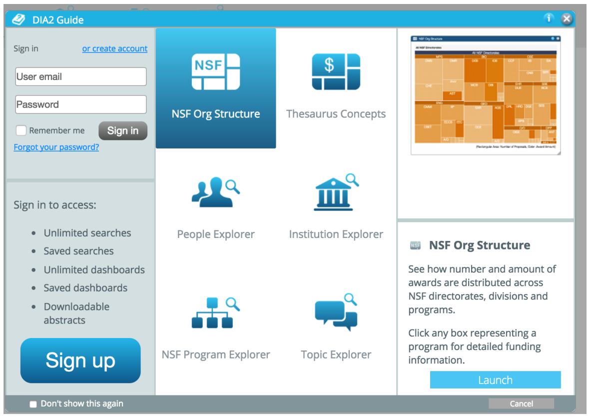
Towards Characterizing Domain Experts as a User Group

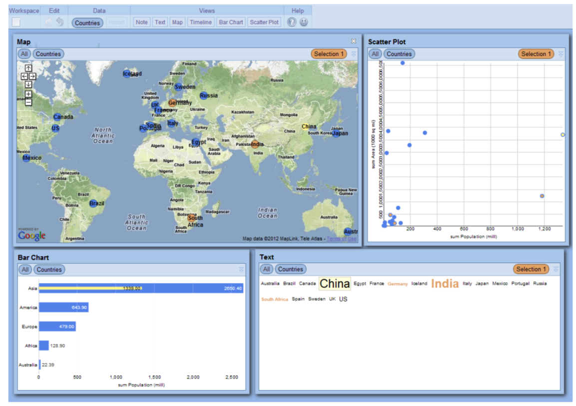
Yuet Ling Wong* Purdue University

Krishna Madhavan† Microsoft Corporation

Niklas Elmqvist‡ University of Maryland, College Park



(a) DIA2: A web-based visual analytics system complies with the philosophy of "no manual, no training" [30].



(b) Choosel: A web-based framework visualization construction tool reflects the design guidelines for domain experts [19].

Figure 1: Two examples of visualization tools designed for domain experts lacking specific visualization expertise. All such tools share common characteristics in minimizing training and instruction, and maximizing accessibility.

Abstract

Visualization is an inherently interdisciplinary research area: visualization researchers are always visualizing other people's data. While recent trends in the field has turned toward a more inclusive audience, particularly within the topic of "visualization for the masses", the traditional user group for our tools have been the *domain expert*: people who are experts in a specific professional domain where they want to apply visualization and analytics, but who often lack high literacy, training, and motivation in visualization and visual analytics. Such domain experts want to opportunistically reap the benefits of visualization, but have no patience for long training, poor interaction or visual design, or complex displays. While domain experts are a familiar user group, surprisingly little effort has been devoted towards characterizing them for both design and evaluation purposes. To help the visualization community better understand this specific group, in this paper, we describe the characteristics of domain experts, discuss existing examples designed for them, and propose possible guidelines to facilitate the design process. We believe this discussion will help visualization researchers better understand this group and uncover more research opportunities.

Keywords: Visualization theory, concepts, evaluation.

1 Introduction

Traditional visualizations were custom software designed for a specific purpose, dataset, and audience, which meant that these tools often required a significant visualization and analytics expertise on behalf of the user. For example, ThemeRiver [21] for time-series data, Trellis Displays [7] for multivariate data, and Bubble Treemaps [18] for hierarchical data, are complex visualization techniques that demand non-trivial visualization literacy [11] on behalf of the users. In other words, these tools historically required the users to not only be *domain experts*, i.e., have expertise in a specific discipline, but also have the time and motivation to become visualization experts. Therefore, applying visualization to scientific fields such as biology, geography, and social science typically required users to become familiar with sophisticated and esoteric visualization techniques. However, acquiring this visualization literacy requires users to spend time for training even if they have no real interest to do so.

Fortunately, recent years have seen the introduction of *casual information visualization* [37] as a new sub-domain of visualization research that designs "visualization for the masses," i.e., where the target audience is more inclusive and involves casual users. Examples include InfoCanvas [50] that maps data to elements in an image, Vizster [22] that creates node-link diagrams of social networks with a user-centric view, and ManyEyes [52], an online visualization tool that allows users to create their desired visualizations by uploading their datasets. These casual information visualizations "*depict personally meaningful data in a visual way that supports users in both everyday life and non-work situation*" [37]. However, a casual user has little interest or time for training or manuals. In order to appeal to such a fickle and easily distracted audience with no intrinsic loyalty to visualization as a tool, designers have to improve their tools so that they are easy, engaging, and self-explanatory to use.

The inevitable backlash from this development, however, has been to make the traditional user group of visualization software—the *domain expert*—much less tolerant of poorly designed interfaces, esoteric visual representations, and questionable encodings. Put differently, domain experts are increasingly eschewing training and manuals as well. They now want to enjoy the fruits of visualization and analytics, but lack the time, training, and motivation to learn complex visualization and analytics techniques. Such unique characteristics require the research community's special attention to design and evaluate effective visualization tools. However, little emphasis

*e-mail: wong64@purdue.edu
†e-mail: krishna.madhavan@microsoft.com
‡e-mail: elm@umd.edu

Table 1: Characteristics of different user groups.

Dimensions \ User Groups	Visualization Experts	Casual Users	Domain Experts
Visualization Literacy	High	Low/None	Low/None
Domain Expertise	Low/None	Low/None	High
Data Analysis Training	High	Low/None	Varies

so far has been dedicated to characterize domain experts as a user group. To address this gap, here we first describe the varying user groups for visualization. Then we elaborate on the characteristics of domain expert and derive several possible design guidelines based on known examples of visualization tools designed for such users.

This work is relevant to the BELIV workshop because it characterizes a user group that requires particular care during the design process. We hope our work can help visualization researchers better understand this user group.

2 BACKGROUND

Information visualization and visual analytics technique have penetrated into every aspect of our daily lives since its emergence. For example, FundExplorer [15] is a visual system that implements a distorted treemap to visualize a persons fund portfolio and market stocks. It enables people to interactively explore diversification possibilities in their stock market portfolio. CopLink [13] is an information and knowledge management system that helps police officers to capture, access, analyze, visualize, and share law enforcement-related information in social and organizational contexts. Even the webpages we browse daily are full of visualizations such as NameVoyager [55], which illustrates the popularity of baby names over time using stacked time visualization, Google Maps, which allows users to search and browse geographic space, and Climate Lab[1], which visualizes weather data from days to years.

Because visualization has been widely used in so many domains, it covers all kinds of users, from engineers to researchers, novices to experts. Based on the work in visualization research, we can identify three target user groups: visualization experts, casual users, and domain experts. Although many possible dimensions can be used to characterize those user groups, we pick three that can be easily derived from their definition: their visual literacy, domain expertise, and data analysis training. Table 1 shows how these three user groups are characterized.

