix highlights the intermediate nodes used in those paths. Individual cells can also be selected so that the contained paths can be inspected in detail in the supplemental views.

4.6.3 Supplemental Views

The supplemental views are meant to provide context (R5) and details (R4) about a selection of paths. They are updated every time a cell in the connectivity matrix or the intermediate node table is selected. We currently provide node-link diagrams—as in Fig. 4.10—and path-list views—as in Fig. 4.4.

We provide two layouts for the node link diagram. The first is a force-directed layout that provides topological context. This layout, for example, lets analyst identify well-connected nodes in the selected paths. The second layout renders the network in a spatial context and can be overlaid with, e.g., a map, as shown in Fig. 4.10.

The path-list views enables analysts to browse the paths and provides details about the individual paths (R4). In particular, it displays a list of the selected paths in a motif hierarchy. For the flight dataset, the motifs describe the airports that flights pass through. The motifs can be expanded to display the underlying paths, e.g., to display information such as their ID, carrier, and departure times.

The spatial layout and the motifs are domain specific, i.e., a map of the US is an appropriate layout for the US flight data, whereas a map of the location of neurons in a microscopy image could be an appropriate layout for the connectomics data. Similarly, the motifs and details displayed in the path list view depend on the dataset. In the flight data,

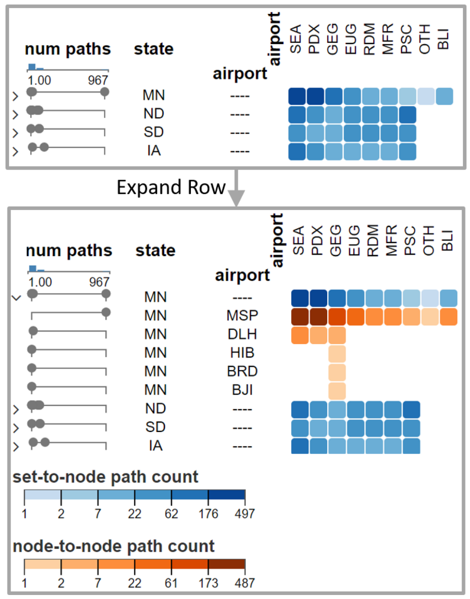

Figure 4.8. Example dynamic aggregation in Graffinity. Here, the connectivity matrix from Fig. 4.4 is aggregated by starting node state; the airports in Minnesota (MN) are then expanded. Different color scales in aggregated cells account for differences in scales and emphasize the aggregation. Dotplots represent quantitative attributes for both the aggregated and expanded rows.

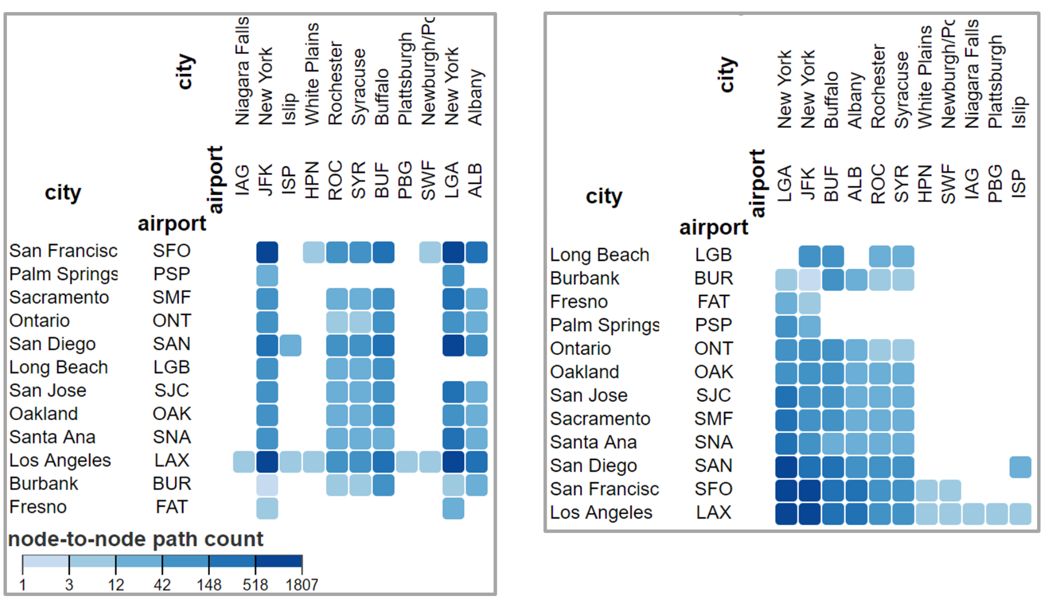

Figure 4.9. Example reordering of the connectivity matrix in Graffinity. The left matrix is in the order that was returned by the database query, and the right matrix is using an optimal leaf-ordering algorithm.

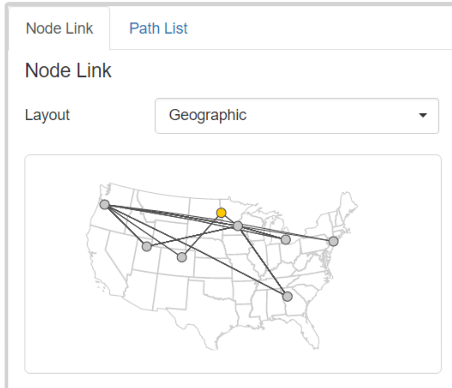

Figure 4.10. Example node-link view with geographic layout in Graffinity. Here, we show the data selected in Fig. 4.4. Graffinity also supports force-directed layouts in this view.

the airport codes provide a meaningful path aggregation, and in the neuroscience data, the classification of neurons provides a meaningful aggregation for our collaborators.

4.6.4 Implementation

Graffinity is a web-based client-server tool that was developed using a combination of web technologies. The visualizations are implemented in ES6 using D3, AngularJS, and Bootstrap. The server is implemented with Python and Flask. It uses the Neo4j graph database, and it executes path queries with a breath-first search strategy. We have included the source code in our supplemental material and made it available on GitHub under an open-source license: `http://bit.ly/Graffinity`.

4.7 Validation

We demonstrate the usefulness of Graffinity through a case study analyzing a *connectome*, a graph of connections between cells. Our collaborators (some of whom are also coauthors) are connectomics researchers studying the connectome of cells in the retina. In this 18-month collaboration, we have leveraged user-centered design methods, such as a CVO workshop and contextual inquiry [68], to understand the analysis needs of this group of neuroscientists. We developed Graffinity to support those needs. In this section, we briefly describe the data involved in retinal connectomics research, followed by a case study where Graffinity was used to detect errors in the connectomics dataset.

The retinal connectome that we worked with, a database called RC1, was generated from a rabbit retina through automated electron microscopy imaging, image processing, and manual annotations [150]. It is a multivariate graph of 15K neurons (nodes) and 26K synapses (directed edges) [151]. The nodes have categorical attributes, such as a label that specifies the type of the cell. They also have quantitative attributes, such as the size of the cell's convex hull. The edges have categorical attributes, such as the type of synapse. It is important to note that the nodes and edges in the graph are annotated based on microscopy images of the retina, i.e., the graph's nodes and edges are an abstraction of the connections observed in the images.

Understanding the connectivity of retinal cells enables researchers to reason about the flow of information through the retina and the functions of various cells. For example, Lauritzen et al. [19] recently identified the winner-take-all, rod-cone crossover networks that switch between pathways for cone-driven bright light vision and those for rod-driven dim light vision. Fast crossover networks are particularly important in mesopic environments where both rods and cones are active and compete for network dominance. This circuitry was discovered through the analysis of approximately 8000 different paths of various lengths in the RC1 connectome.

In one of our sessions for getting feedback on Graffinity, we worked with our collaborator to revisit the cone-rod crossover analysis performed by Lauritzen et al. [19]. One particularly interesting part of this analysis occurred when we discovered an anomalous pathway in the dataset that had not previously been detected. In the remainder of this section, we describe the steps of detecting that anomaly and analyzing its significance.

In the analysis, we queried for two-hop paths that matched the cone-rod crossover circuitry. This resulted in 272 paths that connected 90 cone bipolar cells (denoted with labels that start with *CBb*) to 74 rod bipolar cells (label of *Rod BC*) through 104 intermediate amacrine cells (label containing *YAC* or *AC*).

In these query results, we were interested in connections formed by classes of cells. We aggregated the source nodes (rows of the connectivity matrix) and the intermediate nodes (rows of the intermediate node table) by label. We then inspected the intermediate nodes that connect rods and cones.

In particular, we examined intermediate nodes with the label *YAC Ai*. One of these cells had many more connections than the others of the same label. By expanding the aggregated *YAC Ai* row, we were able to use linked highlighting between the connectivity matrix and intermediate node table to reveal the paths connected by the intermediate nodes. In particular, we noticed that cell *179* received input from a cell with label *CBb3*, shown in Fig. 4.11, which violated the expected connections for that cell type.

The question triggered by this finding is central to all of connectomics: is this anomaly a biological error, which addresses the nature of biological wiring precision, or a technical error inherent in connectomics mapping? By selecting the paths through cell *179* in the intermediate node table, we were able to use the path list view to drill down to the individual synapses responsible for these paths. With these synapse IDs, we accessed the images of the database and discovered that the connection from *CBb3* to *YAC Ai 179* was an error. Although this crossover network had been rigorously analyzed with coarser granularity, fine-scale annotation errors persisted and became apparent when viewed with Graffinity.

4.8 Discussion

In addition to our case study validation, we discuss the qualitative feedback on Graffinity and the scalability of the proposed visualization techniques. We also reflect on the role of these techniques in the larger scope of graph analysis.

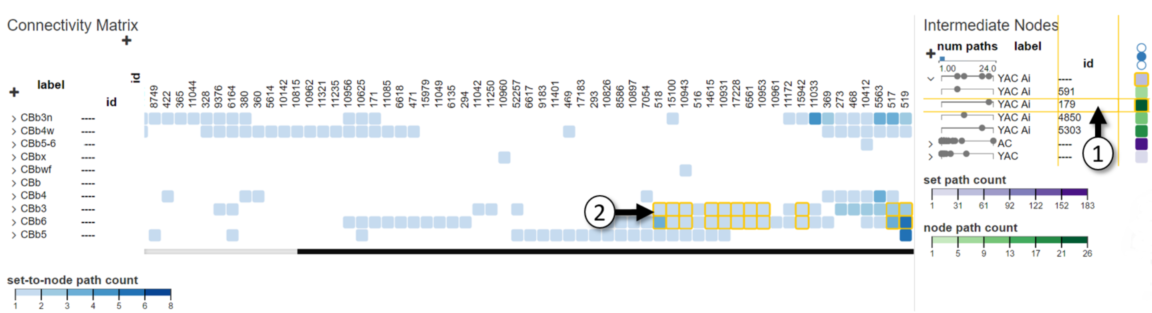

Figure 4.11. Graffinity visualizing cone-rod crossover. Here, the connectivity matrix shows paths of length two connecting cone bipolar cells to rod bipolar cells. 1) The intermediate node 179 with label *YAC Ai* participates in a large number of these crossover paths. Hovering on this row in the intermediate node table reveals the starting and ending nodes of these paths in the connectivity matrix. 2) The yellow boxes around matrix cells for the rows of *CBb3* and *CBb6* show that node 179 receives input from both of these classes. This is surprising as nodes with label *YAC Ai* should not form connections with *CBb3* nodes, although they technically could access them. We questioned whether this anomaly was a biological wiring error or a data collection error. Ultimately, Graffinity guided access to the database images, showing the anomaly to be an annotation error.

Our collaborators provided positive feedback on the range of connectivity analysis supported by Graffinity during 6 hours of informal interviews and demonstrations. One analyst said that Graffinity "generated figures that I didn't think were possible" and that those figures were "exactly what I need" for her on-going research of neuron connectivity. Another analyst referred to the connectivity matrix as "very powerful ... and truly exciting [for connectivity analysis]." Throughout these feedback sessions, we encouraged the analysts to use Graffinity to visualize both novel and previously documented patterns in connectivity. In both cases, they were able to generate new insights about neuron connectivity.

One goal of our feedback sessions was to evaluate whether relatively short paths were sufficient for connectivity analysis. Throughout these sessions, our collaborators expressed interest in querying for paths of length four or less. This supports our assumption that, in practice, relatively short paths are desirable for connectivity analysis, which holds for transportation networks, and we believe is valid for many other analysis scenarios.

As the number of paths connecting two nodes increases exponentially with path length, there are computational limitations regarding query-based analysis. Path queries on the highly connected flight dataset that include paths of length three often require minutes to execute. In contrast, the neuroscience dataset is relatively sparse and supports interactive query results for paths of length four. Graffinity could be improved with streaming query results and progressive visualization updates [152] or with heuristics that predict connectivity.

Informal testing with the flight and neuroscience datasets revealed that the connectivity matrix and intermediate node table scale effectively with the number of paths returned by a query, but they suffer from limitations common to other table-based visualizations as the number of nodes returned by a query increases. Both techniques can interactively display around 100K paths as both visualizations are created in linear time. However, the number of rows and columns in each visualization is limited by screen space, which can require scrolling as seen in Fig. 4.4 and Fig. 4.11.

Our reliance on queries to provide an overview of connectivity of large graphs requires that the analyst have knowledge of the graph and can formulate relevant queries. While this is true in many scenarios, such as the flight dataset, for which it is easy to formulate queries by a typical user, and for the neuroscience dataset, which our collaborators know well, it implies that Graffinity is not well suited to explore a graph that a user does not know much about. Due to this, Graffinity should be used as part of a larger tool chain that supports open exploration, such as through degree-of-interest functions [153] or visual summaries [143].

4.9 Conclusion

In this chapter, we introduced the connectivity matrix and the intermediate node table, two novel visualization techniques for summarizing connectivity relationships in large graphs. The connectivity matrix uses the metaphor of an adjacency matrix generalized to show path-based relationships between start and end nodes. This scalable representation avoids required manual tracing of adjacency matrices. The intermediate node table reveals information about nodes hidden by the connectivity matrix. These two techniques provide an overview of tens of thousands of paths potentially using a variety of connectivity metrics.

We realized these techniques in a prototype system, called Graffinity. This system also contains two supplemental views, a path-list view and node-link diagram view, so that a wide range of connectivity questions can be answered and all our requirements can be addressed. We demonstrated Graffinity's fitness for use in case studies on a retinal connectomics dataset, although more work remains to integrate it into a larger tool chain for graph analysis.

Our prototype implementation illuminated interesting areas of future work focused on the exploration of connectivity metrics and visual encodings to represent these metrics. We have demonstrated a few interesting metrics for path analysis, such as the path count and minimum length, as well as domain-specific metrics such as the percentage of delayed flights. We hope to explore the design space of connectivity metrics and optimal visual encodings for their results in the future.

Finally, the Graffinity system could be extended to support comparison tasks. For example, it would be interesting to compare the flight connectivity using individual airlines, to, e.g., see the differences in connectivity of two airline carriers. Another interesting comparison use case is analyzing inhibitory and excitatory synaptic pathways in the retina.

These comparisons could be achieved using either small multiples of the connectivity matrix and the intermediate node table, or explicit metrics for the differences of these queries, paired with tailored visual encodings.

CHAPTER 5

A FRAMEWORK FOR CREATIVE VISUALIZATION-OPPORTUNITIES WORKSHOPS

This chapter introduces the core contribution of this dissertation: a framework for CVO workshops in applied research. The framework results from a 2-year, cross institution collaboration among five visualization and creativity experts whom we gathered to reflect on our own experiences using a workshop in the formative design study with neuroscientists [2]. More specifically, we compared two workshops — our own workshop with neuroscientists and one with energy analysts [2], [8] — to explore ways in which we could improve future workshops. However, we lacked a holistic framework for thinking about how to use and improve workshops in future projects. We searched existing workshop resources for ideas that could be directly adopted in visualization research, but we found that existing workshop guidance did not appropriately emphasize three visualization specifics — the visualization mindset, visualization methodologies, and visualization methods — as described in Chapter 2. Accordingly, the scope of our collaboration expanded as we set out to understand how and why to use CVO workshops in applied visualization.

We analyzed our collective experiences with 17 workshops in 10 visualization contexts [2], [7]–[9], [29]–[34] and reviewed relevant literature from design [16], [36]–[40], software engineering [20], [41]–[46], and creative problem-solving [15], [22], [23], [47], [48], [51]. Through a messy and iterative research process, we proposed and developed a framework for CVO workshops. The framework is the first meta-analysis of CVO workshops in applied visualization. It provides evidence that CVO workshops are beneficial to projects with a wide range of domain collaborators and intended outcomes. More importantly, it is the first actionable guidance for researchers who want to use CVO workshops in their own projects. It is step toward making CVO workshops a repeatable and dependable method for the early stages of applied work.

This chapter differs from the previous two chapters in that it results from a collaboration of visualization and creativity researchers. We therefore use the term *we* to refer to the collective opinions, experience, and expertise of all the coauthors who contributed to the reflective analysis and resulting CVO workshop framework [3].

5.1 Motivation and Overview

In our experience, CVO workshops provide tremendous value to the stakeholders of applied visualization projects — researchers and the domain specialists with whom they collaborate. CVO workshops provide time for focused thinking about a collaboration, which allows stakeholders to share expertise and explore visualization opportunities. In feedback, one participant reported the workshop was "a good way to stop thinking about technical issues and try to see the big picture" [29].

CVO workshops can also help researchers understand analysis pipelines, work productively within organizational constraints, and efficiently use limited meeting time. As one participant said, "The structured format helped us to keep on-topic and to use the short time wisely. It also helped us rapidly focus on what were the most critical needs going forward. At first I was a little hesitant, but it was spot-on and wise to implement" [31].

Furthermore, CVO workshops can build trust, rapport, and a feeling of co-ownership among project stakeholders. Researchers and collaborators can leave workshops feeling inspired and excited to continue a project, as reported by one participant, "I enjoyed seeing all of the information visualization ideas ... very stimulating for how these might be useful in my work" [29].

Based on these reasons, our view is that CVO workshops have saved us significant amounts of time pursuing problem characterizations and task analysis when compared to traditional visualization design approaches that involve one-on-one interviews and observations. What may have taken several months, we accomplished with several days of workshop preparation, execution, and analysis. In this chapter, we draw upon 10 years of experience using and refining workshops to propose a framework that enables others to use and improve future CVO workshops.

In the remainder of this chapter, we summarize our workshop experience and propose terminology for describing workshops. Then, we summarize the 2-year reflective research process that we used to create the framework. After that, we introduce the framework, which consists of 1) a set of characteristics that contribute to effective workshops; 2) a process model that identifies actions before, during, and after workshops; 3) a structure that describes what happens in the beginning, in the middle, and at the end of effective workshops; 4) a set of 25 actionable guidelines for future workshops; and 4) three example workshops that serve as a starting point for researchers interested in using workshops in their own projects.

The framework is supported by a companion website, an interactive resource that includes: 1) an audit trail summarizing 30 reflective artifacts that show how we created the framework through a process of critically reflective practice; 2) a list of 25 pitfalls to avoid in future workshops that are organized around constructs of the CVO workshop framework; and 3) a set of 15 example methods that we have used or would consider using in future workshops. We provide these details in a website because it allows users to interact with the content in a way that is not possible in printed format. More importantly, the website is a resource of and contribution to the visualization community. It is a living document that we will update as our understanding of workshops evolves. We therefore refer to the website throughout this chapter and encourage readers to explore its content for themselves: `http://www.bit.ly/CVOWorkshops`. The companion website and its source code are also archived with this dissertation in ProQuest.

5.2 Workshop Experience and Terminology

To create this framework, we gathered five researchers who used workshops on three continents over the past 10 years. Our collective experience includes 17 workshops in 10 contexts: 15 workshops in 8 applied collaborations, summarized in Table 5.1 and Table 5.2; and 2 participatory workshops at IEEE VIS that focused on creating visualizations for domain specialists [33], [34].

The ways in which we — a group of five visualization and creativity researchers — use workshops have evolved over 10 years. In three of our projects, we used a series of workshops to explore opportunities, develop and iterate on prototypes, and evaluate the resulting visualizations in collaborations with cartographers [7], energy analysts [8], and

Table 5.1. Summary of the projects in which we have used CVO workshops: six resulted active collaboration that led to publications at major visualization venues [P1] – [P6], one did not result in active collaboration [P7], and one is in progress [P8]. We characterize our involvement in these projects as either the primary researcher or as supporting researchers. The researchers represented by * are colleagues who were involved in each project but not directly involved with developing the CVO workshop framework.

ID	Year	Domain	Summary	Workshops	Result	Prim.	Supp.	Ref.
P1	2009	Cartography	"Reimagining the legend as an exploratory visualization interface"	3	Paper	JD	*	[7]
P2	2012	Smart Homes	Deliver insights into the role of smart homes and new business potential	4	Paper	SG	JD,SJ,*	[8]
P3	2012	Human terrain	"develop [visualization] techniques that are meaningful in HTA"	3	Paper	JD	*	[9]
P4	2015	Neuroscience	Explore problem-driven multivariate graph visualization	1	Paper	EK	MM, *	[2]
P5	2015	Constraint prog.	Design performance profiling methods for constraint programmers	1	Paper	SG	*	[29]
P6	2017	Psychiatry	Support visual analysis of determining or associated factors of suicide	1	Paper	*	EK,*	[30]
P7	2017	Genealogy	Discover opportunities to support visual genealogy analysis	1	—	*	EK,MM,*	[32]
P8	2017	Biology	Support phylogenetic analysis with visualization software	1	In-progress	*	EK,MM,*	[31]

Table 5.2. Summary of a workshop used in each project. We describe workshops by their theme, a concise statement of the topics explored. We characterize workshop stakeholders as facilitators or participants categorized by their affiliation as (v)isualization researchers, (c)ollaborators, or (p)rofessional facilitators. Our workshops included 5 – 14 participants and ranged in length from 1/2 to 2 days.

ID	**Theme**	**Facil.**	**Partic.**	**Hrs**
P1	Explore possibilities for enhancing legends with visualizations	1v	3v / 5c	6
P2	Identify future opportunities for utilising smart home data/technologies	2v / 1p	0v / 5c	6
P3	Identify novel visual approaches most suitable for HTA	1v / 1p	7v / 6c	9
P4	Explore shared user needs for visualization in retinal connectomics	4v	0v / 9c	7
P5	Identify analysis and vis. opportunities for improved profiling of cons. prog.	2v / 1c	0v / 10c	7
P6	Understand the main tasks of psychiatric researchers	2v	1v / 6c	3
P7	Explore opportunities for a design study with genealogists	1v	3v / 7c	3
P8	Explore opportunities for funded collaboration between vis. and bio.	1v / 1c	2v / 12c	7x2

defense analysts [9]. In three additional projects, we used a single workshop to jump-start applied collaborations with neuroscientists [2], constraint programmers [29], and psychiatrists [30]. Recently, we used two workshops to explore opportunities for funded collaboration with genealogists [32] and biologists [31].

Within our broad experience, we have focused our analysis on workshops that are used in the early stages of applied work or as the first in a series of workshops. To describe these workshops, we developed the term CVO workshop because such a workshop deliberately and explicitly fosters creativity while exploring opportunities for applied visualization collaborations. We refer to Chapter 2 for a description of where these workshops fit into existing visualization process and decision models.

Focused on CVO workshops, our experience includes the eight workshops in Table 5.2. Since we analyzed more data than appeared in any resulting publications, including artifacts and experiential knowledge, we refer to workshops and their projects by identifiers, e.g., [P1] refers to our collaboration with cartographers. In projects where we used more than one workshop [P1] – [P3], the identifier corresponds to the *first* workshop in the series, unless otherwise specified.

To describe our experience, we developed terminology for the role of researchers involved in each project. The **primary researcher** is responsible f or deciding to use a CVO workshop, executing it, and integrating its results into a collaboration. Alternatively, **supporting researchers** provide guidance and support to the primary researcher. We have been involved with projects as both primary and supporting researchers (see Table 5.1).

We also adopt terminology to describe CVO workshops. Workshops are composed of **methods**, specific repeatable and modular activities [55]. The methods are designed around a **theme** that identifies the workshop's central topic or purpose [21]. The **facilitators** plan and guide the workshop, and the **participants** carry out the workshop methods. Typically, the facilitators are visualization researchers and participants are domain collaborators, but visualization researchers can participate [P1], [P3], and collaborators can facilitate [P5], [P8]. We adopted and refined this vocabulary during our reflective analysis.

5.3 Research Process

The contributions in this chapter arise from *reflection* — the analysis of experiences to generate insights [27], [28]. More specifically, we applied a methodology of *critically reflective practice* [25], summarized by Thompson and Thompson [26] as "synthesizing experience, reflection, self-awareness and critical thinking to modify or change approaches to practice."

We analyzed our collective experience and our CVO workshop data, which consisted of documentation, artifacts, participant feedback, and research outputs. The analysis methods that we used can be described through three metaphorical lenses of critically reflective practice:

- The lens of our collective experience — we explored and articulated our experiential knowledge through interviews, discussions, card sorting, affinity diagramming, observation listing, and observations-to-insights [16]. We codified our experience, individually and collectively, in both written and diagram form. We iteratively and critically examined our ideas in light of workshop documentation and artifacts.
- The lens of existing theory — we grounded our analysis and resulting framework in the literature of creativity and workshops [15], [23], [36], [38], [47], [48], [51], [52], [57], [60], [154] as well as visualization design theory [6], [10], [17], [24].
- The lens of our learners (i.e., readers) — in addition to intertwining our analysis with additional workshops, we shared drafts of the framework with visualization researchers, and we used their feedback to make the framework more actionable and consistent.

Fig. 5.1 shows a timeline of our messy and iterative reflective analysis. The analysis included periods of focused analysis and writing, followed by reflection on what we had written, which spurred additional analysis and rewriting. Over 2 years, we generated diverse artifacts, including models for thinking about how to use workshops, written reflections on which methods were valuable to workshop success, and collaborative writing about the value of workshops. We recorded our reflective analysis in an audit trail that shows our thinking has evolved over the past 2 years. Fig. 5.2 shows two entries from the audit trail that link to relevant research outputs and artifacts of our reflective analysis.

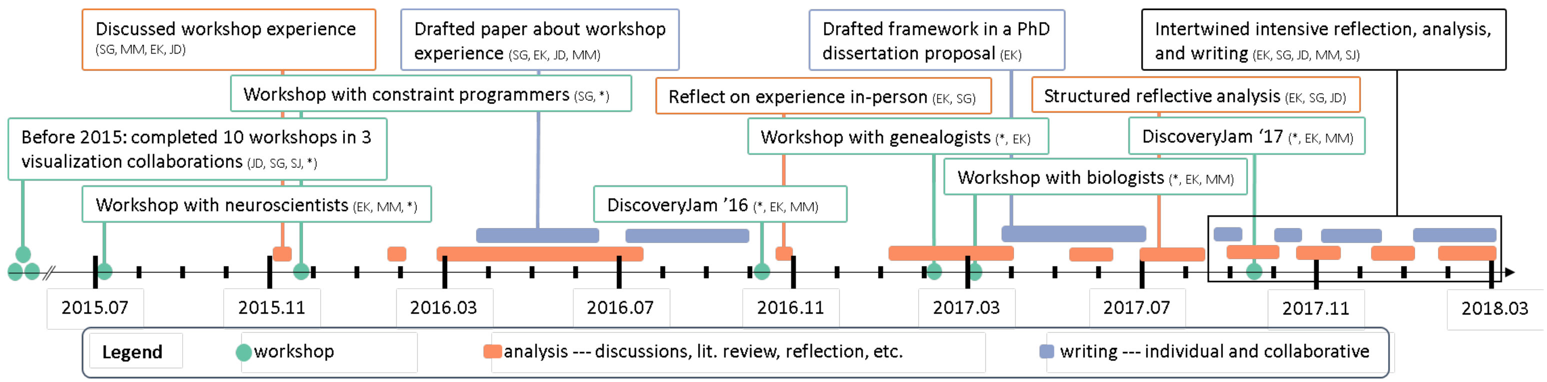

Figure 5.1. Timeline of our reflective analysis. The CVO workshop framework results from a 2-year cross-institution international collaboration that intertwined analysis and writing with additional workshop experience. Throughout this process, we applied a number of analysis methods, including discussions, observation listing, diagramming, and structured written reflection.

Applied workshop with neuroscientists.

2015.06 — 2015.07
workshop

- EK and MM planned, executed, and analyzed a full day workshop with neuroscientists [P4] based on SG's paper [P2].
- The neuroscience workshop inspired development of tools validated through neuroscience and visualization papers.

References

- [P4] Kerzner, E., Lex, A., Sigulinsky, C. L., Urness, T., Jones, B. W., Marc, R.E., Meyer, M. (2017). Graffinity: Visualizing Connectivity In Large Graphs. *Computer Graphics Forum* 36(3).
- [P4] Lauritzen, J.S., Sigulinsky, C. L., Anderson, J. R., Kalloniatis, M., Nelson, N. T., Emrich, D. P., Rapp, C., McCarthy, M., Kerzner, E., Meyer, M., Jones, B. W., Marc, R. E. (2016). Rod-cone Crossover Connectome of Mammalian Bipolar Cells. *Journal of Comparative Neurology*

Applied reflective analysis methods to capture experiential knowledge.

2017.09 — 2017.10
analysis + writing

- Analyzed CVO workshop methods (xlsx) using frameworks for educational and creativity workshops.
- Listed 80+ observations (xlsx) about the process and design of workshops.
  - Connected each observation to our collective experience or existing theory.
  - Considered many approaches to group or rank the observations.
- Applied methods from design — observations-to-insights and insight sorting — to generate a list of actionable insights (pdf).
- Assisted planning and facilitating DiscoveryJam again.

		Intent		Idea Generation			Thought Paradigm			Triggers			Scale		
Method	Sub-method	Develop Trust	Exert Agency	Diverge	Incubation	Converge	Preserve	Bend	Break	Analogy	Iteration		individual	small	large
Opening		■													■
Ice Breaker	*Animal*		■						■						■
Wishful Thinking	*Know/Do/See*	■	■	■			■	■			■		■	■	■
	What Next...		■	■				■	■		■			■	■
Break					■										
Constraint Removal	*Identify*		■				■								■
	Removal			■				■			■			■	■
Lunch					■					■					
Visual Analogies	*Analogies*	■		■					■	■			■		■
	Likes/Dislikes	■	■			■	■	■							
Break					■										
Storyboarding		■	■			■	■	■			■		■	■	■
Closing		■													

Some ideas that do appear in our final framework are shown in this early *matrix of effects* in which we tried to characterize methods by their impact on the workshop ideaspace. After further thinking, some of these characteristics were integrated into the *tactics for effective workshops*.

Figure 5.2. Two example entries from our audit trail. Each audit links to research outputs and reflective artifacts. This dissertation's companion website contains 20 audits that link to more than 30 artifacts, which we have *not* edited for consistency. We have, however, removed some sensitive content.

This dissertation's companion website contains the remainder of the audit trail, including 20 entries that organize and summarize more than 30 reflective artifacts.

Structured written reflection proved particularly useful for generating reflective artifacts that represented our thoughts and opinions about CVO workshops. In this method, one of the paper's coauthors would externalize their ideas into a draft of the framework, which ranged in fidelity from a list of questions to a draft of a paper. We distributed these framework drafts to all of the coauthors along with a list of structured questions, such as "How would you summarize this framework in one sentence?" or "List 3 ideas that you think are missing from the framework and provide concrete details from your experience to support those ideas." As the coauthors responded to these prompts, we created a corpus of reflection documents that captured experiential knowledge from running workshops over the past 10 years. Through further thinking on the reflection documents, we iteratively developed and improved the ideas contained in the framework. As previously stated, we collated and preserved our collective written reflection responses in this dissertation's companion website. Our analysis resulted in the following framework.