2.1 Visualization Experts

Visualization experts, as its name states, are people who have professional knowledge in data visualization. As stated by Windhager and Smuc, visualization experts "*are proven specialists and usually have the broad domain knowledge required to justify an interpretation of the results. Since data exploration is (part of) their job, they share an intrinsic motivation to explore data and hunt for new insights*" [56]. For example, visualization researchers use their own visualization knowledge to analyze existing visualizations in order to design new visualizations or construct frameworks as references for other researchers.

Huang et al. investigated 59 personal visualization (PV) and personal visual analytics (PVA) papers and developed a taxonomy of design dimensions to provide a coherent vocabulary for discussing PV and PVA [23]. In industry, although there exists many visualization tools such as Microsoft Excel, Tableau[2], and Spotfire[3] that provide a set of powerful visualizations for data exploration, data scientists still need to apply their professional knowledge to choose suitable visualization for their data analysis tasks. In addition, user experience designers need to be familiar with and conduct studies upon all kinds of visualizations to learn about users' experience for visualization improvement. Lam et al. reviewed over 800 evaluation publications and presented seven scenarios to distinguish different study goals and types of research questions of evaluation methods [28]. All of these examples require visualization knowledge to accomplish their tasks.

2.2 Casual Users

Walsh and Hall [54] describe a casual user as someone "*who has just stumbled across [the digital] collection in the same way that they would wander into the [cultural heritage] institution's physical space*". Casual users in visualization "*do not analyze the data in a detailed way, but become aware of basic patterns, gain a feeling for the data, and reflect on its social and personal relevance*", as described by Pousman et al. [37]. Based on these descriptions, we define **casual users** as users who are not trained in data analysis, visualization, or statistics, but use visualization for casual purpose such as social activities, entertainment, or just for curiosity.

Visualizations designed for this group of users are called *casual visualizations* [37] or *personal visualizations* [23]. In recent decades, this area has attracted many researchers' attention and much work has been done to design useful and attractive visualization for casual users. For instance, Data Memes [14] is an artistic visualization that users can merge charts of their personal data within an image. FinVis [41] is a visual analytic tool that helps casual users to interpret the return, risk, and correlation of financial data and make decisions. In addition, LastHistory [5] visualizes users' music listening histories for analysis and reminiscing purposes. All of these kinds of visualizations are interesting, easy to use, and understandable to attract casual users' attention, lower their usage hurdle, and enjoy the aesthetics of visualizations.

2.3 Domain Experts

In Grammel's dissertation [19], he defines visualization novices as "*users who create visualizations to support their primary tasks, but who are typically not trained in data analysis, information visualization, and statistics*". However, since users' goal of creating visualizations is to support their primary tasks, we note that the visualization novices he defines are actually domain experts.

Here we extend his definition to our *domain expert* definition: **Domain experts** are users who **create or consume** visualizations to support their primary tasks, but are typically not trained in data analysis, visualization, and statistics. This kind of users includes police officers who are experts in public safety, sociologists whose job involves analyzing social phenomena, and logistics experts who schedule and deliver goods in desired times. All of these users have advanced expertise in their own domains, but typically have little knowledge in visualization techniques. Much has been done for this user group that we discuss in Section 4. In next section we give a detailed discussion of domain experts' characteristics.

3 CHARACTERISTICS OF DOMAIN EXPERTS

Domain experts have many characteristics in common with casual users, including having basic knowledge of commonly used visualizations such as histograms, bar charts, and line graphs, yet lacking

[1]http://www.climate-lab-book.ac.uk

[2]https://www.tableau.com/

[3]https://spotfire.tibco.com/

specialized visualization skills as well as the time and motivation to learn novel visualizations. However, they also have their own specific characteristics: primarily being experts in a specific domain, being willing to try the techniques that can alleviate their workload, and being restricted to a specific working environment. In this section, we are going to discuss these characteristics in details.

3.1 Strengths

Domain experts have many strengths and skills that casual users—and even visualization experts—lack.

- **Subject matter experts:** Since domain experts are familiar with the knowledge in their expertise domains, they have clear understanding of the attributes, relationships, and possible operations of the data they are analyzing. So while designing visualization tools to support their tasks, visualization designers can focus on the problem-solving aspect of visualizations other than aesthetic aspects, as for casual users who may just browse the data or "*do not analyze the data in a detailed way*" [31].

- **Stronger reading skills in visualization than casual users:** By virtue of being professionals, domain experts typically have stronger skills in reading visualizations in their specific domain than casual users. This has been demonstrated by an eye-tracking experiment of comparing two groups of students viewing nuclear magnetic resonance (NMR) spectroscopic signals and choosing the corresponding molecular structure from the candidates. [51]. The results show that domain experts illustrated more efficient scan patterns in the experiment than the novice group. Although domain experts have strong visualization skills in their domain, they are still not visualization experts, which must be considered in the design process [47].

- **Certain level of learning ability in domain-specific visualization:** Gegenfurtner and Seppänen [17] conducted a mixed method study that displayed three different computer-based imaging technologies—a familiar, a semi-familiar, and an unfamiliar imaging technology—to domain experts and recorded their performance. Their results show that experts can transfer their process from the familiar to semi-familiar, but not to the unfamiliar imaging technology. This indicates domain experts have certain abilities to transfer their domain-specific visualization knowledge to new visualizations.

Overall, while designing visualization for this group of users, visualization researchers should consider all of the strengths discussed above so that the visualizations contain suitable functionalities that can fully support the task requirements.

3.2 Weaknesses

Although domain experts' strengths provide many advantages for visualization designers, their weaknesses and missing skills also bring many challenges that cannot be ignored.

- **Time limitation:** The number one priority for most domain experts is to achieve their daily tasks, and they often lack the time to invest in any extraneous activities. As a rule, the more specialized the expert, the more demands they have on their time. For example, Sedlmair et al. explain that in large companies [45], experts "*work under heavy time pressure and are bound to strict deadlines.*" These time limitations have several impacts. First, domain experts often lack the time to even participate in research studies since the payoff from such is not immediate. According to Sedlmair et al., spending time with visualization researchers "*is an extraordinary task without necessarily direct evidence of impact on their actual work tasks.*". Similarly, for the persona Amy created in the DIA2 [30] project which is designed for government organizations, she is highly qualified for her job, however, "*she spends most of her day acting like a human search engine, manually parsing search results from databases that are difficult to query*".

  Even given a finished visualization tool, many domain experts have little time to spend on manuals or training. Again, such training would be seen as an extraordinary task that would increase to their workload. Worse, after they learn the technique, they may find that the technique does not help them because, e.g., they may need to "*double-check against older, more trusted tools to ensure the results are accurate*" [30], which increases their workload further.