5.4 Fundamentals of the Framework

The framework proposed in this dissertation describes how and why to use CVO workshops. We use the term *framework* because what we have created provides an interpretive understanding and approach to practice instead of causal or predictive knowledge [49]. The framework is a thinking tool to navigate the process of planning, running, and analyzing a workshop, but we note that it cannot resolve every question about workshops because the answers will vary with local experience, preference, and context. In this section, we describe a set of factors that contribute to workshop effectiveness, as well as introduce the workshop process model and structure. We intend for the framework to be complemented by existing workshop resources from outside of visualization [21]–[23],[51].

5.4.1 Tactics for Effective Workshops

Reflecting on our experience and reviewing the relevant literature [15], [36], [52], [57], [60] enables us to identify several key factors that contribute to the effectiveness of workshops: focusing on the `topic` of visualization, data and analysis, while fostering, maintaining, and potentially varying the levels of `agency`, `collegiality`, `trust`, `interest`, and `challenge` associated with each. We term these factors **`TACTICs` for effective workshops**:

- (**T**)opic — the space of ideas relevant to data, visualization, and domain challenges in the context of the workshop theme.
- (**A**)gency — the sense of stakeholder ownership in the workshop contributions, outcomes, and the research project.
- (**C**)ollegiality — the degree to which communication and collaboration occur among stakeholders.
- (**T**)rust – the confidence that stakeholders have in each other, the workshop, the design process, and the researchers' expertise.
- (**I**)nterest — the amount of attention, energy, and engagement to workshop methods by the stakeholders.
- (**C**)hallenge — the stakeholders' barrier of entry to, and likelihood of success in, workshop methods.

The `TACTICs` are not independent, consistent, or measurable. The extent to which they are fostered depends upon the context in which they are used, including various characteristics of the workshop — often unknown in advance, although perhaps detectable by facilitators. Yet, selecting methods to maintain appropriate levels of `agency`, `interest`, and `trust` — while varying levels of `challenge` and approaching the `topic` from different perspectives — likely helps workshops to have a positive influence on the mindset of stakeholders and to generate ideas that move forward the methodology of the project. Hence, we refer to the `TACTICs` throughout this framework.

5.4.2 Process Model and Structure

The framework proposes two models for describing how to use CVO workshops: a process model and a workshop structure. The models were adapted from the extensive literature that describes how to use workshops outside of visualization [15], [21]–[23], [48], [51], [155].

The process model (Fig. 5.3 [left]) consists of three stages that describe the actions of using CVO workshops:

1. **Before: define & design.** Define the workshop theme and design workshop methods, creating a flexible workshop plan.

2. **During: execute & adapt.** Perform the workshop plan, adapting it to participants' reactions in light of the `TACTICs`, generating workshop output as a set of rich and descriptive artifacts and documentation.

3. **After: analyze & act.** Make sense of the workshop output and use it in the downstream design process.

Nested within the process is the CVO workshop structure (Fig. 5.3 [right]) that identifies key aspects of the methods used in the beginning, middle, and end of workshops:

1. **Opening.** Establish shared context and `interest` while promoting `trust`, `agency`, and `collegiality`.

2. **Core.** Promote creative thinking about the `topic`, potentially varying `challenge` to maintain `interest`.

3. **Closing.** Provide time for reflection on the `topic` and promote continued `collegiality` in the collaboration.

The process model and structure are closely connected as shown by the orange box in Fig. 5.3. As part of the workshop process, we design and execute a workshop plan. This plan follows the workshop structure because it organizes methods into the opening, core, and closing. In other words, the process is about how we use a workshop; the structure is about how methods are organized within a workshop. We use the process model and structure to organize the following four sections of this chapter. In these sections, we use lists to summarize 25 actionable workshop guidelines. Also, Fig. 5.3 and Appendix A summarize the 25 actionable guidelines in the context of the CVO workshop process model and structure.

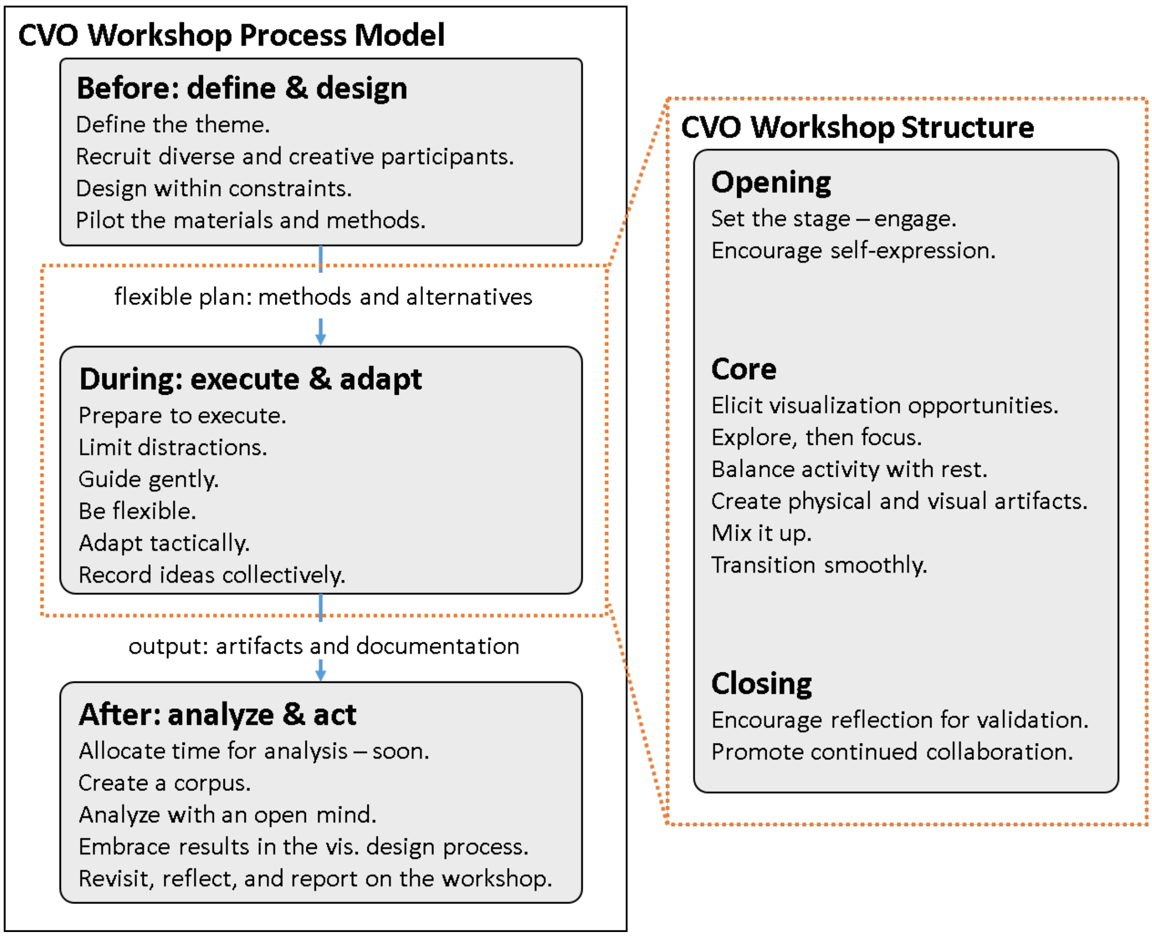

Figure 5.3. Summary of the CVO workshop framework. The framework's two models are 1) a process model (left) that describes the common actions before, during, and after workshops; and 2) a structure (right) that describes principles for methods used in the beginning, in the middle, and at the end of workshops. In these models, we propose 25 guidelines for future workshops.

5.5 Before the Workshop: Define & Design

Creating an effective CVO workshop is a design problem: there is no single correct workshop, the ideal workshop depends on its intended outcomes, and the space of possible workshops is practically infinite. Accordingly, workshop design is an iterative process of defining a goal, testing solutions and evaluating their effectiveness, and improving ideas. The framework we have developed here is part of this process. In this section, we introduce four guidelines — described in a list — for workshop design.

- *Define the theme.* Just as design starts with defining a problem, creating a CVO workshop starts with defining its purpose, typically by articulating a concise theme. An effective theme piques `interest` in the workshop through a clear indication of the `topic`. It encourages a mindset of mutual learning among stakeholders. It also focuses on opportunities that exhibit the appropriate *task clarity* and *information location* of the design study methodology [5]. Examples from our work emphasize visualization opportunities (e.g., "enhancing legends with visualizations" [P1]), domain challenges (e.g., "identify analysis and visualization opportunities for improved profiling of constraint programs" [P5]), or broader areas of mutual interest (e.g., "explore opportunities for a funded collaboration with phylogenetic analysts" [P8]).

  Although we can improve the theme as our understanding of the domain evolves, posing a theme early can ground the design process and identify promising participants.

- *Recruit diverse and creative participants.* We recruit participants who have relevant knowledge and diverse perspectives about the `topic`. We also consider their openness to `challenge` and potential `collegiality`.

  Examples of effective participants include a mix of frontline analysts, management, and support staff [P4]; practitioners, teachers, and students [P5]; or junior and senior analysts [P6]. We recommend that participants attend the workshop in person because remote participation proved distracting in one workshop [P8]. Recruiting *fellow tool builders* [5] as participants should be approached with caution because their perspectives may distract from the `topic` — this happened in our workshop that did not result in active collaboration [P7].

- *Design within constraints.* Identifying constraints can help winnow the possibilities for the workshop. Based on our experience, the following questions are particularly useful for workshop design:

  – Who will use the workshop results? Identifying the primary researcher early in the process is important because he or she will be responsible for the workshop and ultimately use its results. In a workshop in which we did not identify the primary researcher, the results went unused [P7].

  – How many participants will be in the workshop? We typically recruited 5 - 15 participants — a majority domain collaborators but sometimes designers and researchers [P1], [P3], [P6] – [P8].

  – Who will help to facilitate the workshop? We have facilitated our workshops as the primary researcher, with the assistance of supporting researchers or professional workshop facilitators. Domain collaborators can also be effective facilitators, especially if the domain vocabulary is complex and time is limited [P5], [P8].

  – How long will the workshop be? Although we have run workshops that range from 2 hours [P6], [P7] to 2 days [P8], these extremes either feel rushed or require significant commitment from collaborators. We recommend that an effective workshop lasts about 1/2 to 1 working day.

  – Where will the workshop be run? Three factors are particularly important for determining the workshop venue: a mutually convenient location, a high-quality projector for visualization examples, and ample space to complete the methods. We have had success with workshops at offsite locations [P2], [P3], our workplaces, and our collaborators' workplaces [P4] – [P6].

  – What are additional workshop constraints? Examples include the inability of collaborators to share sensitive data [P3], [P6] and the available funding.

- *Pilot the methods and materials.* Piloting (i.e., testing) methods can ensure that the workshop will generate ideas relevant to the `topic` while maintaining appropriate levels of `interest` and `challenge`. We have piloted methods to evaluate how understandable they are [P2], [P4], to test whether they create results that can be used to advance visualization methodologies [P6], [P8], to find mistakes in their prompts [P2], [P4], [P6], [P8], and to ensure that the materials are effective — e.g., sticky notes are the correct size and visualizations are readable on the projector.

It is also useful to pilot workshops with proxy participants, such as researchers [P4] or collaborators [P8]. Feedback from collaborators during pilots has helped us revise the theme, identify promising participants, and refine the workshop methods.

5.6 Workshop Structure and Methods

This section describes guidelines for the methods used in the three phases of the CVO workshop structure (described in Sec. 5.4.2) — the opening, core, and closing. It concludes with summaries of three example workshops and resources for additional workshop methods.

5.6.1 Workshop Opening

The workshop opening communicates the goals and guidelines for participants, but it can be more than that. It can foster `agency` by encouraging self-expression and idea generation. It can encourage `collegiality` and `trust` by promoting open communication, acknowledging expertise, and establishing a safe co-owned environment. It can also garner `interest` by showing that the workshop will be useful and enjoyable. Two guidelines contribute to an effective opening.

- *Set the stage — engage.* CVO workshops typically open with a short introduction that reiterates the theme and establishes shared context for participants and facilitators. We have introduced workshops as “guided activities that are meant to help us understand: what would you like to do with visualization?”[P4]. We have also used graphics that summarize the goals of our project, potentially priming participants to engage with the `topic` of visualization [P3].

  The opening can establish principles for creativity [15], [23], potentially fostering `trust` and `collegiality`. We used the following principles in one of our workshops [P2]: 1) all ideas are valid — express and record them; 2) let everyone have their say; 3) be supportive of others; 4) instead of criticizing, create additional ideas; 5) think 'possibility' – not implementation; 6) speak in headlines and follow with detail; and 7) switch off all electronic devices.

Introduction presentations should be kept short to maintain `interest`. Passive methods, such as lectures and presentations, can discourage participation at the outset. For example, we started one workshop [P8] with a presentation on the current state of analysis tools. This presentation encouraged participants to passively listen rather than actively explore. From this experience, we learned that an effective opening engages participants.

- *Encourage self-expression.* We use methods that encourage self-expression to support interpersonal leveling and to act on the creativity principles — *all ideas are valid* and *be supportive of others*. Such interpersonal methods help to establish an atmosphere of `trust` and `collegiality` among participants and facilitators. They can also provide participants with a sense of `agency` [21].

  We have used interpersonal methods that ask participants to sketch ideas while suspending judgment [34] or to introduce themselves through analogies as a potential primer for creativity (see Sec. 5.6.4). Overall, we use interpersonal methods in the opening to engage participants and facilitators, preparing them for the workshop core.

5.6.2 Workshop Core

In the workshop core, we harness the active and engaged mindset of participants by encouraging them to explore a wide ideaspace before selecting the more promising ideas. The methods in the core potentially generate hundreds of post-it notes, sketches, and other artifacts. Analysis of our experience and relevant literature leads us to suggest five guidelines for an effective core.

- *Elicit visualization opportunities.* We select workshop methods relevant to the `topic` that asks participants about their current analysis challenges, limitations of existing tools, characteristics of their data, or the ways in which they would like to use visualization. In one workshop [P3], for example, we used a method that "developed user stories, considered relevant datasets, discussed alternative scenarios, and sketched solutions" with our domain collaborators. In practically all of our workshop methods, we incorporate ideas about data, visualization, and analysis.

- *Explore, then focus.* We organize the workshop core to first generate ideas using divergent methods that expand the ideaspace. Then, we evaluate ideas using convergent methods that winnow the ideaspace [15]. Using divergent methods early in the core allows us to consider many possibilities while also promoting `agency` and maintaining `interest`. Then, convergent methods can narrow the ideaspace to the more promising ideas.

  Classifying methods as either divergent or convergent risks oversimplification as individual methods often include both divergent and convergent aspects. Consider our use of brainstorming [15] during one workshop [P1]. We asked participants to record "problems and successes associated with the current clients on sticky notes" (divergent) and then to share the more interesting ideas (convergent). We classify this method as divergent because it creates ideas, despite the convergent discussion. In contrast, a convergent method may only involve grouping sticky notes notes from previous methods. Overall, in line with existing workshop guidance [15], [23], [48], [51], we judge methods by their intended impact on the ideaspace and organize the core with phases of divergent and convergent methods.

- *Create physical and visual artifacts.* We select methods by how they encourage participants to write, draw, or otherwise externalize their ideas. Externalizing ideas creates artifacts for us to analyze after the workshop. It aids creative thinking because expressing an idea forces the creator to elaborate it [57] and promotes idea sharing that encourages `collegiality`.

  We consider the artifact materials to be important. Sticky notes are particularly useful because they enable participants to group or rank ideas and potentially to discover emergent concepts in the ideaspace [155]. We have used sticky notes in almost all of our workshops, often using sticky note color to encode information about which method generated an idea and their positions to relate, differentiate, or rank ideas, which can help establish consensus. It can also aid postworkshop analysis by recording how ideas evolved and were valued throughout the workshop. Additional materials effective for externalizing ideas include handouts with structured prompts, butcher paper, and poster boards. Using whiteboards is tempting, but ideas are easily lost if the boards are erased.

We also consider the form of ideas to be important. Effective methods create artifacts relevant to the theme and `topic` of visualization, which can be achieved through the use of visual language (see wishful thinking in Sec. 5.6.4) and by encouraging participants to sketch or draw, such as in storyboarding [P2], [P4], [P5]. We see many opportunities to create useful artifacts with existing methods, such as sketching with data [77] or constructive visualizations [76].

- *Balance activity with rest.* Because continuously generating or discussing ideas can be tiring for participants, we structure workshop methods to provide a balance between activity and rest. Specifically, we incorporate passive methods that provide time for incubation, the conscious and unconscious combination of ideas [57].

  Passive methods can include short breaks with food and coffee, informal discussions over meals, or methods where participants listen to presentations. When using methods that present ideas, asking participants to record their thoughts and reactions can promote `interest` and maintain a feeling of `agency`. We have typically used passive methods in full-day workshops [P2], [P4], [P5], [P8], but we rely on breaks between methods for shorter workshops [P6].

- *Mix it up.* We consider the relationships among methods to be important as we strive to balance exploration with focus and activity with rest, while also using many materials for externalizing ideas. Considering methods that vary these factors can provide different levels of `challenge` because, for example, methods that require drawing ideas may be more difficult than discussing ideas. Using a variety of methods may also maintain `interest` because participants may become bored if too much time is spent on a specific idea.

- *Transition smoothly.* We avoid potentially jarring transitions between methods to preserve participant `interest`. Convergent discussions can be used to conclude individual methods by highlighting the interesting, exciting, or influential ideas. These discussions can promote `collegiality` by encouraging communication of ideas, `agency` by validating participants' contributions, and `interest` in the ideas generated. Convergent discussions also highlight potentially important ideas for researchers to focus on after the workshop.

Convergent methods can also conclude the workshop core by grouping or ranking key ideas. We have used storyboarding to encourage the synthesis of ideas into a single narrative [P2], [P4], [P5]. We have also asked participants to rank ideas, providing cues for analyzing the workshop results [P3]. Convergent methods provide a sense of validation, potentially helping to build `trust` among researchers and collaborators as we transition to the closing.

5.6.3 Workshop Closing

The workshop closing sets the tone for continued collaboration. It is an opportunity to promote `collegiality` by reflecting on the shared creative experience. It also allows for analysis that can potentially identify the more interesting visualization opportunities. The following two guidelines apply to effective closings:

- *Encourage reflection for validation.* We use discussions at the end of workshops to encourage reflection, potentially providing validation to participants and generating information valuable for workshop analysis. We encourage participants to reflect on how their ideas have evolved by asking, "What do you know now that you did not know this morning?" [P5] or "What will you do differently tomorrow, given what you have learned today?" [P2]. Responses to these questions can provide validation for the time committed to the workshop. One participant, for example, reported, "I was surprised by how much overlap there was with the challenges I face in my own work and those faced by others" [P5].

- *Promote continued collaboration.* We conclude the workshop by identifying the next steps of action — continuing the methodology of the collaboration. We can explain how the ideas will be used to move the collaboration forward, often with design methods, as we describe in Sec. 5.8.

We can also ask participants for feedback about the workshop to learn more about their perceptions of visualization and to evaluate the effectiveness of workshop methods — encouraging the visualization mindset. E-mailing online surveys immediately after a workshop is effective for gathering feedback [P4], [P8].

5.6.4 Examples Workshops and Methods

The workshop structure is a thinking tool to help researchers select and organize methods into a coherent and effective CVO workshop. We intentionally omit guidelines that focus on specific methods because the effectiveness of a method depends on local context, preference, and experience. Nevertheless, to illustrate the CVO workshop structure, in this subsection we propose three example CVO workshops and a detailed description of three workshop methods. The workshops and methods described in this subsection, however, are not meant to be complete or exhaustive — we hope that they inspire researchers to think creatively about how to select and organize workshop methods.

Fig. 5.4 summarizes a full-day workshop that we have used successfully in three of our projects [P2], [P4], [P5]. It consists of eight methods that transition smoothly from the workshop opening through the closing. The workshop starts with an opening presentation to establish creativity principles [23], followed by an analogy introduction that promotes interpersonal leveling [8]. Next, wishful thinking elicits opportunities for visualization [8], [156], which are expanded and explored further in the barrier removal method [8], [20]. After that, a break for lunch and an excursion provides time for rest and incubation [8], [47]. After lunch, the visualization analogies method encourages participants to specify requirements by example [8], [14]. Then, storyboarding is used to summarize key ideas in a graphic narrative [157]. Lastly, a reflective discussion highlights potentially interesting ideas for postworkshop analysis [21].

Fig. 5.5 shows two half-day workshops based on methods that we have used successfully a number of times [P2] – [P8], [34]. Fig. 5.5 (left) shows a workshop composed of a subset of the methods described in the full-day workshop. We selected this particular subset of methods because they effectively produced interesting visualization opportunities and fostered rapport among project stakeholders. However, two methods of this workshop may require stakeholders to invest time in preparing for the workshop. For the

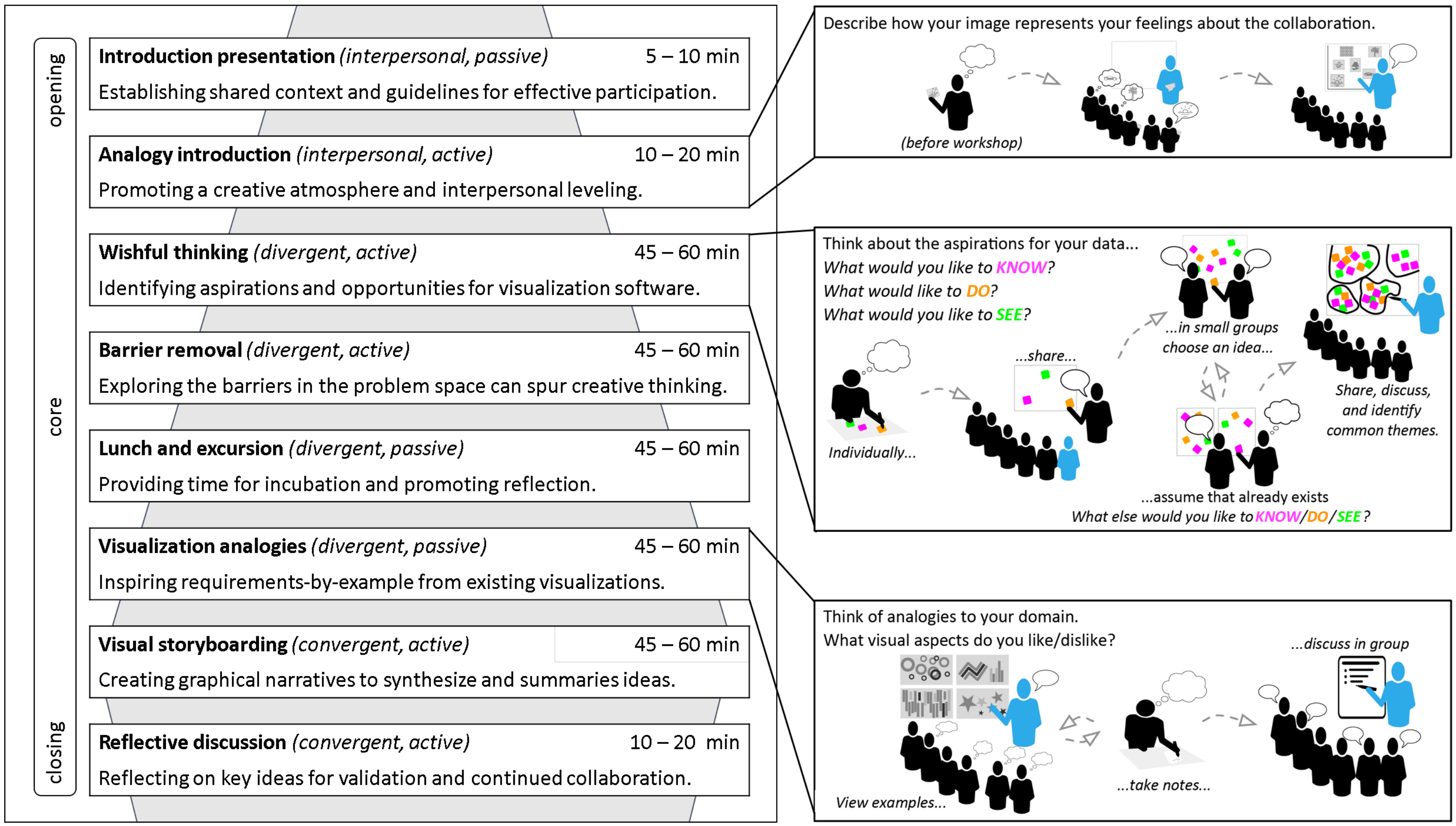

Figure 5.4. Example full-day CVO workshop. The eight methods of the full-day, example CVO workshop (left) with the process of three methods summarized graphically (right). The gray background represents the workshop ideaspace that expands as participants explore ideas, and then contracts to focus on the more promising or interesting ideas.

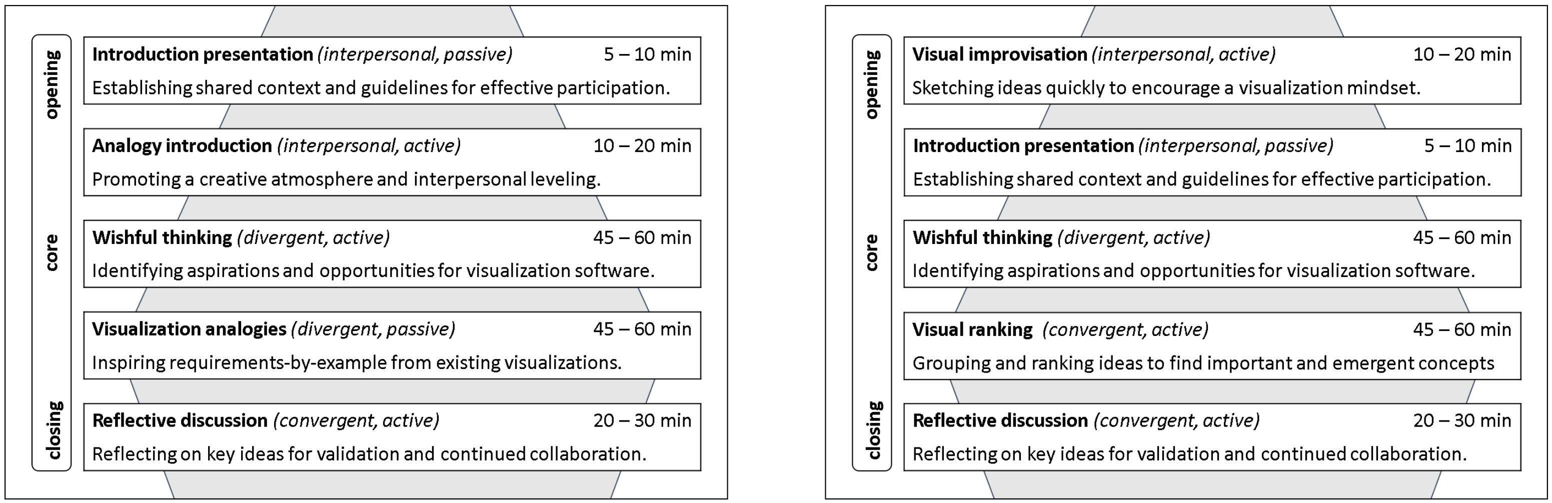

Figure 5.5. Two example half-day CVO workshops. Similar to the full-day workshop in Fig. 5.4, the two half-day workshops explore a broad space of ideas before focusing on the more interesting ideas. We intend for these workshops to serve as a starting place for designing future workshops.

analogy introduction, stakeholders may prepare graphical materials, such as sketches or photographs, that show their feelings about the collaboration. Although this can encourage `agency` among participants, it may also be a `challenge` for participants to introduce themselves with visualization — especially if they have not yet embraced a visualization mindset. Similarly, the visualization analogies method requires that facilitators spend time preparing and rehearsing a presentation of visualization examples to stimulate analogies and discussion. Furthermore, careful attention should be paid during potentially passive methods — the introduction presentation and visualization analogies — to maintain `agency` and `collegiality` in the workshop. Nevertheless, with appropriate preparation and facilitation, the methods in this workshop can elicit a wealth of ideas that move forward collaborations in new and interesting ways. The workshop concludes with a reflective discussion that highlights interesting next steps for the collaboration.

Fig. 5.5 (right) shows a workshop composed of methods that are not part of the full-day example. It opens with visual improvisation that encourages participants to rapidly sketch ideas of increasing complexity [34], which establishes a creative atmosphere and promotes a visualization mindset. This method can be surprising, perhaps encouraging `interest` in the workshop and workshop methods. Next, an introduction presentation can establish focus on the `topic` and workshop theme. After that, wishful thinking can encourage self-expression and elicit visualization opportunities. Then, the visual ranking method encourages participants to group and rank ideas into important concepts or topics [P2], [P3], [P6], [P8]. Visual ranking requires less preparation than visualization analogies, but it may have quite high `challenge` because participants will be asked to think deeply about ideas generated in wishful thinking. We encourage using this method in cases where participants are already familiar with a wide range of visualizations or when facilitators do not have time to prepare a presentation for visualization analogies. Again, this workshop concludes with a reflective discussion. Yet, we encourage researchers to invent new methods to effectively close workshops and set the direction for continued collaboration.

The three example workshops show how methods can be assembled into a coherent workshop, and they summarize methods at a high-level of abstraction. Next, we describe our experiences with three methods that we have found particularly useful in our workshops: analogy introduction, wishful thinking, and visualization analogies. We refer to this dissertation's companion website for details on additional workshop methods. To explain the methods, we refer to their process — the steps of execution [38]. This process description abstracts and simplifies the methods because, during execution, the methods are adapted based on participant reactions and our judgment of the `TACTICs`.

- *Analogy introduction.* We have used this active, interpersonal, and potentially divergent method in the workshop opening. A process of this method, shown in Fig. 5.4 (right, top), starts with a facilitator posing the analogy introduction prompt, e.g., "If you were to describe yourself as an animal, what would you be and why?" [P2]. The facilitators and participants then respond to the prompt in turn — expressing themselves creatively.

  Because everyone responds to the eccentric prompt, this method supports interpersonal leveling, which helps develop `trust` and `collegiality` among stakeholders. Using analogy can prime participants to think creatively [47].

  This method is simple to execute, and participants report that it has a profound impact on the workshop because of the leveling that occurs. The method helps to establish `trust` and that all ideas should be accepted and explored [P4].

  A more topical alternative requires more preparation. We have asked participants to come to the workshop with an image that represents their feelings about the project. Participants have created realistic images, clip-art, and sketches to present and discuss [P3]. A visual analogy introduction can help establish the `topic` of visualization early in the workshop.

- *Wishful thinking.* We have used this divergent, active method early in the workshop core. It is based on creativity methods to generate aspirations [156]. We tailored these methods to visualization by prompting participants with a domain scenario and asking questions: "What would you like to know? What would you like to do? What would you like to see?"

  One process of this method is shown in Fig. 5.4 (right, middle). First, we introduce the prompt, and participants answer the know/see/do questions individually on sticky notes. Next, participants share ideas in a large group to encourage `collegiality` and cross-pollination of ideas. Then, participants form small groups and try to build on their responses by selecting interesting ideas, assuming that they have been completed, and responding to the know/see/do questions again — increasing the `challenge`. Finally, we lead a convergent discussion to highlight interesting ideas and transition to the next method.

We encourage participants to record answers to the know/see/do questions on different color sticky notes because each prompt provides information that is useful at different points in the design process. Participants describe analysis tasks that they would like *to do* or envisaged insights they would like *to know*. Asking what participants would like *to see* is often more of a `challenge` but ensures that a `topic` of visualization is established early.

We tailor the prompt to the workshop theme and project goals. For example, we asked energy analysts about long-term goals for their project — "aspirations for the Smart Home programme..." They generated forward-thinking ideas, e.g., to better understand the value of the data [P2]. In contrast, we asked neuroscientists about their current analysis needs — "suppose you are analyzing a connectome..." They generated shorter term ideas, e.g., to see neuron connectivity [P4].