- **Lack of motivation:** This weakness is caused by the time limitation and conventional tool attachment. This is particularly problematic in large companies where employees are highly specialized in a small subset of a highly specific collaborative task set, and are very accustomed to and effective with traditional tools to finish their tasks in required time [45]. "*This effectiveness naturally leads to attachment to the traditional tool and results in a reluctance to learn a new systems*" [45]. If a new visual technique is introduced, they would lose their work efficiency, or even become stripped of their respected expert status [45].

- **Environmental constraints:** Although this weakness is not a characteristic specific to domain experts, visualization designers still need to take it into account when designing visualization tools for them. In company or government organization settings, there are many work environment constraints that visualization researchers need to overcome. For example, confidentiality of information will prevent researchers from video-, audio- and screen-recording while designing visualizations for domain experts [45]. Confidentiality constraints may also increase the difficulty for visualization researchers to publish their research [30].

  In addition, it is often challenging to convince stakeholders to adopt new visualization tools since it involves the time-profit tradeoff that stakeholders most care about [45]. Moreover, designing visualization for domain experts can be time-sensitive because the whole design process is often long [12, 47], so that when the design cycle ends, domain experts' requirements may have already changed. Finally, since domain experts are used to their daily work process, integrating new tools in this daily process involves technical, political, and organizational issues such as integrating new visualization tools to existing analysis process, authorization of software, and the amount of bureaucracy involved for this design process.

4 Supporting Examples

In order to come up with possible design guidelines that take advantage of domain experts' strengths and alleviate their weaknesses, we conducted a literature survey to analyze existing examples from visual design aspect. Since design studies have become increasingly popular in designing visualization for domain experts, we conducted our literature survey from the examples mentioned in Sedlmair et al.'s design study methodology paper [47], the papers that cite this methodology paper, and other papers from TVCG (including VIS papers) that target users are domain experts. Relevance was based on meeting all of the following criteria:

- The target users must be domain experts;

- Experts' domain knowledge of each paper should be different so that as many domains as possible can be covered; and

Table 2: Design dimensions and corresponding examples from the literature.

Examples	Design Layouts	Visual Rep.	Interaction Tech.	Help Info
DIA2 [30]	multi single views/customized	simple/advanced	direct/indirect	guide/place holder/tooltips
VASA [25]	multi vis in one view/fixed	simple/advanced	direct/indirect	linked
BallotMaps [57]	multi vis in one view/fixed	simple	–	–
VisMOOC [49]	multi vis in one view/fixed	simple/advanced	direct/indirect	tooltips
TimeLineCurator [16]	multi vis in one view/fixed	simple/advanced	direct/indirect	highlight/linked
VisOHC [26]	multi vis in one view/fixed	simple/advanced	direct/indirect	highlight/tooltips
RelEx [44]	multi single views/customized	simple/advanced	direct/indirect	linked
Autobahn [46]	multi single views/customized	simple/advanced	direct/indirect	linked
MostVis [43]	multi single views/fixed	simple/advanced	direct/indirect	highlight/color/linked
code_swarm [35]	multi vis in one view/fixed	simple/advanced	–	–
PowerSetViewer [34]	multi vis in one view/fixed	simple/advanced	direct/indirect	text/linked
Choosel [19]	multi single views/customized	simple	direct/indirect	tooltip/color/linked
dotlink360 [3]	multi single views/customized	simple/advanced	direct/indirect	color/linked
MizBee [33]	multi vis in one view/fixed	simple/advanced	direct/indirect	highlight/linked
Vismon [10]	multi vis in one view/fixed	simple/advanced	direct/indirect	color/linked
TiMoVA [9]	multi vis in one view/fixed	simple	direct/indirect	color/linked
Poemage [32]	multi vis in one view/fixed	simple/advanced	direct/indirect	color/linked
WeaVER [39]	multi vis in one view/fixed	simple/advanced	direct/indirect	color/linked
SimilarityExplorer [36]	multi single views/customized	simple/advanced	direct/indirect	color/linked
VisInfo [8]	multi single views/customized	simple	direct/indirect	highlight

- The goal of the designed visualization should be to support domain experts rather than mere visual aesthetics.

This resulted in an initial set of 20 papers (clearly not a complete list). The purpose of our survey is to characterize a subset of possible design guidelines among the visualization in most domains. Therefore, the papers in our list contain visualization tools designed for governments, industry, and science. We coded these papers using four design dimensions: design layouts, visual representation types, interaction techniques, and types of help information. Table 2 summarizes these dimensions and corresponding examples from the literature and Fig. 2 demonstrates how these dimensions reflect in the examples.

Design Layouts From the literature, we detected two major design layouts: multiple separated views where each view contains one visual representation, and one single view that contains several visual representations. For the multiple separated view layout, the position, size, or visibility of each view can be customized by users. Some of the views may only appear after users' specific operations. For the single view layout, there is usually a main visualization that consumes a large portion of the space and is accompanied by several small visualizations. In our list of related work, 8 papers that adopt the multiple view layout, while 12 adopt the other. Although there are two different layouts, they are just different forms of containers for the visualizations that provide different flexibilities for the users; both contain multiple visualizations.

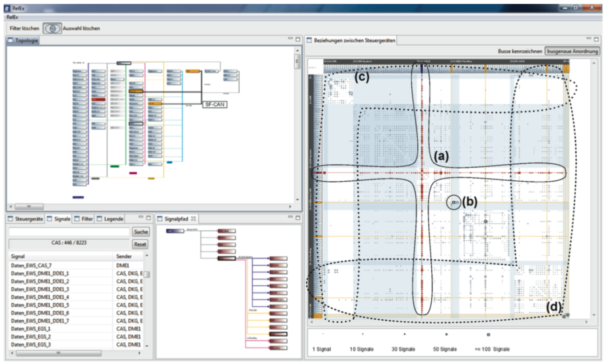
Visual Representations We make a distinction between "simple" visualizations such as bar chart, line graph, and histogram that are commonly seen in classic statistical graphics. Other visualizations outside this list are coded as "advanced" visualizations. We counted the numbers of simple and advanced visualizations for each example and found that at least one simple visualization is used in each example and four of them only contain simple visualizations.

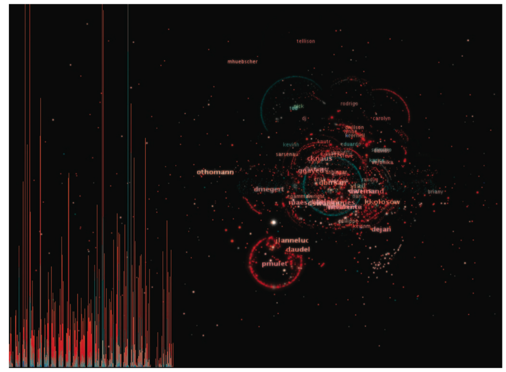
Interaction Techniques Interaction techniques provide the ways for users to interact with the visualizations and further explore the data. These techniques include direct operations on the visualization such as select, pan, zoom, and so on, and indirect operations such as searching for the input items, filtering the data according to specific criteria, and sorting the data items in the visualizations. In our paper list, except for two examples, all others deploy both types of interaction techniques. For those two examples without interactions, BallotMaps [57] is used to show the electoral result of each candidate to detect if there is name bias in alphabetically ordered ballot papers, while code_swarm [35] displays how the code documents evolve as committers change the code over time. Both of them can achieve the goals without the support of interaction techniques.