- *Visualization analogies.* We have used this divergent, initially passive method later in the workshop core because it promotes incubation while allowing participants to specify visualization requirements by example. Similar to analogy-based creativity methods [47] and the visualization awareness method [14], we present a curated collection of visualizations and ask participants to individually record analogies to their domain and to specify aspects of the visualizations that they like or dislike. We have used this method repeatedly, iteratively improving its process by reflecting on what worked in a number of our workshops [P1] – [P5], [P8]. Although this method is primarily passive, participants report that it is engaging and inspiring to see the possibilities of visualization and think about how such visualizations apply to their domain.

  One process of this method is shown in Fig. 5.4 (right, bottom). First, we provide participants with paper handouts that contain a representative image of each visualization. (We have encouraged participants to annotate the handouts, externalizing their ideas [P4], [P5], [P8].) Next, we present the curated visualizations on a projector and ask participants to think independently about how each visualization could apply to their domain and record their ideas. Then, we discuss these visualizations and analogies in a large group.

We curate the example visualizations to increase `interest` and establish participants' `trust` in our visualization expertise. We have used visualizations that we created (to show authority and credibility); those that we did not create (for diversity and to show knowledge of the field); older examples (to show depth of knowledge); challenging examples (to stretch thinking); playful examples (to support engagement and creativity); closely related examples (to make analogies less of a `challenge`); and unrelated examples (to promote more challenging divergent thinking).

The discussions during this method have expanded the workshop ideaspace in surprising ways, including "What does it mean for legends to move?" [P1], "What does it mean for energy to flow?" [P2], and "What does it mean for neurons to rhyme?" [P4]. Because this method is initially passive, it gives participants room to think individually. They reported that it is engaging and inspiring to see the broad possibilities of visualization and discuss how such visualizations apply to their domain.

We provide these three example workshops and three example methods as a starting point for future workshops. Yet, the workshop design space is practically infinite and design should be approached with creativity in mind. To help researchers navigate the design space, this dissertation's companion website contains a list of 15 example methods that we have used or would consider using in future workshops. For these methods, we describe their process, their influence on the workshop ideaspace, their level of activity, and their potential impact on the `TACTICs` for effective workshops.

We have also found other resources particularly useful while designing workshops. These include books [16], [22], [23], [51], [158], [159] and research papers [6], [160], [161]. Although these resources target a range of domains outside of visualization, they can be customized to encourage a visualization mindset and focus on the `topic` of visualization opportunities.

5.7 During the Workshop: Execute & Adapt

Continuing the CVO workshop process model (shown in Fig. 5.3), we execute the workshop plan. This section proposes five guidelines for workshop execution.

- *Prepare to execute.* We prepare for the workshop in three ways: resolving details, reviewing how to facilitate effectively, and checking the venue. We encourage researchers to prepare for future workshops in the same ways.

  We prepare by resolving many details, such as inviting participants, reserving the venue, ordering snacks for breaks, making arrangements for lunch, etc. Brooks-Harris and Stock-Ward [21] summarize many practical details that should be considered in preparing for execution. Our additional advice is simply to promote the visualization mindset in workshop preparation and execution.

  We prepare by reviewing principles of effective facilitation, such as acting professionally, demonstrating acceptance, providing encouragement, and using humor [21]–[23], [51], [162]. We also assess our knowledge of the domain because, as facilitators, we will need to lead discussions. Effectively leading discussions can increase `collegiality` and `trust` between stakeholders as participants can feel that their ideas are valued and understood. In cases where we lacked domain knowledge, we recruited collaborators to help facilitate the workshop [P5], [P8].

  We also prepare by checking the venue for necessary supplies, such as a high-quality projector, an Internet connection (if needed), and ample space for group activity. Within the venue, we arrange the furniture to promote a feeling of co-ownership and to encourage `agency` — a semicircle seating arrangement works well for this [163]. A mistake in one of our workshops was to have a facilitator using a podium, which implied a hierarchy between facilitators and participants, hindering `collegiality` [33].

- *Limit distractions.* Workshops provide a time to step away from normal responsibilities and to focus on the `topic`. Accordingly, participants and facilitators should be focused on the workshop without distractions, such as leaving for a meeting.

Communicating with people outside of the workshop — e.g., through e-mail — commonly distracts participants and facilitators. It should be discouraged in the workshop opening (e.g., "Switch off all electronic devices."). Principles in the workshop opening, however, should be justified to participants. Also, facilitators should lead by example at the risk of eroding `trust` and `collegiality`.

- *Guide gently.* While starting execution, the workshop opening can establish an atmosphere in which participants take initiative in completing methods. It is, however, sometimes necessary to redirect the participants in order to stay focused on the `topic`. Conversations that deviate from the workshop theme should be redirected. In one workshop [P4], participants were allowed to discuss ideas more freely, and they reported in feedback that "We had a tendency to get distracted [during discussions]." In a later workshop [P8], we more confidently guided discussions, and participants reported "We were guided and kept from going too far off track . . . this was very effective."

  However, guiding participants requires judgment to determine whether a conversation is likely to be fruitful. It also requires us to be sensitive to the `TACTICs` — e.g., how would redirecting this conversation influence `collegiality` or `agency`? Redirection can be jolting and can contradict some of the guidelines (e.g., "All ideas are valid"). We may prepare participants for redirection with another guideline during the workshop opening: "Facilitators may keep you on track gently, so please be sensitive to their guidance."

- *Be flexible.* As we guide participants to stay on topic, it is important to be flexible in facilitation. For example, we may spend more time than initially planned on fruitful methods or cut short methods that bore participants.

  Following this guideline can also blur the distinction between participants and facilitators. In one workshop [P3], participants proposed a method that was more useful than what was planned. Thus, they became facilitators for this part of the workshop, which reinforced `agency` and maintained the `interest` of all stakeholders in the project. In the future, we may explore ways to plan this type of interaction, perhaps encouraging participants to create their own workshop methods.

- *Adapt tactically.* As we guide the workshop, we interpret group dynamics and adapt methods to the changing situation. We can be forced to adapt for many reasons, such as a failing method (*nobody feels like an animal this morning*, *sticky notes don't stick*), a loss of `interest` (*there is no energy*; *the room is too hot*; *we had a tough away day yesterday*), a lack of `agency` (*some participants dominate some tasks*), or an equipment failure (*projector does not work*; *no WiFi connection to present online demos* [P2]). Designing the workshop with alternative methods in mind — perhaps with varying degrees of `challenge` — can ensure that workshop time is used effectively.

- *Record ideas collectively.* Remember: conversations are ephemeral and anything not written down will likely be forgotten. We therefore encourage facilitators and participants to document ideas with context for later analysis. Selecting methods to create physical artifacts can help with recording ideas. As described in Sec. 5.6, externalizing ideas on sticky notes and structured prompts has been effective in our workshops and addresses the visualization mindset.

  We are uncertain about the use of audio recording to capture workshop ideas. Although it can be useful for shorter workshops [P6], it can require tremendous time to transcribe before analysis [13]. Also, recording audio effectively can be challenging as participants move around during the workshop.

  Ensuring that facilitators know that they are expected to help document ideas can be useful, and a pilot workshop can help with this. In at least one of our projects [P5], a pilot workshop may have reduced the note taking pressure on the primary researcher during execution.

5.8 After the Workshop: Analyze & Act

After the CVO workshop, we analyze its output and use the results of that analysis to influence the on-going collaboration. In this section, we describe five guidelines for this analysis and action.

- *Allocate time for analysis — soon.* Effective CVO workshops generate rich and inspiring artifacts that can include hundreds of sticky notes, posters, sketches, and other documents. The exact output depends on the methods used in the workshop. Piloting methods can help prepare researchers for the analysis. Regardless, making sense of this output is labor intensive, often requiring more time than the workshop itself. Thus, it is important that we allocate time for analysis, particularly within a day of the workshop, so that we can analyze the workshop output while the experiences are still fresh in our memory.

- *Create a corpus.* We usually start the analysis by creating a digital corpus of the workshop output. We type or photograph the artifacts, organizing ideas into digital documents or spreadsheets. Through this process, we become familiar with key ideas contained in the artifacts. The corpus also preserves and organizes the artifacts, potentially allowing us to enlist diverse stakeholders — such as facilitators and collaborators — in analysis [P4], which can help in clarifying ambiguous ideas or adding context to seemingly incomplete ideas.

- *Analyze with an open mind.* Because the ideas in the workshop output will vary among projects, there are many ways to analyze this corpus of artifacts. We have used qualitative analysis methods — open coding, mindmapping, and other less formal processes — to group artifacts into common themes or tasks [P2], [P4] – [P7]. Quantitative analysis methods should be approached with caution as the frequency of an idea provides little information about its potential importance.

  We have ranked the themes and tasks that we discovered in analysis according to various criteria, including novelty, ease of development, potential impact on the domain, and relevance to the project [P2], [P4] – [P6]. In other cases, workshop methods generated specific requirements, tasks, or scenarios that could be edited for clarity and directly integrated into the design process [P1], [P3].

  We encourage that analysis be approached with an open mind because of the many ways to make sense of the workshop data, including some approaches that we may not yet have considered.

- *Embrace results in the visualization design process.* Similarly, CVO workshop results can be integrated into visualization methodologies and design processes in many ways. We have, for example, run additional workshops that explored the possibilities for visualization designs [P1], [P2]. We have applied traditional user-centered design methods, such as interviews and contextual inquiry, to better understand collaborators' tasks [P4]. We have created prototypes of varying fidelity, from sketches to functioning software [P4] – [P6], and we have identified key aims in proposals for funded collaboration [P8].

In all of these cases, our actions were based on the reasons we ran the workshops, and the workshop results profoundly influenced the direction of our collaboration. For example, in our collaboration with neuroscientists [P4], the workshop helped us focus on graph connectivity, a topic that we were able to explore with technology probes and prototypes of increasing fidelity, ultimately resulting in new visualization tools and techniques (see Chapter 4).

- *Revisit, reflect, and report on the workshop.* The CVO workshop output is a trove of information that can be revisited throughout (and even beyond) the project. It can be used to document how ideas evolve throughout applied collaborations. It can also be used to evaluate and validate design decisions by demonstrating that any resulting software fulfills analysis needs that were identified in the workshop data [P1] – [P6]. Revisiting workshop output repeatedly throughout a project can continually inspire new ideas.

  In our experience creating this dissertation, revisiting output from our own workshops allowed us to discover new insights about how and why to use CVO workshops. We encourage researchers to reflect and report on their experiences using CVO workshops, the ways in which workshops influence collaborations, and ideas for future workshops. We hope this framework provides a starting point for research into these topics.

5.9 Conclusion

This chapter introduced a framework for using CVO workshops in the early, formative stages of applied visualization research. The framework consists of two models for CVO workshops — a process model and a workshop structure. It also identifies 25 actionable guidelines for future workshops and provides three example workshops.

To enhance the actionability of the framework, we provide a companion website that contains 25 pitfalls to avoid in future workshops and 15 example workshop methods. The website also contains 30 documents that show how we created the framework during a 2-year, cross-institution, international collaboration. We provide these resources and the CVO workshop framework as a starting point for researchers who are interested in using workshops in their own collaborations.

CHAPTER 6

DISCUSSION

In this chapter, we first reflect on how CVO workshops influence applied collaborations. Then, we describe the implications and limitations of the CVO workshop framework as well as the research methodology of critically reflective practice that we used to create the framework.

6.1 CVO Workshops in Applied Research

Our experience in diverse domains — from cartography to neuroscience — provides evidence that CVO workshops are a valuable method for advancing visualization methodologies and fostering the visualization mindset among all stakeholders in applied collaborations. In this section, we describe how the use of a CVO workshop can increase the likelihood of successful collaborations, speculate on how a workshop would have influenced our formative design study with defense analysts, and summarize the limitations of CVO workshops.

We speculate that using CVO workshops early in the design process has led to visualization tools that support a wide range of user needs due to the diversity of perspectives that the workshops elicit and explore. Existing visualization process models frame visualization design as a search problem in which researchers should explore a broad space of potential solutions before selecting the more promising ones [5], [6]. Because visualization solutions are closely linked to the problem formulation [24], we argue that using a CVO workshop to explore a broad space of potential domain problems — i.e., visualization opportunities — can increase the likelihood of successful collaborations.

More concretely, CVO workshops create tremendous amounts of artifacts and data about visualization opportunities in a relatively short amount of time. It would require months of interviews for researchers to have the same amount of person-hours with collaborators. Furthermore, based on our experiences with defense analysts [1], researchers spend significant time maintaining engagement when interviews are spread over weeks or months. CVO workshops effectively use the limited time and energy of project stakeholders by providing a forum for focused thinking about the direction of the collaboration. The use of structured methods contributes to the effectiveness of workshops. For example, a fellow visualization researcher described the effect of a workshops on her collaboration with psychiatrists [P6], concluding that "We got a lot more out of the workshop than unstructured meetings" including a better understanding of "what our collaborators were trying to do and what information they needed to do it."

The characteristics engendered by effective CVO workshops — `agency`, `collegiality`, `trust`, and `interest`— are important to the mindset of stakeholders in applied collaborations [5],[17],[61]. Workshops provide a forum that can allow all project stakeholders to contribute to the design requirements by using methods that explicitly encourage `trust` among stakeholders while promoting individual `agency`. These characteristics extend beyond the workshop. For example, in one project [P4], after a successful workshop, members of an academic laboratory were more willing to meet with us regularly and provide us with help accessing and parsing their data. We attribute this shift in accessibility to the workshop experience, and we speculate that it would not have happened using more traditional collaboration techniques [12].

Further thinking about our formative design study with defense analysts suggests that a CVO workshop could have helped us navigate the challenges of working with specialized analysts in a large organization with sensitive data. In the formative design study, we concluded that it is important to 1) sample the perspectives of diverse collaborators involved in data analysis pipelines and 2) recognize the limitations of working with sensitive data by budgeting time to debug visualizations when transitioning from test data to real data [1]. Our reflective analysis suggests that CVO workshops can provide a forum to elicit ideas from collaborators throughout data analysis pipelines [P1] – [P8]. Furthermore, workshops can provide opportunities to identify surrogate data when real data cannot be made public [P3] and to characterize data by scale or relevant attributes [P8]. Therefore, we speculate that using a workshop in collaboration with defense analysts could have more effectively used limited meeting time with collaborators and reduced the time and energy required to piece together perspectives of disparate stakeholders. By spending less time establishing consensus of project stakeholders, we could have used our energy designing and building visualizations, potentially allowing us to create more audacious and transformative visualization tools.

Despite their benefits, workshops may not be appropriate in some scenarios. Because using workshops requires researchers to ask interesting questions and potentially lead discussions about their collaborators' domain, we caution the use of workshops as the first method in a project. Traditional user-centered methods should be used to learn domain vocabulary and explore the feasibility of collaboration. In our project that did not result in an ongoing collaboration [P7], we lacked the domain knowledge needed to effectively design the workshop. Also, our collaborators were too busy to meet with us before the workshop, which should have been a warning about the nature of the project. Accordingly, we recommend researchers evaluate the preconditions of design studies [5] in projects for which they are considering workshops.

We also recognize that workshops may not be well received by all project stakeholders. In a full-day workshop [P4], one participant reported that "Overall, it was good, but a bit long and slightly repetitive." Similarly, after another full-day workshop [P5], one participant said, "There was too much time spent expanding and not enough focus ... discussions were too shallow and nonspecific." Nevertheless, both workshops were generally well received by stakeholders as they allowed us to explore a broad space of visualization opportunities. We can, however, improve future workshops by ensuring that the methods are closely related to the `topic` and that we facilitate workshops in a way that provides appropriate `agency` to all of the stakeholders.

More generally, whether workshops can enhance creativity is an open question [57], [60]. Creativity is a complex phenomenon studied from many perspectives, including design [36], psychology [57], sociology [164], and biology [165]. The results of several controlled experiments indicate that group-based methods can reduce creativity [90],[166]. Yet, critics of these studies argue that they rely on contrived metrics and lack ecological validity [56],[91]. Experimentally testing the relationship between workshops and creativity is beyond the scope of this dissertation. Instead, we approach workshops as reflective practitioners who are trying to understand how to improve existing visualization and workshop practices.