Help Information Help information provide guidance for how a visualization can be viewed and interacted with. This includes separated user guides and any kinds of visual elements that provide hints for the users. In our survey, only one example—DIA2 [30] (Figure 1(a)) provides a guide widget that describes the function of each widget. Other examples embed different visual elements in the visualization to provide hints to users. These visual elements include placeholder to promote what to input, brushing technique to link different views, legends to illustrate the meaning of elements in the visualization, and so on.

5 DESIGN GUIDELINES

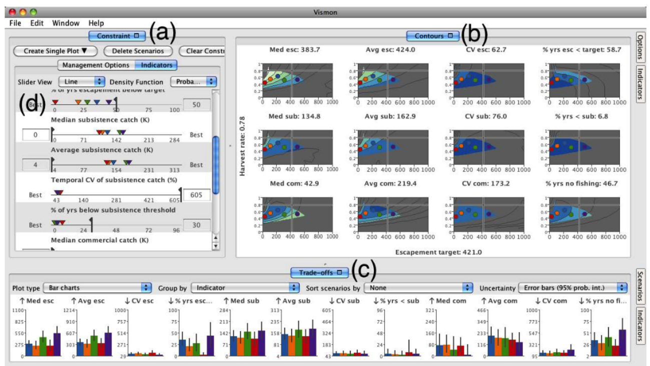
It is challenging to formulate generic design guidelines for domain experts since the actual domains often vary widely and have correspondingly varied requirements. Requirements can even be different for the same domain given different tasks. For example, in the automotive industry examples, three tools (RelEx [44], Autobahn [46], and MostVis [43]) are designed to accomplish different tasks for automotive engineers. Nevertheless, Sedlmair et al. [47] propose a methodological framework and provide practical guidance for conducting **design studies**, which constitute a specific process of designing visualization mainly for domain experts. This design study framework has nine stages that are classified into three top-level phases: precondition phase, core phase, and analysis phase. Each stage can also go backward to previous stages according to current stage's result. Moreover, some stages in the precondition phase and analysis phase are time-consuming since they involve meeting and interviewing with domain experts and conducting user studies, and it is usually hard to find a common time for all experts, delaying the whole process.





(a) RelEx: multiple views to show traffic patterns of in-car communication networks [44]. (Source: Sedlmair et al., 2012)

(b) code_swarm: shows the code document changes over time [35]. (Source: Ogawa and Ma, 2009)



(c) Vismon: multiple views to support data analysis of simulation results [10]. (Source: Booshehrian et al., 2012)

Figure 2: Three visualization examples demonstrate different design layouts, visual representations, and interaction techniques.

To accelerate the design study process, we can reduce the time cost in the core phase which contains four stages: discover, design, implement, and deploy. The purpose of the discover stage is to discover the needs, problems, and requirements of domain experts and to characterize and abstract the problems. The time cost in this stage can be different due to the domain problems we need to address. While the purpose of the "implement" and "deploy" stage is the implementation and deployment of the visualizations, its time cost is also varied and depends on the developers. So the only possible stage that we can improve is the design stage by providing some design guidelines that researchers can follow, and thus save time. Therefore, based on the examples we analyzed above, we characterize several possible guidelines from the coded dimensions to help researchers design visualizations for domain experts. However, due to the limited number (20) of examples in our literature list, it is clear that our design guidelines are only a subset of the possible guideline space. In addition, because domain experts work in different domains, their specific terminology, tasks, and workflows are different. Therefore, our design guidelines are expressed in a general form in order to fit the requirements across all domains.

We propose the following design guidelines (and describe them in detail in the following subsections):

DG1 Refactor complex views into manageable components;

DG2 Design visualizations with domain experts' terminologies;

DG3 Progressively refine exploration to promote learning; and

DG4 Embed help information in the display to minimize training.

5.1 Refactor Complex Views

As the tasks for domain experts become complicated, embedding all the information in one single visualization would make the visualization too complex. This may overloads user perception and mislead their decisions. To address this complexity problem, one possible solution is to refactor those complex views into multiple manageable components to keep each component as simple and illustrative as possible that each of them only presents parts of the information rather than the whole. In our survey examples, all of them contain more than one visualization, regardless of whether it is in multiple or single view layouts. Another reason to adopt this design is that those manageable components can be organized to match domain expert workflows [12] and adapt different customs.

However, multiple visualization components can make users lose the navigation between views and the global sense of the data. In order to provide connection between different components, brushing visual techniques [6] as a direct manipulation technique mentioned by Roberts [40] can be applied so that users can still have general idea of the whole information while they are navigating in one view. This brushing technique has already been applied to Keshif [58], a systematic design to simplify authoring of dashboards composed of summaries and individual records from raw data using fluid interaction. As the guidelines developed by Baldonado and Kuchinsky [2] and the study conducted by Qu and Hullman [38] show, keeping consistency between multiple views is important to help users to interpret the visualizations. This brushing technique provides a

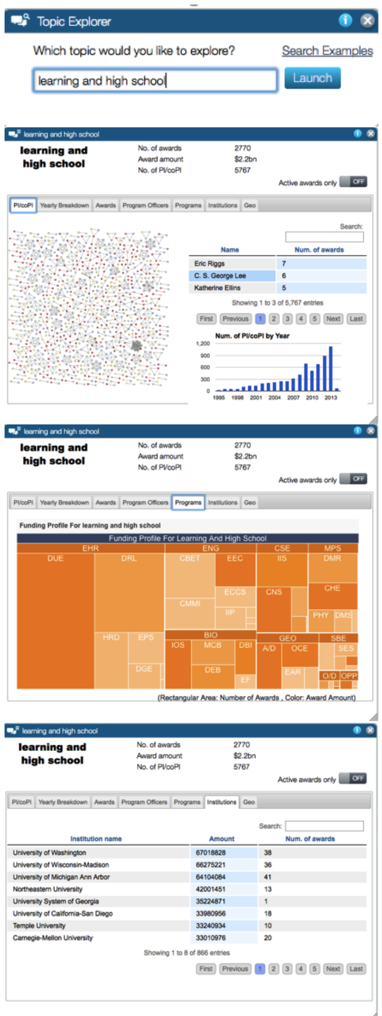


Figure 3: A progressive refinement process of the DIA2 platform shows how users explore the data starting from searching a term [30].

concise design that can remedy domain experts' characteristics of lacking visual literacy by reducing their perception load.