6.2 CVO Workshop Framework

We created the CVO workshop framework to understand and communicate how and why to use CVO workshops in applied research collaborations. In this section, we describe the value of the workshop framework as well as its limitations and implications for future research.

The key value of the framework is that it is a step toward making CVO workshops a repeatable and dependable method. It describes common actions and decisions in the process of using workshops. It also provides guidance for what happens within effective workshops. We have shared drafts of the framework with three researchers who have successfully used it to design, execute, and analyze workshops with collaborators who were interested in virtual reality, music recital composition, and indoor air quality monitoring. One researcher described the framework as "really helpful for designing a [CVO] workshop ... especially for those who aren't lucky enough to have someone like you come in and explain everything to them."

The framework proposes guidelines in which we — a group of five visualization and creativity researchers [3] — agree on the interpretation of our experiences. Nevertheless, we failed to reach consensus, or lacked information, about a number of workshop concepts. For example, we omit a detailed description of what belongs in the workshop plan because it depends on local preference — some of the coauthors preferred highly specific and detailed plans; others preferred more improvisational plans. Both approaches can result in successful workshops. We also omit a guideline about the value of recruiting help from a professional workshop facilitator because we worked with a professional facilitator in only two of our eight workshops. We do not have enough experience or information to create a guideline on this topic. We have also avoided guidelines on specific workshop methods because we intend for the framework to support the adaption and invention of new workshop methods.

More generally, we also did not reach consensus on how to approach workshops used throughout the design process. In some projects [P1] – [P3], we used a series of workshops spaced over weeks or months to allow for development of visualization tools between workshops. Because visualization design is a messy and iterative process [6], it is challenging to describe the intended outcomes of workshops dispersed throughout applied collaborations. We therefore focused explicitly on workshops used in the early, formative stages of applied collaborations or as the first in a series of workshops. This decision allowed us to reach consensus on the intended outcomes of workshops — exploring visualization opportunities. We speculate that many aspects of our framework, e.g., the process model and the `TACTICs` for effective workshops, generalize to workshops used to design prototypes or evaluate completed systems. However, understanding the role of workshops throughout the design process requires more experience, research, and reflection.

While creating the framework, we considered ranking guidelines by the amount of supporting evidence or our perception of their importance, but we discovered that using CVO workshops is more complicated than that. The importance or impact of ideas contained in the framework depends on the context in which they are being applied. Researchers using the framework must account for their own experiences and preferences as well as the larger context of their projects, including the intellectual, interpersonal, and organizational aspects [5],[17]. Because the framework cannot account for every context in which workshops will be used, and the importance of ideas is context dependent, we have avoided ranking specific guidelines. Instead, we encourage researchers to approach the framework as critically reflective practitioners and to use it as a starting point for future CVO workshops.

The CVO workshop framework also provides a language for future research on the use of workshops in applied research. In discussions and feedback on the framework, we have been asked vague questions about future workshops. For example, a number of fellow researchers have asked us about flipping workshops, i.e., allowing collaborators to facilitate the workshop while researchers participate in it. Using terminology from the framework, we could investigate how asking collaborators to facilitate the workshop contributes to their sense of `agency`, the `trust` they have in the research process, or the `collegiality` among all stakeholders. The value of the framework is not that it prescribes a course of action in every situation, but rather that it provides thinking tools to help researchers consider new ways to use CVO workshops.

Lastly, the framework demonstrates the use of critical reflection to learn about visualization in practice. Applied visualization research stresses the importance of reflection to create knowledge that is transferable, rather than generalizable [5]. Yet, it is unclear how we can evaluate and validate knowledge generated through reflection [167]. We contribute the following discussion of critically reflective practice as a step toward understanding the role of reflection in applied visualization.

6.3 Critically Reflective Practice

Throughout this dissertation, we wrestled with a fundamental question: how can we rigorously learn from our diverse, collective experience? We first examined measurable attributes of workshops, such as their length, number of participants, and quantity of ideas generated. However, our workshops were conducted over 10 years in applied settings with no experimental controls. More importantly, it is difficult, if not impossible, to measure how ideas influence collaborations. Quantitative analysis, we decided, would not produce useful knowledge about how to use CVO workshops.

We also considered qualitative research methodologies and methods, such as grounded theory [168] and thematic analysis [169]. These approaches focus on extracting meaning from externalized data, but the most meaningful and useful information about workshops resided in our collective, experiential knowledge. We therefore abandoned analysis methods that ignore (or seek to suppress) the role of experience in knowledge generation.

We found critically reflective practice to be an appropriate approach because it provides a methodology to learn from the analysis of experience, documentation, and existing theory, while allowing for the use of additional analysis methods [25], [26]. Yet, the use of reflective practice may raise questions about this work's validity. After all, can the framework be validated without experimental data?

We emphasize our choice of the term *framework* [49] because we intend for it to be evaluated by whether it provides an interpretive understanding of CVO workshops. Our position is that it achieves this goal because it enables us to learn from our experience using workshops on three continents over the past 10 years. For example, we used the framework to identify and organize 25 pitfalls to avoid in future workshops — they are described in this dissertation's companion website. However, this framework is a snapshot of our current understanding of CVO workshops, which will evolve with continued research, practice, and reflection.

Given that this work results from the subjective analysis of our experience, we recognize questions about its trustworthiness could arise. Therefore, to increase the trustworthiness of our results, we provide an audit trail of our work that contains a timeline of our analysis and our experience as well as diverse artifacts, including comparative analysis of our workshops, presentations outlining the framework, early written drafts of our framework, and structured written reflection to elicit ideas from all of the paper's coauthors [3]. This audit trail — also in the companion website — summarizes and includes 30 of the reflective artifacts that show how our thinking evolved throughout the 2-year reflective collaboration.

In future reflective projects we plan to establish guidelines that encourage transparency of reflective artifacts through mechanisms to flag documents as on or off the record. Because our research and meta-analysis would have been impossible without well-preserved documentation, we hope that the audit trail inspires future thinking on how to document and preserve the decisions in visualization collaborations. We put forth both the audit trail and our documented use of critically reflective practice as contributions of this dissertation.

CHAPTER 7

CONCLUSION

In this chapter, we reiterate the contributions of this dissertation, which arise from two years of reflective analysis and two formative design studies. We conclude by describing interesting areas for future research into CVO workshops.

7.1 Summary

This dissertation is about a framework that describes how and why to use CVO workshops in the early, formative stages of applied visualization research [3]. To create the framework, we analyzed CVO workshops conducted over the past 10 years, on three continents, that were used in collaborations with cartographers [P1], defense analysts [P3], academic researchers [P4], [P5], [P7], [P8], and clinicians [P6]. Because we analyzed CVO workshops from applied collaborations with no experimental controls, it is difficult to establish a causal link between workshops and any specific project outcome. Nevertheless, our diverse experience gives us confidence that CVO workshops can engender engagement and rapport among project stakeholders while providing time for them to think creatively about the directions of a collaboration. The core contribution of this dissertation — a CVO workshop framework — represents our current collective, experiential knowledge about how and why to use CVO workshops.

More specifically, the framework proposes six `TACTICs` for effective CVO workshops and two workshop models: 1) a process model that identifies the common actions before, during, and after CVO workshops; and 2) a workshop structure that describes how methods are organized within effective CVO workshops. To support the two models, we provide 25 actionable guidelines, three example workshops, and detailed descriptions of three workshop methods. To increase the trustworthiness of our results, we also provide — in this dissertation's companion website — an audit trail that summarizes 30 artifacts, showing how our ideas evolved through 2 years of reflective analysis. We hope that the

CVO workshop framework and the audit trail inspire others to use and report on CVO workshops in applied visualization research.

The CVO workshop framework was motivated by and grounded in two formative design studies, which provide a number of this dissertation's secondary contributions. In the first formative design study, we applied user-centered design methods to understand how ballistic vulnerability analysts reason about combined spatial and nonspatial simulation data [1]. This design study contributes task analysis and data abstraction for the domain of vulnerability analysis, as well as ShotViewer, a validated visualization tool that uses linked views to support analysis of ballistic simulation data. Throughout this design study, we spent significant time and energy piecing together an understanding of domain challenges with traditional user-centered design methods. To avoid a similar time sink in our second formative design study, we used a CVO workshop to jump-start a collaboration with neuroscientists, which ultimately resulted in a novel task analysis for analyzing connectivity in large graphs as well as two new visualization techniques and their open-source implementation in a prototype system called Graffinity [2]. Our successful use of a workshop with neuroscientists inspired us to create the CVO workshop framework.

7.2 Future Work

The CVO workshop framework reveals many opportunities for future research, including the development of methods that encourage the visualization mindset or that incorporate data into the workshop. For example, we could encourage the visualization mindset by using methods that promote new ways of thinking about visualization. Inspired by the Dear Data project [170], we could ask participants to create graphics that reveal something about their daily lives in the week before the workshop. The Dear Data Postcard Kit [171] offers possibilities here, providing materials for sharing graphics of data about weekly behaviors.

With respect to using data in CVO workshops, visualization methodologies stress the importance of using real data early in collaborative projects [5], [13]. However, our workshops have focused on participants' perceptions of data rather than using real data because working with data is time consuming and unpredictable. In some projects, we incorporated data into the design process by using a series of workshops spaced over weeks or months, providing time for developers to design prototypes between workshops [P1] – [P3]. This development between workshops was expensive in terms of time and effort. But time moves on, and technologies and approaches that may provide quick and reliable ways of using data in workshops are emerging, such as high-level visualization design tools [172], declarative visualization languages [173], constructive visualization [76], and sketching [77]. We would also like to examine synergies between CVO workshops and design practices from outside visualization, such as design sprints [83] or parallel prototyping approaches [96].

Lastly, this dissertation focuses on workshops to elicit visualization opportunities in the early stages of applied work. Exploring how the framework could be influenced by and extended for workshops that correspond to other stages of applied work — including the creation and analysis of prototypes, the exploration of data, or the deployment, training and use of completed systems — may open up opportunities for creative visualization design and research.

APPENDIX

LIST OF WORKSHOP GUIDELINES

This appendix summarizes the 25 actionable guidelines for CVO workshops, which we present in the context of the CVO workshop process model and structure.

A.1 CVO Workshop Process

- Before: Define & Design
  1. Defined the theme.
  2. Recruit diverse and creative participants.
  3. Design within constraints.
  4. Pilot the methods and materials.
- During: Execute & Adapt
  5. Prepare to execute.
  6. Limit distractions.
  7. Guide gently.
  8. Be flexible.
  9. Adapt tactically.
  10. Record ideas collectively.
- After: Analyze & Act
  11. Allocate time for analysis — soon.
  12. Create a corpus.
  13. Analyze with an open mind.
  14. Embrace results in the visualization design process.
  15. Revisit, reflect, and report on the workshop.

A.2 CVO Workshop Structure

- Workshop Opening
  16. Set the stage — engage.
  17. Encourage self-expression.
- Workshop Core
  18. Elicit visualization opportunities.
  19. Explore, then focus.
  20. Balance activity with rest.
  21. Create physical and visual artifacts.
  22. Mix it up.
  23. Transition smoothly.
- Workshop Closing
  24. Encourage reflection for validation.
  25. Promote continued collaboration.

REFERENCES

[1] E. Kerzner, L. Butler, C. Hansen, and M. Meyer, "A shot at visual vulnerability analysis," *Comput. Graph. Forum*, vol. 34, no. 3, pp. 391–400, 2015.

[2] E. Kerzner, A. Lex, T. Urness, C. L. Sigulinsky, B. W. Jones, R. E. Marc, and M. Meyer, "Graffinity: Visualizing connectivity in large graphs," *Comput. Graph. Forum*, vol. 34, no. 3, pp. 251–260, 2017.

[3] E. Kerzner, S. Goodwin, J. Dykes, S. Jones, and M. Meyer, "A framework for creative visualization opportunities workshops," *IEEE Trans. Vis. Comput. Graphics, to be published.*, vol. 25, no. 1, 2019.

[4] H. Lam, E. Bertini, P. Isenberg, and C. Plaisant, "Empirical studies in information visualization: Seven scenarios," *IEEE Trans. Vis. Comput. Graphics*, vol. 18, no. 9, pp. 1520–1536, 2012.

[5] M. Sedlmair, M. Meyer, and T. Munzner, "Design study methodology: Reflections from the trenches and the stacks," *IEEE Trans. Vis. Comput. Graphics*, vol. 18, no. 12, pp. 2431–2440, 2012.

[6] S. McKenna, D. Mazur, J. Agutter, and M. Meyer, "Design activity framework for visualization design," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 12, pp. 2191–2200, 2014.

[7] J. Dykes, J. Wood, and A. Slingsby, "Rethinking map legends with visualization," *IEEE Trans. Vis. Comput. Graphics*, vol. 16, no. 6, pp. 890–899, 2010.

[8] S. Goodwin, J. Dykes, S. Jones, I. Dillingham, G. Dove, D. Allison, A. Kachkaev, A. Slingsby, and J. Wood, "Creative user-centered design for energy analysts and modelers," *IEEE Trans. Vis. Comput. Graphics*, vol. 19, no. 12, pp. 2516–2525, 2013.

[9] R. Walker, A. Slingsby, J. Dykes, K. Xu, J. Wood, P. H. Nguyen, D. Stephens, B. L. W. Wong, and Y. Zheng, "An extensible framework for provenance in human terrain visual analytics," *IEEE Trans. Vis. Comput. Graphics*, vol. 19, no. 12, pp. 2139–2148, 2013.

[10] M. Tory and T. Moller, "Human factors in visualization research," *IEEE Trans. Vis. Comput. Graphics*, vol. 10, no. 1, 2004.

[11] N. McCurdy, J. Dykes, and M. Meyer, "Action design research and visualization design," in *Proc. Workshop on Beyond Time and Errors on Novel Evaluation Methods for Vis. (BELIV)*. ACM, 2016, pp. 10–18.

[12] J. Wood, R. Beecham, and J. Dykes, "Moving beyond sequential design : Reflections on a rich multi-channel approach to data visualization," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 12, pp. 2171–2180, 2014.

[13] D. Lloyd and J. Dykes, "Human-centered approaches in geovisualization design: Investigating multiple methods through a long-term case study," *IEEE Trans. Vis. Comput. Graphics*, vol. 17, no. 12, pp. 2498–2507, 2011.

[14] L. C. Koh, A. Slingsby, J. Dykes, and T. S. Kam, "Developing and applying a user-centered model for the design and implementation of information visualization tools," in *Proc. Int. Conf. Inform. Vis.* IEEE, 2011, pp. 90–95.

[15] A. Osborn, *Applied Immagination: Principles and Procedures of Creative Problem Solving*. New York, NY, USA: Charle Scribener's Sons, 1953.

[16] V. Kumar and V. LaConte, *101 Design Methods: A Structured Approach to Driving Innovation in Your Organization*. San Francisco, CA, USA: Wiley, 2012.

[17] M. Sedlmair, P. Isenberg, D. Baur, and A. Butz, "Evaluating information visualization in large companies: Challenges, experiences and recommendations," in *Proc. Workshop on Beyond Time and Errors on Novel Evaluation Methods for Vis. (BELIV)*. ACM, 2010, pp. 79–86.

[18] C. Plaisant, "The challenge of information visualization evaluation," in *Proc. Conf. Advanced Visual Interfaces*. ACM, 2004, pp. 109–116.

[19] J. S. Lauritzen, C. L. Sigulinsky, J. R. Anderson, M. Kalloniatis, N. T. Nelson, D. P. Emrich, C. Rapp, N. Mccarthy, E. Kerzner, M. Meyer, B. W. Jones, and R. E. Marc, "Rod-cone crossover connectome of mammalian bipolar cells," *J. Comparative Neurology*, 2016.

[20] S. Jones and N. Maiden, "RESCUE: An integrated method for specifying requirements for complex socio-technical systems," in *Requirements Engineering for Sociotechnical Systems*, J. L. Mate and A. Silva, Eds. Arlington, VA, USA: Information Resources Press, 2005, pp. 245–265.