Based on the genres defined by Segel and Heer [48] for narrative visualization, multiple component visualization falls in the Partitioned Poster genre that users can begin the exploration with the view most attractive to them. As the examples we choose in Fig. 2 show, there are two of them have a view larger than other views in their multiple view design. This salient design can help users focus on the larger view to start their exploration for their first time use.

5.2 Design Visualizations Using Domain Terminologies

As mentioned in Grammel et al.'s study about how domain experts construct visualization, they found that "*participants strongly preferred visualization types that they are familiar with*" [20]. This finding can also be applied to domain experts who only consume visualizations. However, how can we define familiar visualization? One idea is to define familiar visualization as those that users commonly see in their daily used software. However, based on the result shown in the "Visualization Techniques" column of Table 2, we find that except four examples, all other examples contain both simple and advanced examples. This finding contradicts our initial hypotheses.

After further analyzing the design flows of all the examples, we determine that familiar visualizations are visualizations designed based on a domain terminology. They do not have to be simple visualization from classic statistical graphics. As described in the DIA2 example [30], domain experts can "*use their knowledge of the organizations structure to infer the meanings associated with block size and color saturation on the treemap visualization even when none of them seemed to be familiar with visualizations*". Another example is how Brehmer et al. redesigned the Energy Manager system with "*the juxtaposition of a matrix and a boxplot, two unfamiliar encodings, together with coordinated interaction and highlighting, received more positive feedback than either of these encodings in isolation*" [12].

This familiar visualization also implies that users tend to favor those visualizations that match their mental models since the schematic, semantic or item-specific information preserved in mental models allow users to construct and simulate a problem and thereby aid reasoning as argued by Liu & Stasko [29]. Because of the "lack of motivation" of domain experts, using familiar visualizations that match their mental models can lower their learning barrier, attract their attention, and as a result provide them with the motivation to understand how to use the visualization tool.

5.3 Progressively Refine Visuals to Promote Learning

As recent studies in visualization accuracy conducted by Wakeling et al. [53] show, if users' preferred visual representation is not the correct visual form, their performance is degraded. Thus, it is necessary to use even unfamiliar visualization in a visualization tool if they represent the correct visual form that is best suited to the task. To help domain experts learn these unfamiliar visualizations, interaction techniques can play an important role by progressively introducing the unfamiliar visual representation.

We propose a *progressive reveal* as a step-by-step learning process where domain experts can transfer their existing visualization knowledge to learn more advanced visual techniques. As a case in point, Kwon and Lee [27] conducted a study investigating the efficacy of multimedia learning environments for data visualization education where including an interaction tutorial helped users understand the visualization better and have a more engaging experience. Applying interaction techniques to show more and more details as the user explores simulates the process of storytelling that can attract users' interest, and, as a result, promote the learning process.

In practice, progressive reveal starts with a familiar visualization that a person knows, and then the visual representation is transformed in a step-by-step manner into something rich and more complex. For example, a parallel coordinate display can be explained in a step by step manner starting from a scatterplot through moving the horizontal axis to vertical axis and changing the dots to lines (shown in Fig. 6). Provided users know the original state, and since each change is a simple animation, users should be able to understand the final state too. Yoon et al. [59] conducted eye-tracking studies and found that this "progressive revealing" technique does improve task accuracy and other measures of performance. This design also falls in the Interactive Slideshow structure defined by Segel and Heer [48] that allows users to explore particular points of a visualization or proceed to more detailed parts. Fig. 3 illustrates a progressive refinement process of the DIA2 platform that shows how users explore the data starting from searching a term step by step.

From the existing examples, we find several interaction techniques are used for this refinement process. These techniques include clicking on an object in the visualization that usually leads to a more detailed view or a totally different visualization of the data; clicking on an option to filter data and refresh the visualization; searching for a specific item which would be highlighted on the visualization; dragging an item to the visualization to create connections with existing data; and filtering operation to reduce cluttering of visualizations. Table 2 summarizes the interaction techniques used in each example.

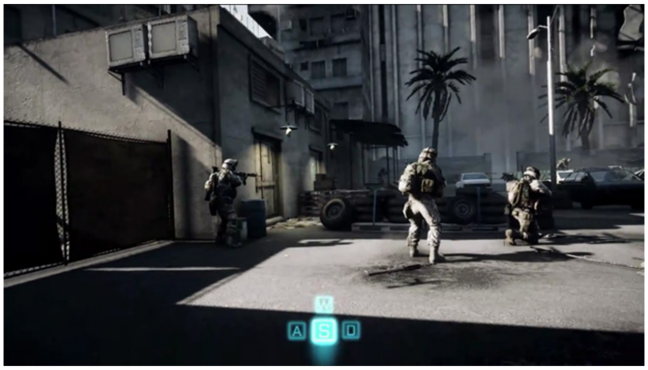


Figure 4: Help information to remind the player how to perform a move. (Source: *Battlefield 3* by DICE)

5.4 Embed Help to Avoid Training and Manuals

Besides the interaction techniques we discussed to support progressive refinement of visualizations, there are also some other visual elements that we label as help information. User guides and text are the two most commonly used visual elements that provide direct help information, which require users to spend a little bit time for the reading, and sometimes may be ignored by users, as mentioned in the evaluation of the DIA2 project [30]. Tooltips, besides extending the visual space to provide more information, is also another type of help information that only displays as needed to provide hints. Color is usually used to categorize different types of objects. Finally, highlight provides connections between objects in the same view or between different views, as in the brushing technique discussed in DG 1. In general, help information can be classified using the design strategies presented by Segel and Heer for narrative visualization [48]. For example, tooltip falls in the category of *details-on-demand*, user guide, text, and placeholder fall in the category of *annotation*, and color and highlight fall in the category of *visual highlighting*.

Embedding help information in the display can remind users how to explore the visualization when they need it to avoid training and manuals. This guideline was inspired from computer games, which are typically very complex with many time-sensitive controls, but where casual players get constantly reminded of how to achieve complex tasks when it is contextually relevant. As we can see in Fig. 4, which is from *Battlefield 3* by DICE, the player is reminded of which keys to press using an on-screen instruction when the game character is supposed to be moving forward. This help information also brings in the "just in time" (JIT) concept [24] that instructions are to be shown in the context when they are needed, and only then. This means that users do not see the help information when they are not needed. This form of contextual help will not increase the visual clutter of the display and will thus not affect the resulting cognitive load on the user. Our guideline condenses the domain expert's learning time for new visualization since they can use their existing knowledge for the visual parts they already know and learn the parts they do not know from help information.

6 A Step-by-step Example Using Design Guidelines

It is hard to create manageable components without knowing domain experts' terminologies and workflows. For this reason, in this step-by-step example, we simplify the steps of DG1 and DG2, supposing that we already have each manageable component needed (DG2) under the scenario of designing a visual tool for car engineers to evaluate how different car attributes affect its performance. Since we already have those necessary components, what we need first is to figure out a design layout to organize those components so that they are manageable by domain experts. Here "manageable" means that the positions and the orders of those components match the workflows of the domain experts. According to DG1, there are many possible options to organize those components. Fig. 5 shows two of the possible layout options, one with equal size views and one with a larger view accompanied with several small views. Let us assume the domain expert workflow is to get an overview first, then explore the details by demand. In this scenario, choosing the layout similar to the one on the right in Fig. 5 would be more suitable since it provides a larger space to display the overview visualization.

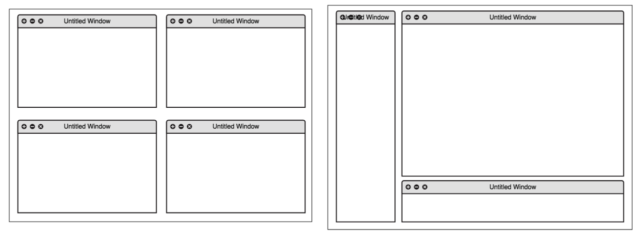


Figure 5: Two possible layout options of visual tools.

After filling the components into the layout, our next step is to deploy interaction techniques to guide the refining process and embed help information to support learning. Suppose we already have a list view to show all the car models, a table to list the attribute values for each car model, and a scatterplot view to display the relationship between two attributes picked by the user. According to DG3, there are two interaction techniques can be applied in this scenario to support progressive refinement. As shown in Fig. 6, one is the input box that allows user to search for a specific car model. The further interaction after the search can be that users select the data points they searched and the visualization change to a more detailed view to show the information of the data they clicked. Another interaction technique is the "progressively revealing" technique that helps users to understand how to transform a scatterplot to parallel coordinate after the user chooses two attributes as the vertical axis. According to DG4, there is some possible help information can be applied to current visualization such as a tooltip when a user hovers on a data point (Fig. 7), or the brushing technique that highlights the hovered data that is not shown in this simple mockup.

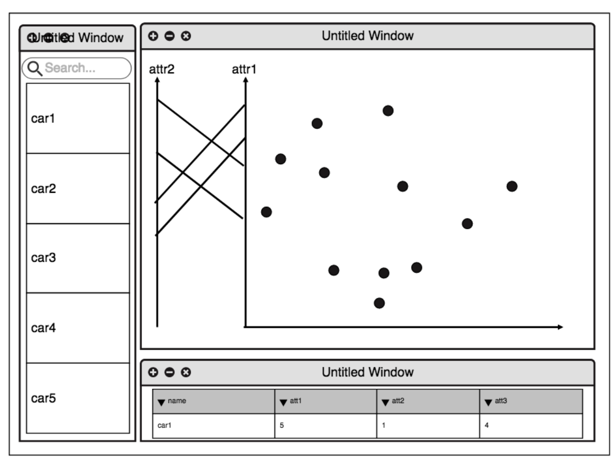


Figure 6: Progressively transform scatterplot to parallel coordinate by changing two dimensions of data each time.

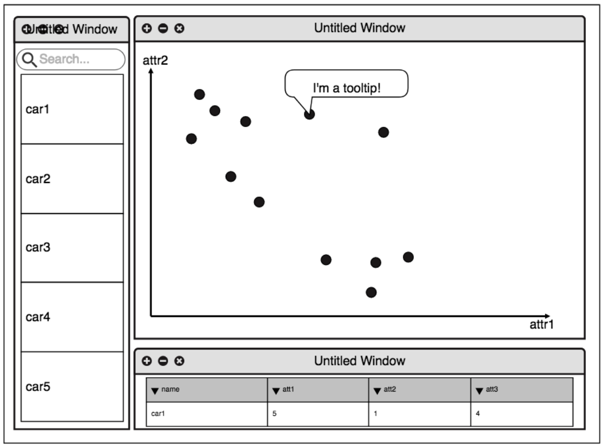


Figure 7: Tooltips show details on-demand of a selected data point as help information.

7 Evaluation for Domain Experts

Due to the identified characteristics and the specific design study methodology of domain experts, multiple evaluation methods need to be combined during the whole design process. Using the correct evaluation methods at the right stage of the design procedure can detect and avoid some of the pitfalls listed by Sedlmair et al. [47].

Lam et al. conducted empirical studies to identify seven evaluation scenarios discussing what questions those scenarios have and what are the possible evaluation methods to answer those questions [28]. According to these scenarios and the purposes during different design stages, we characterize evaluation methods used at each stage.

At the precondition stage, the goal is to understand domain experts' requirements, task flows, and terminologies they use during their work. The evaluation methods used at this stage are interviews (e.g., Autobahn [46], MizBee [33], and WeaVer [39]), field observation (e.g., DIA2 [30] and dotlink360 [3]), and ethnographic and observational studies (e.g., contextual inquiry applied in RelEx [44] and the work by Batch and Elmqvist [4]). This stage is time-consuming due to domain experts' time limitation, but play an important role for the whole process. If there is not enough time for the evaluation during this stage, the design is unlikely to succeed, causing either pitfall PF-5 (insufficient time available from potential collaborators) or PF-8 (no need for research; engineering vs. research project) [47].

At the core stage, the goal is to design the requested visualizations and test the usability of those tools developed. Due to the challenge of complex interface and visual representation, domain experts may not be able to work with early versions of tools or complex tools. To design their ideal visualizations, one possible solution is to use a Wizard of Oz approach where the research team invisibly translates the domain expert's commands into actual commands for the visualization system, like the work done by Salber and Coutaz [42] that applied the Wizard of Oz technique to study the multimodal systems. Another approach is to use pair analytics [1] where the visualization expert is the driver, and the domain expert is the passenger explaining to the driver what they want and let the driver perform the expert operations to achieve that. To test the usability of the tools developed, case studies are conducted to recruit several domain experts and ask them to finish tasks from their daily work to evaluate the visual design, efficiency, and efficacy of the tool.

Since the final goal of this design study is the adoption of developed tools by domain experts, the last stage would be to deploy the designed tool to the company environment and conduct the field observation evaluation. Although the case studies from previous stage have validated the usability of the tools, external deployment environment may not be exactly the same as the internal company environment due to the characteristic of environmental constraints, so field observation will be necessary to further validate the tools. To confirm the long term usage of developed tools, longitudinal acceptance testing and field testing in the organization may be required.

8 Conclusion

In this paper, we have presented the characteristics, existing examples and a subset of possible design guidelines for a special group of users: *domain experts*. We hope our work can help other researchers understand, design, and evaluate visualizations for this kind of users. In the future, we will explore how to accommodate these guidelines in visualization tools and toolkits.

Acknowledgments

This work was partially supported by U.S. National Science Foundation award DUE-1123108. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agency.