[21] J. E. Brooks-Harris and S. R. Stock-Ward, *Workshops: Designing and Facilitating Experiential Learning*. Thousand Oaks, CA, USA: SAGE Publications, Inc, 1999.

[22] D. Gray, J. Macanufo, and S. Brown, *Gamestorming: A Playbook for Innovators, Rulebreakers, and Changemakers*. Sebastopol, CA, USA: O'Reilly Media, 2010.

[23] *Creative Problem-Solving Resource Guide*. Scituate, MA, USA: Creative Education Foundation, 2015.

[24] T. Munzner, "A nested model for visualization design and validation," *IEEE Trans. Vis. Comput. Graphics*, vol. 15, no. 6, pp. 921–928, 2009.

[25] S. Brookfield, "Critically reflective practice," *J. of Continuing Edu. in the Health Professions*, vol. 18, no. 4, pp. 197–205, 1998.

[26] S. Thompson and N. Thompson, *The Critically Reflective Practioner*. New York, NY, USA: Palgrave Macmillan, 2008.

[27] D. Boud, R. Keogh, and D. Walker, *Reflection: Turning Experience into Learning*. London, UK: Routledge Taylor and Francis Group, 1985.

[28] D. A. Schon, *The Reflective Practitioner*. New York City, NY, USA: Basic Books, 1988.

[29] S. Goodwin, C. Mears, T. Dwyer, M. Garcia de la Banda, G. Tack, and M. Wallace, "What do constraint programming users want to see? Eexploring the role of visualisation in profiling of models and search," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 1, pp. 281–290, 2016.

[30] C. Nobre, N. Gehlenborg, H. Coon, and A. Lex, "Lineage: Visualizing multivariate clinical data in genealogy graphs," *IEEE Trans. Vis. Comput. Graphics, to be published.*, 2018.

[31] C. Lisle, E. Kerzner, A. Lex, and M. Meyer, "Arbor summit workshop (workshop, University of Utah)," unpublished, 2017.

[32] E. Kerzner, A. Lex, and M. Meyer, "Utah population database workshop (workshop, University of Utah)," unpublished, 2017.

[33] D. H. Rogers, C. Aragon, D. Keefe, E. Kerzner, N. McCurdy, M. Meyer, and F. Samsel, "Discovery Jam," in *IEEE Vis (Workshops)*, 2016.

[34] D. H. Rogers, F. Samsel, C. Aragon, D. F. Keefe, N. McCurdy, E. Kerzner, and M. Meyer, "Discovery Jam," in *IEEE Vis (Workshops)*, 2017.

[35] M. Meyer, M. Sedlmair, and T. Munzner, "The four-level nested model revisited: Blocks and guidelines," in *Proc. Workshop on Beyond Time and Errors on Novel Evaluation Methods for Vis. (BELIV)*. ACM, 2012, pp. 1–7.

[36] B. Shneiderman, G. Fischer, M. Czerwinski, and B. Myers, *NSF Workshop Report on Creativity Support Tools*. National Science Foundation, 2005.

[37] B. Shneiderman, "Creating creativity: User interfaces for supporting innovation," *ACM Trans. on Comput.-Human Interaction*, vol. 7, no. 1, pp. 114–138, 2000.

[38] M. M. Biskjaer, P. Dalsgaard, and K. Halskov, "Understanding creativity methods in design," in *Proc. Conf. Designing Interactive Syst.* ACM SIGCHI, 2017, pp. 839–851.

[39] E. B.-N. Sanders, E. Brandt, and T. Binder, "A framework for organizing the tools and techniques of participatory design," in *Proc. Participatory Des. Conf.*, 2010, pp. 195–198.

[40] G. Dove, "CoDesign with data," PhD Dissertation, City University London, 2015.

[41] S. Jones, N. Maiden, and K. Karlsen, "Creativity in the specification of large-scale socio-technical systems," in *Conf. Creative Inventions, Innovations and Everyday Des. HCI*, 2007.

[42] S. Jones, P. Lynch, N. Maiden, and S. Lindstaedt, "Use and influence of creative ideas and requirements for a work-integrated learning system," in *Int. Requirements Eng. Conf.* IEEE, 2008, pp. 289–294.

[43] N. Maiden, S. Manning, S. Robertson, and J. Greenwood, "Integrating creativity workshops into structured requirements processes," in *Proc. Conf. Designing Interactive Syst.* ACM SIGCHI, 2004, pp. 113–122.

[44] N. Maiden, S. Jones, K. Karlsen, R. Neill, K. Zachos, and A. Milne, "Requirements engineering as creative problem solving: A research agenda for idea finding," in *Int. Requirements Eng. Conf.* IEEE, 2010, pp. 57–66.

[45] N. Maiden, C. Ncube, and S. Robertson, "Can requirements be creative? Experiences with an enhanced air space management system," in *Int. Conf. Software Eng.* IEEE, 2007, pp. 632–641.

[46] J. Horkoff, N. Maiden, and J. Lockerbie, "Creativity and goal modeling for software requirements engineering," in *Proc. Conf. Creativity and Cognition*. ACM, 2015, pp. 165–168.

[47] J. Gordon, William, *Synectics - the Developmnent of Creative Capacity*. New York, NY, USA: Harper and Row, 1961.

[48] E. de Bono, *Lateral Thinking for Management*. Middlesex, UK: Pelican Books, 1983.

[49] Y. Jabareen, "Building a conceptual framework: Philosophy, definitions, and procedure," *Intern. J. of Qualitative Methods*, vol. 8, no. 4, pp. 49–62, 2008.

[50] S. Bennet, *The Workshop Book: New Systems and Structures for Classrooms That Read, Write, and Think*. Portsmouth, NH, USA: Heinemann, 2007.

[51] P. Hamilton, *The Workshop Book: How to Design and Lead Succesful Workshops*. Upper Saddle River, NJ, USA: FT Press, 2016.

[52] K. R. Sawyer, *Group Creativity: Music, Theater, Collaboration*. Mahwah, NJ, USA: Lawrence Erlbaum Associates, 2003.

[53] D. Hornung, *Color: A Workshop for Artists and Designers*. London, UK: Laurence King Publishing, 2012.

[54] Merriam-Webster, *The Merriam-Webster Dictionary New Edition*. Springfield, MA, USA: Merriam Webster Inc, 2016.

[55] M. Crotty, *The Foundations of Social Research*. London, UK: SAGE Publications, Inc, 1998.

[56] R. Mayer, "Fifty years of creativity research," in *Handbook of Creativity*, R. J. Sternberg, Ed. Cambridge, UK: Cambridge University Press, 1999, pp. 449–460.

[57] K. R. Sawyer, *Explaining Creativity - the Science of Human Innovation*. New York, NY, USA: Oxford University Press, 2006.

[58] J. Löwgren, "Applying design methodology to software development," in *Proc. Conf. Designing Interactive Syst.* ACM SIGCHI, 1995, pp. 87–95.

[59] J. Hadamard, *The Psychology of Invention in the Mathematical Field*. New York City, NY, USA: Dover Publications, Inc, 1945.

[60] R. S. Nickerson, "Enhancing creativity," in *Handbook of Creativity*, R. J. Sternberg, Ed. Cambridge, UK: Cambridge University Press, 1999, pp. 392–430.

[61] B. Shneiderman and C. Plaisant, “Strategies for evaluating information visualization tools,” in *Proc. Workshop on Beyond Time and Errors on Novel Evaluation Methods for Vis. (BELIV)*. ACM, 2006, pp. 1–7.

[62] M. Brehmer, S. Carpendale, B. Lee, and M. Tory, “Pre-design empiricism for information visualization,” in *Proc. Workshop on Beyond Time and Errors on Novel Evaluation Methods for Vis. (BELIV)*. ACM, 2014, pp. 147–151.

[63] S. P. McKenna, “Design activity framework: Investigating the data visualization design process,” PhD Dissertation, University of Utah, 2017.

[64] M. Meyer, “Visualization Course Notes,” in *http://www.sci.utah.edu/˜miriah/cs6630/*. The University of Utah, 2013, series Title: http://www.sci.utah.edu/˜miriah/cs6630/.

[65] G. E. Marai, “Activity-centered domain characterization for problem-driven scientific visualization,” *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 1, pp. 913–922, 2018.

[66] J. Blomberg, J. Giacomi, A. Mosher, and P. Swenton-Wall, “Ethnographic field methods and their relation to design,” in *Participatory Design: Principles and Practice*, D. Schuler and A. Namioka, Eds. Hillsdale, NJ, USA: Lawrence Erlbaum Associates, 1993, pp. 122–155.

[67] M. Polanyi, *Personal Knowledge*. Chicago, IL, USA: University of Chicago Press, 1974.

[68] K. Holtzblatt and S. Jones, “Contextual inquiry: A participatory technique for system design,” in *Participatory Design: Principles and Practice*, D. Schuler and A. Namioka, Eds. Hillsdale, NJ, USA: Lawrence Erlbaum Associates, 1993, pp. 177–210.

[69] K. Holtzblatt and H. Beyer, *Contextual Design: Evolved*, J. M. Carroll, Ed. Williston, VT, USA: Morgan & Claypool Publishers, 2015.

[70] R. Sakai and J. Aerts, “Card sorting techniques for domain characterization in problem-driven visualization research,” in *Eurographics Conf. Vis. (Short Papers)*. Eurographics, 2015.

[71] T. Hogan, U. Hinrichs, and E. Hornecker, “The elicitation interview technique: Capturing people’s experiences of data representations,” *IEEE Trans. Vis. Comput. Graphics*, vol. 22, no. 12, pp. 2579–2593, 2015.

[72] J. C. Roberts, C. Headleand, and P. D. Ritsos, “Sketching designs using the five design-sheet methodology,” *IEEE Trans. Vis. Comput. Graphics*, vol. 22, no. 1, pp. 419–428, 2016.

[73] S. McKenna, D. Staheli, and M. Meyer, “Unlocking user-centered design methods for building cyber security visualizations,” in *IEEE Symp. Vis. for Cyber Security (VizSec)*, 2015.

[74] S. Huron, S. Carpendale, J. Boy, and J. D. Fekete, “Using VisKit: A manual for running a constructive visualization workshop,” in *Pedagogy of Data Vis. Workshop at IEEE Vis*, 2016.

[75] S. He and E. Adar, "VizItCards: A card-based toolkit for infovis design education," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 1, pp. 561–570, 2017.

[76] S. Huron, S. Carpendale, A. Thudt, A. Tang, and M. Mauerer, "Constructive visualization," in *Proc. Conf. Designing Interactive Syst.* ACM SIGCHI, 2014, pp. 433–442.

[77] J. Walny, S. Huron, and S. Carpendale, "An exploratory study of data sketching for visual representation," *Comput. Graph. Forum*, vol. 34, no. 3, pp. 231–240, 2015.

[78] D. A. Norman and S. W. Draper, *User Centered System Design; New Perspectives on Human-Computer Interaction*. Hillsdale, NJ, USA: L. Erlbaum Associates Inc, 1986.

[79] M. Muller and S. Kuhn, "Participatory design," *Commun. ACM*, vol. 36, no. 6, pp. 24–28, 1993.

[80] E. B.-N. Sanders and P. J. Stappers, "Co-creation and the new landscapes of design," *CoDesign: Int. J. of CoCreation in Des. and the Arts*, vol. 4, no. 1, pp. 5–18, 2008.

[81] L. Sanders and P. J. Stappers, *Convivial Toolbox: Generative Research for the Front End of Design*. Amsterdam, The Netherlands: BIS Publishers, 2013.

[82] B. Buxton, *Sketching User Experiences: Getting the Design Right and the Right Desing*. San Francisco, CA, USA: Morgan Kaufmann, 2010.

[83] J. Knapp, J. Zeratsky, and B. Kowitz, *Sprint: How to Solve Big Problems and Test New Ideas in Just Five Days*. New York, NY, USA: Simon & Schuster, 2016.

[84] B. Laural, Ed., *Design Research: Methods and Perspectives*. Cambridge, MA, USA: MIT Press, 2003.

[85] B. J. D. Couger, L. F. Higgins, and C. Scott, "(Un)structured creativity in information systems organizations," *MIS Quarterly*, vol. 17, no. 4, pp. 375–397, 1993.

[86] G. Dove, S. Jones, J. Dykes, A. Brown, and A. Duffy, "Using data visualization in creativity workshops: A new tool in the designer's kit," in *Proc. Conf. Creativity and Cognition*. ACM, 2013, pp. 304–307.

[87] G. Dove and S. Jones, "Using data to stimulate creative thinking in the design of new products and services," in *Proc. Conf. Designing Interactive Syst.* ACM SIGCHI, 2014, pp. 443–452.

[88] N. Maiden and S. Robertson, "Developing use cases and scenarios in the requirements process," in *Proc. Intern. Conf. Software Eng.* ACM, 2005, pp. 561–570.

[89] B. Hollis and N. Maiden, "Extending agile processes with creativity techniques," *IEEE Software*, vol. 30, no. 5, pp. 78–84, 2013.

[90] B. Mullen, C. Salas, and E. Johnson, "Productivity loss in brainstorming groups: A meta-analytical integration," *Basic and Appl. Social Psychology*, vol. 12, no. 1, pp. 3–23, 1991.

[91] T. Hewett, M. Czerwinski, M. Terry, J. Nunamaker, L. Candy, B. Kules, and E. Sylvan, "Creativity support tool evaluation methods and metrics," in *NSF Workshop Report on Creativity Support Tools*, 2005.

[92] E. Haug, "Engineering safety analysis via destructive numerical experiments," *EUROMECH 121, Polish Acaemy of Sciences, Eng. Trans.*, vol. 29, no. 1, pp. 39–49, 1981.

[93] R. E. Ball, *The Fundamentals of Aircraft Combat Survivability Analysis and Design*, 2nd ed. Reston, VA, USA: American Institute of Aeronautics and Astronautics, Inc., 2003.

[94] M. Sedlmair, P. Isenberg, and D. Baur, "Cardiogram: Visual Analytics for Automotive Engineers," in *Proc. Conf. Human Factors Comput. Syst.* ACM SIGCHI, 2011, pp. 1727–1736.

[95] P. Drucker, "The coming of the new organization," *Harvard Bus. Rev.*, vol. 66, no. 1, 1988.

[96] S. P. Dow, A. Glassco, J. Kass, M. Schwarz, D. L. Schwartz, and S. R. Klemmer, "Parallel prototyping leads to better design results, more divergence, and increased self-efficacy," *ACM Trans. Comput.-Human Interaction*, vol. 17, no. 4, pp. 1–24, 2010.

[97] L. Institute, "Vision statement: A taxonomy of innovation," *Harvard Bus. Rev.*, vol. 01, 2014.

[98] W. Guber, R. Nagel, R. Goldstein, P. S. Mittelman, and M. H. Kalos, "A geometric description technique suitable for computer analysis of both nuclear and conventional vulnerability of armored military vehicles," Mathematical Applications Group Inc., White Plains, NY, USA, Tech. Rep., 1967.

[99] L. A. Butler and A. Stephens, "Bullet ray vision," in *Proc. Symp. on Interactive Ray Tracing*. IEEE, 2007, pp. 167–170.

[100] W. E. Vesely, F. F. Goldberg, N. H. Roberts, and D. F. Haasl, "Fault Tree Handbook," U.S. Nuclear Regulatory Commission, Washington D.C., USA, Tech. Rep., 1981.

[101] SURVICE Engineering, "The Advanced Joint Effectiveness Model," 2013.

[102] D. F. Keefe, M. Ewert, W. Ribarsky, and R. Chang, "Interactive coordinated multiple-view visualization of biomechanical motion data," *IEEE Trans. Vis. Comput. Graphics*, vol. 15, no. 6, pp. 1383–1390, 2009.

[103] A. G. Landge, J. A. Levine, K. E. Isaacs, A. Bhatele, T. Gamblin, M. Schulz, S. H. Langer, P.-t. Bremer, and V. Pascucci, "Visualizing network traffic to understand the performance of massively parallel simulations," *IEEE Trans. Vis. Comput. Graphics*, vol. 8, no. 12, pp. 2467–2476, 2012.