References

[1] R. Arias-Hernández, L. T. Kaastra, T. M. Green, and B. D. Fisher. Pair analytics: Capturing reasoning processes in collaborative visual analytics. In D. A. Keim and S. Wrobel, eds., *Scalable Visual Analytics*, number 10471 in Dagstuhl Seminar Proceedings. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany, Dagstuhl, Germany, 2011.

[2] M. Q. W. Baldonado, A. Woodruff, and A. Kuchinsky. Guidelines for using multiple views in information visualization. In *Proceedings of the ACM Conference on Advanced Visual Interfaces*, pp. 110–119, 2000. doi: 10.1145/345513.345271

[3] R. C. Basole, T. Clear, M. Hu, H. Mehrotra, and J. Stasko. Understanding interfirm relationships in business ecosystemswith interactive visualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2226–2235, 2013. doi: 10.1109/TVCG.2013.209

[4] A. Batch and N. Elmqvist. The interactive visualization gap ininitial exploratory data analysis. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):278–287, 2018. doi: 10.1109/TVCG.2017.2743990

[5] D. Baur, F. Seifert, M. Sedlmair, and S. Boring. The streams of our lives: Visualizing listening histories in context. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1119–1128, 2010. doi: 10.1109/TVCG.2010.206

[6] R. A. Becker and W. S. Cleveland. Brushing scatterplots. *Technometrics*, 29(2):127–142, 1987.

[7] R. A. Becker, W. S. Cleveland, and M.-J. Shyu. The visual design and control of Trellis display. *Journal of Computational and Graphical Statistics*, 5(2):123–155, 1996.

[8] J. Bernard, D. Daberkow, D. Fellner, K. Fischer, O. Koepler, J. Kohlhammer, M. Runnwerth, T. Ruppert, T. Schreck, and I. Sens. Visinfo: a digital library system for time series research databased on exploratory searcha user-centered design approach. *International Journal on Digital Libraries*, 1(16):37–59, 2015. doi: 10.1007/s00799-014-0134-y

[9] M. Bögl, W. Aigner, P. Filzmoser, T. Lammarsch, S. Miksch, and A. Rind. Visual analytics for model selection in time series analysis. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2237–2246, 2013. doi: 10.1109/TVCG.2013.222

[10] M. Booshehrian, T. Möller, R. M. Peterman, and T. Munzner. Vismon: Facilitating analysis of trade-offs, uncertainty, and sensitivity in fisheries management decision making. *Computer Graphics Forum*, 31(3):1235–1244, 2012. doi: 10.1111/j.1467-8659.2012.03116.x

[11] J. Boy, R. A. Rensink, E. Bertini, and J.-D. Fekete. A principled way of assessing visualization literacy. *IEEE Transactions on Visualization Computer Graphics*, 20(12):1963–1972, 2014. doi: 10.1109/TVCG.2014.2346984

[12] M. Brehmer, J. Ng, K. Tate, and T. Munzner. Matches, mismatches, and methods: Multiple-view workflows for energy portfolio analysis. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):449–458, 2015. doi: 10.1109/TVCG.2015.2466971

[13] H. Chen, D. Zeng, H. Atabakhsh, W. Wyzga, and J. Schroeder. CopLink: Managing law enforcement data andknowledge. *Communication of the ACM*, 46(1):28–34, 2003. doi: 10.1145/602421.602441

[14] D. Coelho, A. Kumar, and K. Mueller. Data memes for personal visualization. In *Proceedings of the Personal Visualization Workshop*, 2015.

[15] C. Csallner, M. Handte, O. Lehmann, and J. Stasko. FundExplorer: Supporting the diversification of mutual fund portfolios using context treemaps. In *Proceedings of the IEEE Symposium on Information Visualization*, pp. 203–208, 2003. doi: 10.1109/INFVIS.2003.1249027

[16] J. Fulda, M. Brehmel, and T. Munzner. TimeLineCurator: Interactive authoring of visual timelines from unstructured text. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):300–309, 2015. doi: 10.1109/TVCG.2015.2467531

[17] A. Gegenfurtner and M. Seppänen. Transfer of expertise: An eye tracking and think aloud study using dynamic medical visualizations. *Computers & Education*, 63:393–403, 2012. doi: 10.1016/j.compedu.2012.12.021

[18] J. Görtler, C. Schulz, D. Weiskopf, and O. Deussen. Bubble treemaps for uncertainty visualization. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):719–728, 2018. doi: 10.1109/TVCG.2017.2743959

[19] L. Grammel. *User Interfaces Supporting Information Visualization Novices in Visualization Construction*. PhD thesis, RWTH Aachen University, Department of Computer Science, Germany, 2012.

[20] L. Grammel, M. Tory, and M.-A. Storey. How information visualization novices construct visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):943–952, 2010. doi: 10.1109/TVCG.2010.164

[21] S. Havre, B. Hetzler, and L. Nowell. ThemeRiver: Visualizing theme changes over time. In *Proceedings of the IEEE Symposium on Information Visualization*, pp. 115–123, 2000. doi: 10.1109/INFVIS.2000.885098

[22] J. Heer and danah boyd. Vizster: visualizing online social networks. In *Proceedings of the IEEE Symposium on Information Visualization*, pp. 33–40, 2005. doi: 10.1109/INFVIS.2005.1532126

[23] D. Huang, M. Tory, B. A. Aseniero, L. Bartram, S. Bateman, S. Carpendale, A. Tang, and R. Woodbury. Personal visualization and personal visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 21(3):420–433, 2015. doi: 10.1109/TVCG.2014.2359887

[24] D. Hutchins, ed. *Just-in-Time*. Aldershot: Gower, 1988.

[25] S. Ko, J. Zhao, J. Xia, S. Afzal, X. Wang, G. Abram, N. Elmqvist, L. Kne, D. V. Riper, K. Gaither, S. Kennedy, W. Tolone, W. Ribarsky, and D. S. Ebert. VASA: Interactive computational steering of large asynchronous simulation pipelines for societal infrastructure. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1853–1862, 2014. doi: 10.1109/TVCG.2014.2346911

[26] B. C. Kwon, S.-H. Kim, S. Lee, J. H. Jaegul Choo, , and J. S. Yi. VisOHC: Designing visual analytics for online health communities. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):71–80, 2016. doi: 10.1109/TVCG.2015.2467555

[27] B. C. Kwon and B. Lee. A comparative evaluation on online learning approaches using parallel coordinate visualization. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 993–997, 2016. doi: 10.1145/2858036.2858101

[28] H. Lam, E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale. Empirical studies in information visualization: Seven scenarios. *IEEE Transactions on Visualization and Computer Graphics*, 18(9):1520–1536, 2011. doi: 10.1109/TVCG.2011.279

[29] Z. Liu and J. T. Stasko. Mental models,visual reasoning and interaction in information visualization: A top-down perspective. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):999–1008, 2010. doi: 10.1109/TVCG.2010.177

[30] K. Madhavan, N. Elmqvist, M. Vorvoreanu, Y. Wong, H. Xian, Z. Dong, and A. Johri. DIA2: Web-based cyberinfrastructure for visual analytics of funding portfolios. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1823–1832, 2014. doi: 10.1109/TVCG.2014.2346747

[31] E. Mayr, P. Federico, S. Miksch, G. Schreder, M. Smuc, and F. Windhager. Visualization of cultural heritage data for casual users. In *the 1st Workshop for Visualization for the Digital Humanities*, 2016.

[32] N. McCurdy, J. Lein, K. Coles, and M. Meyer. Poemage: Visualizing the sonic topology of a poem. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):439–448, 2016. doi: 10.1109/TVCG.2015.2467811

[33] M. Meyer, T. Munzner, and H. Pfister. MizBee: A multiscale synteny browser. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):897–904, 2009. doi: 10.1109/TVCG.2009.167

[34] T. Munzner, Q. Kong, R. T. Ng, J. Lee, J. Klawe, D. Radulovic, and C. K. Leung. Visual mining of power sets with large alphabets. *Technical report TR-2005-25, UBC, Canada*, 2005.

[35] M. Ogawa and K.-L. Ma. code_swarm: A design study in organic software visualization. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1097–1104, 2009. doi: 10.1109/TVCG.2009.123

[36] J. Poco, A. Dasgupta, Y. Wei, W. Hargrove, C. Schwalm, R. Cook, E. Bertini, and C. Silva. SimilarityExplorer: A visual inter-comparison toolfor multifaceted climate data. *Computer Graphics Forum*, 3(33):1923–1932, 2014. doi: 10.1111/cgf.12390

[37] Z. Pousman, J. T. Stasko, and M. Mateas. Casual information visualization: Depictions of data in everyday life. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1145–1152, 2007. doi: 10.1109/TVCG.2007.70541

[38] Z. Qu and J. Hullman. Keeping multiple views consistent: Constraints, validations, and exceptions in visualization authoring. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):1077–2626, 2018. doi: 10.1109/TVCG.2017.2744198

[39] P. S. Quinan and M. Meyer. Visually comparing weather features in forecasts. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):389–398, 2016. doi: 10.1109/TVCG.2015.2467754

[40] J. C. Roberts. State of the art: Coordinated & multiple views in exploratory visualization. In *Proceedings of the International Conference on Coordinated and Multiple Views*, 2007. doi: 10.1109/CMV.2007.20

[41] S. Rudolph, A. Savikhin, and D. S. Ebert. FinVis: Applied visual analytics for personal financial planning. pp. 195–202, 2009. doi: 10.1109/VAST.2009.5333920

[42] D. Salber and J. Coutaz. Applying the wizard of oz techniqueto the study of multimodal systems. In *International Conference on Human-Computer Interaction*, pp. 219–230, 1993.

[43] M. Sedlmair, C. Bernhold, D. Herrscher, S. Boring, and A. Butzr. MostVis: An interactive visualization supporting automotive engineers in MOST catalog exploration. In *Proceedings of the International Conference on Information Visualisation*, pp. 173–182, 2009. doi: 10.1109/IV.2009.95

[44] M. Sedlmair, A. Frank, T. Munzner, and A. Butz. Relex: Visualization for actively changing overlay network specifications. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2729–2738, 2012. doi: 10.1109/TVCG.2012.255

[45] M. Sedlmair, P. Isenberg, D. Baur, and A. Butz. Information visualization evaluation in large companies: Challenges, experiences and recommendations. *Information Visualization*, 10(3):248–266, 2011.

[46] M. Sedlmair, B. Kunze, W. Hintermaier, and A. Butz. User-centered development of a visual exploration system for in-car communication. In *Proceedings of the International Symposium on Smart Graphics*, pp. 105–116, 2009. doi: 10.1007/978-3-642-02115-2_9

[47] M. Sedlmair, M. Meyer, and T. Munzner. Design study methodology: Reflections from the trenches and the stacks. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2431–2440, 2012. doi: 10.1109/TVCG.2012.213

[48] E. Segel and J. Heer. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1139–2010, 2010. doi: 10.1109/TVCG.2010.179

[49] C. Shi, S. Fu, Q. Chen, and H. Qu. VisMOOC: Visualizing video clickstream datafrom massive open online courses. In *Proceedings of the IEEE Pacific Visualization Symposium 2015*, pp. 159–166, 2015. doi: 10.1109/PACIFICVIS.2015.7156373

[50] J. Stasko, T. Miller, Z. Pousman, C. Plaue, and O. Ullah. Personalized peripheral information awareness through information art. In *Proceedings of the ACM Conference on Ubiquitous Computing*, pp. 18–35, 2004. doi: 10.1007/978-3-540-30119-6_2

[51] H. Tang, J. J. Topczewski, A. M. Topczewski, and N. J. Pienta. Permutation test for groups of scanpaths using normalized levenshtein distances and application in nmr questions. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, pp. 169–172, 2012. doi: 10.1145/2168556.2168584

[52] F. B. Viégas, M. Wattenberg, F. van Ham, J. Kriss, and M. M. McKeon. ManyEyes: a site for visualization at internet scale. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1121–1128, 2007. doi: 10.1109/TVCG.2007.70577

[53] S. Wakeling, P. D. Clough, and J. Wyper. Graph literacy and business intelligence: Investigating user understanding of dashboard data visualizations. *Business Intelligence Journal*, 20(4):8–19, 2015.

[54] D. Walsh and M. M. Hall. Just looking around: Supporting casual users initial encounters with digital cultural heritage. In *Proceedings of the International Workshop on Supporting Complex Search Tasks*, 2015.

[55] M. Wattenberg and J. Kriss. Designing for social data analysis. *IEEE Transactions on Visualization and Computer Graphics*, 12(4):549–557, 2006. doi: 10.1109/TVCG.2006.65

[56] F. Windhager and M. Smuc. The arts of the possible information visualization in the field of politics. *eJournal of eDemocracy and Open Government*, 6(2):151–165, 2014.

[57] J. Wood, D. Badawood, J. Dykes, and A. Slingsby. Ballotmaps: Detecting name bias in alphabetically ordered ballot papers. *IEEE Transactions on Visualization and Computer Graphics*, 17(2):2384–2391, 2011. doi: 10.1109/TVCG.2011.174

[58] M. Yalçın, N. Elmqvist, and B. B. Bederson. Keshif: Rapid and expressive tabular data exploration for novices. *IEEE Transactions on Visualization and Computer Graphics*, PP(99):1077–2626, 2017. doi: 10.1109/TVCG.2017.2723393

[59] D. Yoon, N. H. Narayanan, S. Lee, and O.-C. Kwon. Exploring the effect of animation and progressive revealing on diagrammatic problem solving. In *Proceedings of the International Conference on Theory and Application of Diagrams*, pp. 226–240, 2006. doi: 10.1007/11783183_31