[104] G. H. Weber, O. Rübel, M.-Y. Huang, A. H. DePace, C. C. Fowlkes, S. V. E. Keränen, C. L. Luengo Hendriks, H. Hagen, D. W. Knowles, J. Malik, M. D. Biggin, and B. Hamann, "Visual exploration of three-dimensional gene expression using physical views and linked abstract views," *IEEE/ACM Trans. Comput. Biol. Bioinform.*, vol. 6, no. 2, pp. 296–309, 2009.

[105] M. Meyer, T. Munzner, A. DePace, and H. Pfister, "MulteeSum: A tool for comparative spatial and temporal gene expression data," *IEEE Trans. Vis. Comput. Graphics*, vol. 16, no. 6, pp. 908–917, 2010.

[106] A. K. Al-Awami, J. Beyer, H. Strobelt, N. Kasthuri, J. W. Lichtman, H. Pfister, and M. Hadwiger, “NeuroLines: A subway map metaphor for visualizing nanoscale neuronal connectivity,” *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 12, pp. 2369–2378, 2014.

[107] H. Doleisch, M. Gasser, and H. Hauser, “Interactive feature specification for focus+context visualization of complex simulation data,” in *Proc. Symp. Data Vis.* Eurographics Association, 2003, pp. 239–248.

[108] A. Maries, N. Mays, M. Hunt, K. F. Wong, W. Layton, R. Boudreau, C. Rosano, and G. E. Marai, “GRACE: A visual comparison framework for integrated spatial and non-spatial geriatric data,” *IEEE Trans. Vis. Comput. Graphics*, vol. 19, no. 12, pp. 2916–2925, Dec. 2013.

[109] M.-W. Chang and C. Collins, “Exploring entities in text with descriptive non-photorealistic rendering,” in *Pacific Vis. Symp.* IEEE, 2013, pp. 9–16.

[110] M. Sedlmair, C. Bernhold, D. Herrscher, S. Boring, and A. Butz, “MostVis: An interactive visualization supporting automotive engineers in MOST catalog exploration,” in *IEEE Conf. Inform. Vis.*, 2009, pp. 173–182.

[111] M. Sedlmair, A. Frank, T. Munzner, and A. Butz, “RelEx: Visualization for actively changing overlay network specifications,” *IEEE Trans. Vis. Comput. Graphics*, vol. 18, no. 12, pp. 2729–2738, 2012.

[112] M. Gleicher, D. Albers, R. Walker, I. Jusufi, C. D. Hansen, and J. C. Roberts, “Visual comparison for information visualization,” *Inform. Vis.*, vol. 10, no. 4, pp. 289–309, 2011.

[113] J. a Ferstay, C. B. Nielsen, and T. Munzner, “Variant view: Visualizing sequence variants in their gene context,” *IEEE Trans. Vis. Comput. Graphics*, vol. 19, no. 12, pp. 2546–2555, 2013.

[114] W. Li, M. Agrawala, B. Curless, and D. Salesin, “Automated generation of interactive 3D exploded view diagrams,” *ACM Trans. Graph.*, vol. 27, no. 3, pp. 1–7, 2008.

[115] B. Johnson and B. Schneiderman, “Tree-Maps: A space-filling approach to the visualization of hierarchical information structures,” in *Proc. Conf. Vis.* IEEE, 1991, pp. 284–291.

[116] E. R. Gansner and S. C. North, “An open graph visualization system and its applications to software engineering,” *Software — Practice & Experience*, vol. 30, no. 11, pp. 1203–1233, 2000.

[117] I. Wald, S. Woop, C. Benthin, G. S. Johnon, and M. Ernst, “Embree: A kernel framework for efficient CPU ray tracing,” *ACM Trans. Graph.*, vol. 33, no. 4, pp. 1–8, 2014.

[118] T. Munzner, *Visualization Analysis and Design*. Boca Raton, FL, USA: CRC Press, 2015.

[119] J. L. Hennessy and D. A. Patterson, *Computer Architecture: A Quantitative Approach*, 5th ed. Burlington, MA, USA: Morgan Kaufmann, 2011.

[120] K. Madhavan, N. Elmqvist, and M. Vorvoreanu, "DIA2: Web-based cyberinfrastructure for visual analysis of funding portfolios," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 1, pp. 1823–1832, 2014.

[121] B. Lee, C. Plaisant, C. S. Parr, J.-D. Fekete, and N. Henry, "Task taxonomy for graph visualization," in *Proc. Workshop on Beyond Time and Errors on Novel Evaluation Methods for Vis. (BELIV)*. ACM, 2006, pp. 1–5.

[122] S. van den Elzen and J. J. van Wijk, "Multivariate network exploration and presentation: From detail to overview via selections and aggregations," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 12, pp. 2310–2319, 2014.

[123] A. J. Pretorious, H. C. Purchase, and J. T. Stasko, "Tasks for multivariate network analysis," in *Multivariate Network Visualization*, ser. Lecture Notes in Computer Science, A. Kerren, H. C. Purchase, and M. O. Ward, Eds. Berlin, Heidelberg, Germany: Springer, 2013, vol. 8380, pp. 79–95.

[124] A.-L. Barabasi, *Network Science*. Cambridge, UK: Cambridge University Press, 2016.

[125] B. Shneiderman and A. Aris, "Network visualization by meaningful substrates," *IEEE Trans. Vis. Comput. Graphics*, vol. 12, no. 5, pp. 733–740, 2006.

[126] M. Ghoniem, J. D. Fekete, and P. Castagliola, "On the readability of graphs using node-link and matrix-based representations: A controlled experiment and statistical analysis," *Inform. Vis.*, vol. 4, no. 2, pp. 114–135, 2005.

[127] B. Aleman-Meza, C. Halaschek-Wiener, I. B. Arpinar, C. Ramakrishnan, and A. P. Sheth, "Ranking complex relationships on the semantic web," *IEEE Internet Comput.*, vol. 9, no. 3, pp. 37–44, 2005.

[128] P. Heim, S. Hellmann, J. Lehmann, S. Lohmann, and T. Stegemann, "RelFinder: Revealing relationships in RDF knowledge bases," in *Semantic Multimedia*, ser. Lecture Notes in Computer Science, T. Chua, Y. Kompatsiaris, B. Mérialdo, W. Haas, G. Thallinger, and W. Bailer, Eds. Berlin, Heidelberg, Germany: Springer, 2009, vol. 5887, pp. 182–187.

[129] Y. Zhang, G. Cheng, and Y. Qu, "RelClus: Clustering-based relationship search," in *Intern. Semantic Web Conf.*, vol. 1035, 2013, pp. 1–4.

[130] Y. Tu and H. W. Shen, "GraphCharter: Combining browsing with query to explore large semantic graphs," in *Pacific Vis. Symp.* IEEE, 2013, pp. 49–56.

[131] C. Partl, S. Gratzl, M. Streit, A. M. Wassermann, H. Pfister, D. Schmalstieg, and A. Lex, "Pathfinder: Visual analysis of paths in graphs," *Comput. Graph. Forum*, vol. 35, no. 3, pp. 71–80, 2016.

[132] Y. Yang, T. Dwyer, S. Goodwin, and K. Marriott, "Many-to-many geographically-embedded flow visualisation: An evaluation," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 1, pp. 411–420, 2016.

[133] T. von Landesberger, A. Kuijper, T. Schreck, J. Kohlhammer, J. J. V. Wijk, J. Fekete, and D. W. Fellner, "Visual analysis of large graphs: State-of-the-art and future research challenges," *Comput. Graph. Forum*, vol. 30, no. 6, pp. 1719–1749, 2011.

[134] F. Beck, M. Burch, S. Diehl, and D. Weiskopf, "A taxonomy and survey of dynamic graph visualization," *Comput. Graph. Forum*, vol. 36, no. 1, pp. 133–159, 2017.

[135] A. Kerren, H. C. Purchase, and M. O. Ward, Eds., *Multivariate Network Visualization*, ser. Lecture Notes in Computer Science. Berlin, Heidelberg, Germany: Springer, 2014, vol. 8380.

[136] D. Auber, "Tulip: A huge graph visualisation framework," in *Graph Drawing Software*, ser. Mathematics and Visualization, P. Mutzel and M. Junger, Eds. Berlin, Heidelberg, Germany: Springer, 2004, pp. 105–126.

[137] M. Bastian, S. Heymann, and M. Jacomy, "Gephi: An open source software for exploring and manipulating networks," in *Intern. Conf. Weblogs Social Media*. Association for the Advancement of Artificial Intelligence, 2009, pp. 361–362.

[138] J. Webber, "A programmatic introduction to Neo4j," in *Proc. Conf. Syst. Programming Appl.: Software for Humanity*. ACM, 2016, pp. 217–218.

[139] N. Henry and J.-d. Fekete, "MatLink: Enhanced matrix visualization for analyzing social networks," in *Human-Computer Interaction*, ser. Lecture Notes in Computer Science, C. Baranauskas, P. Palanque, J. Abascal, and S. Barbosa, Eds. Berlin, Heidelberg, Germany: Springer, 2007, vol. 4663, pp. 288–302.

[140] Z. Shen and K.-l. Ma, "Path visualization for adjacency matrices," in *Proc. Conf. Vis.* Eurographics / IEEE, 2007, pp. 83–90.

[141] J. Zhao, M. Glueck, F. Chevalier, Y. Wu, and A. Khan, "Egocentric analysis of dynamic networks with EgoLines," in *Proc. Conf. Human Factors Comput. Syst.* ACM SIGCHI, 2016, pp. 5003–5014.

[142] F. van Ham, H.-j. Schulz, and J. M. Dimicco, "Honeycomb: Visual analysis of large scale social networks," in *Human-Computer Interaction*, ser. Lecture Notes in Computer Science, T. Gross, Ed. Berlin, Heidelberg, Germany: Springer, 2009, vol. 5727, pp. 429–442.

[143] M. Wattenberg, "Visual exploration of multivariate graphs," in *Proc. Conf. Human Factors Comput. Syst.* ACM SIGCHI, 2006, pp. 811–819.

[144] J. P. Bagrow, E. M. Bollt, J. D. Skufca, and D. Ben-Avraham, "Portraits of complex networks," *Europhysics Letters*, vol. 81, no. 6, 2007.

[145] W. Czech, "Graph descriptors from B-matrix representation," in *Graph-Based Representations in Pattern Recognition*, ser. Lecture Notes in Computer Science, X. Jiang, M. Ferrer, and A. Torsello, Eds. Berlin, Heidelberg, Germany: Springer, 2011, vol. 6658, pp. 12–21.

[146] S. Kairam, D. MacLean, M. Savva, and J. Heer, "GraphPrism: Compact visualization of network structure," in *Proc. Intern. Conf. Advanced Visual Interfaces*. ACM, 2012, pp. 498–505.

[147] N. Elmqvist, T. N. Do, H. Goodell, N. Henry, and J. D. Fekete, "ZAME: Interactive large-scale graph visualization," in *Pacific Vis. Symp.* IEEE, 2008, pp. 215–222.

[148] M. Behrisch, B. Bach, N. Henry Riche, T. Schreck, and J. D. Fekete, "Matrix reordering methods for table and network visualization," *Comput. Graph. Forum*, vol. 35, no. 3, pp. 693–716, 2016.

[149] J.-d. Fekete, "Reorder.js : A JavaScript library to reorder tables and networks," in *IEEE Vis (Posters)*, 2015.

[150] J. R. Anderson, S. Mohammed, B. Grimm, B. W. Jones, P. Koshevoy, T. Tasdizen, R. Whitaker, and R. E. Marc, "The Viking Viewer for connectomics: Scalable multi-user annotation and summarization of large volume data sets," *J. Microscopy*, vol. 241, no. 1, pp. 13–28, 2010.

[151] J. R. Anderson, B. W. Jones, C. B. Watt, M. V. Shaw, J.-H. Yang, D. DeMill, J. S. Lauritzen, D. M. P. K. Yanhua Lin, Kevin D. Rapp, B. Grimm, T. Tasdizen, R. Whitaker, and R. E. Marc, "Exploring the retinal connectome," *Molecular Vision*, vol. 17, pp. 355–379, 2011.

[152] D. Fisher, I. Popov, S. Drucker, and M. Schraefel, "Trust me, I'm partially right: Incremental visualization lets analysts explore large datasets faster," in *Proc. Conf. Human Factors Comput. Syst.* ACM SIGCHI, 2012, pp. 1673–1682.

[153] F. Van Ham and A. Perer, ""Search, show context, expand on demand": Supporting large graph exploration with degree-of-interest," *IEEE Trans. Vis. Comput. Graphics*, vol. 15, no. 6, pp. 953–960, 2009.

[154] W. C. Miller, *The Creative Edge: Fostering Innovation Where You Work*. New York City, NY, USA: Basic Books, 1989.

[155] G. Dove, S. Julie, M. Mose, and N. Brodersen, "Grouping notes through nodes: The functions of post-it notes in design team cognition," in *Des. Thinking Res. Symp.*, Copenhagen Business School, 2016.

[156] M. J. Hicks, *Problem Solving and Decision Making: Hard, Soft, and Creative Approaches*. London, UK: Thomson Learning, 2004.

[157] K. N. Truong, G. R. Hayes, and G. D. Abowd, "Storyboarding: An empirical determination of best practices and effective guidelines," in *Proc. Conf. Designing Interactive Syst.* ACM SIGCHI, 2006, pp. 12–21.

[158] L. Hohmann, *Innovation Games: Creating Breakthrough Products Through Collaborative Play*. Boston, MA, USA: Addison-Wesley, 2007.

[159] M. Michalko, *Thinkertoys: A Handbook for Creative-Thinking Techniques*. Emeryville, CA, USA: Ten Speed Press, 2006.

[160] E. McFadzean, "The creativity continuum:towards a classification of creative problem solving techniques," *J. of Creativity and Innovation Manage.*, vol. 7, no. 3, pp. 131–139, 1998.

[161] E. B.-N. Sanders, "Information, insipiration, and co-creation," in *Conf. European Academy of Des.*, 2005.

[162] R. B. Stanfield, *The Workshop Book: From Individual Creativity to Group Action*. Gabriola Island, BC, Canada: New Society Publishers, 2002.

[163] R. S. Vosko, "Where we learn shapes our learning," *New Directions for Adult and Continuing Edu.*, vol. 50, no. Summer, pp. 23–32, 1991.

[164] T. I. Lubart, "Creativity across cultures," in *Handbook of Creativity*, R. J. Sternberg, Ed. Cambridge, UK: Cambridge University Press, 1999, pp. 339–350.

[165] C. Martindale, "Biological bases of creativity," in *Handbook of Creativity*, R. J. Sternberg, Ed. Cambridge, UK: Cambridge University Press, 1999, pp. 137–152.

[166] T. J. Bouchard, "Personality, problem-solving procedure, and performance in small groups," *J. Appl. Psychology*, vol. 53, no. 1, pp. 1–29, 1969.

[167] J. Dykes, M. Meyer, U. Hinrichs, N. Henry Riche, R. Chang, P. Isenberg, H. Lam, and T. Munzner, "Reflection on Reflection in Design Studies," in *IEEE Vis (Workshops)*, 2017.

[168] J. Corbin and A. Strauss, "Grounded theory research: Procedures, canons, and evaluative critera," *Qualitative Sociology*, vol. 13, no. 1, pp. 3–21, 1990.

[169] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative Res. Psychology*, vol. 3, no. 2, pp. 77–101, 2006.

[170] G. Lupi and S. Posavec, *Dear Data: The Story of a Friendship in Fifty-Two Postcards*. London, UK: Penguin, 2016.

[171] ——, *Dear Data Postcard Kit: For Two Friends to Draw and Share (Postcards)*. New York City, NY, USA: Princeton Architectural Press, 2017.

[172] K. Wongsuphasawat, D. Motitz, J. Mackinlay, B. Howe, and J. Heer, "Voyager: Exploratory analysis via faceted browsing of visualization recommendations," *IEEE Trans. Vis. Comput. Graphics*, vol. 22, no. 1, pp. 649–658, 2016.

[173] A. Satyanarayan, D. Moritz, and K. Wongsuphasawat, "Vega-Lite: A grammar of interactive graphics," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 1, pp. 341–350, 2017.