



King's Research Portal

Document Version
Peer reviewed version

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Windhager, F., Abdul-Rahman, A., Bludau, M.-J., Hengesbach, N., Lamqaddam, H., Meirelles, I., Speckmann, B., & Correll, M. (in press). Complexity as Design Material. In *BELIV: evaluation and BEyond - methodoLogical approaches for Visualization*

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.









- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Complexity as Design Material

Position Paper

Florian Windhager , Alfie Abdul-Rahman , Mark-Jan Bludau , Nicole Hengesbach , Houda Lamqaddam , Isabel Meirelles , Bettina Speckmann , and Michael Correll 

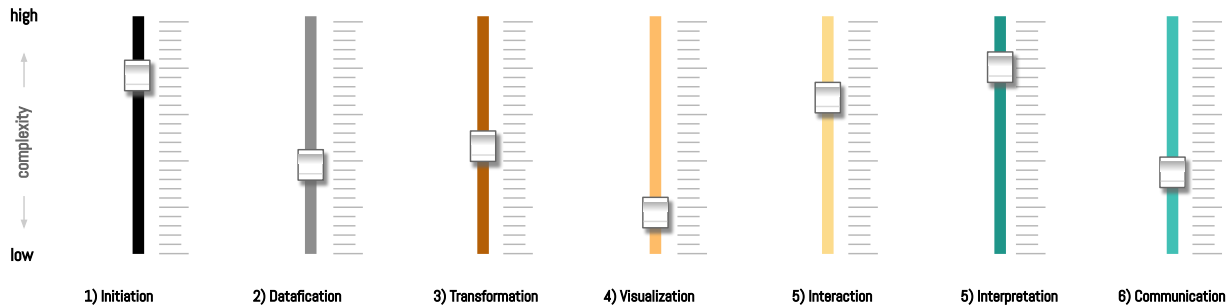


Fig. 1: Axes of complexity and complexity transformation in visualization design, bridging from project initiation complexity to the complexity of interpretation and communication activities, using the metaphor of a mixing board. A designer might strategically employ higher or lower levels of complexity across these axes to achieve a desired effect. Likewise, changes to one type of complexity *shift* complexity to other parts of the pipeline.

Abstract—Complexity is often seen as an inherent negative in information design, with the job of the designer being to reduce or eliminate complexity, and with principles like Tufte’s “data-ink ratio” or “chartjunk” to operationalize minimalism and simplicity in visualizations. However, in this position paper, we call for a more expansive view of complexity as a *design material*, like color or texture or shape: an element of information design that can be used in many ways, many of which are beneficial to the goals of using data to understand the world around us. We describe complexity as a phenomenon that occurs not just in visual design but in every aspect of the sensemaking process, from data collection to interpretation. For each of these stages, we present examples of ways that these various forms of complexity can be used (or abused) in visualization design. We ultimately call on the visualization community to build a more nuanced view of complexity, to look for places to usefully integrate complexity in multiple stages of the design process, and, even when the goal is to reduce complexity, to look for the non-visual forms of complexity that may have otherwise been overlooked.

Index Terms—Complexity, Design, Visualization

1 INTRODUCTION

Much ink has been spilled and experimental data collected in service of what is perceived as a key conflict in data visualization research and practice: the role of complexity. On one side, a “minimalist” [69] view of data visualization design towards reducing complexity as much as possible, influenced by notions from Tufte such as “chartjunk” and the

“data-ink ratio” [132]. On the other side, a view touting the benefits of adornment and ornamentation on grounds of memorability [16], pedagogy, or a host of other desiderata [14]. In short, to what extent is the goal of data visualization to reduce complexity?

We posit that certain kinds of complexity, especially when thinking of complexity *beyond the visual*, is rarely, if ever, actually reduced, but merely *moved* to other parts of the process of visualization creation or interpretation. We draw inspiration here from maxims like Tesler’s Law that states that “every application has an inherent amount of irreducible complexity. The only question is who will have to deal with it, the user or the developer” [117]. In line with the adage “as simple as possible, but not simpler” we consider an adapted *law of requisite complexity* [15] as a noteworthy addition to reductionist visualization design guidelines: in the face of complex things, useful representations have to actually *retain* a sufficient degree of complexity, to inform related controlling, governing and problem solving efforts. Lastly, we hold that complexity is not an intrinsic deficiency, or even an overall negative component with occasional counter-intuitive “beneficial difficulties” [67], but that complexity is (beyond being an inescapable and often necessary facet of human experience [100]) frequently a *useful* tool for accomplishing important design goals.

We therefore propose that neither the thesis of data visualization minimalism nor the antithesis of data visualization adornment capture the phenomenon of interest or the language in which we should be articulating design principles. We instead propose a synthesis: *complexity as a design material that can be strategically employed by designers at all stages of the design pipeline*. Just as a designer can make judicious use

- Florian Windhager is with University for Continuing Education Krems. E-mail: florian.windhager@donau-uni.ac.at
- Alfie Abdul-Rahman is with King’s College London. E-mail: alfie.abdulrahman@kcl.ac.uk
- Mark-Jan Bludau is with University of Applied Sciences Potsdam. E-mail: mark-jan.bludau@fh-potsdam.de
- Nicole Hengesbach is with Warwick University. E-mail: n.hengesbach@warwick.ac.uk
- Houda Lamqaddam is with University of Amsterdam. E-mail: h.lamqaddam@uva.nl
- Isabell Meirelles is with OCAD University Toronto. E-mail: imeirelles@ocadu.ca
- Bettina Speckmann is with Eindhoven University of Technology. E-mail: b.speckmann@tue.nl
- Michael Correll is with Northeastern University. E-mail: m.correll@northeastern.edu

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxx/TVCG.201x.xxxxxx

(or non-use) of color, interactivity, or animation to accomplish design goals in data visualization, so too can a designer thoughtfully employ (or even “dance with” [49]) complexity. To continue the analogy with color, when, for example, Stone proposes that chart designers seeking to employ color first “get it right in black and white” [126], this is not to divide the community into pro-color and anti-color camps, but to call for intentionality and strategy in the use of color as a design material. Similarly, we hold that designers should be mindful and strategic in how they employ complexity, rather than a knee-jerk assumption that complexity can (or should) be unilaterally minimized in a funnel-like process of reduction from data to chart to insight.

To guide our discussion of complexity’s role in the design process, and to focus on forms of complexity other than merely the visual, we propose a set of *axes of complexity* (Figure 1) that represent complexities at all stages of the design process, and show how complexity changes form, purpose, and role as it moves along these axes. We focus especially on aspects of complexity that have been given short shrift by the over-emphasis on the pros and cons of strictly *visual* notions of visualization complexity. We suggest that there are many existing productive or fruitful designs of visualizations that employ strategies other than reducing complexity as much as possible, as quickly as possible, as prescribed by many of the standard visualization pipelines.

We also suggest that often what is perceived as a *reduction* in complexity is in fact a *shifting* of this complexity to other points along the design and sensemaking process. For instance, while a technique like principal component analysis could be used to reduce a high-dimensional dataset down to a two-dimensional scatterplot—a seeming reduction in data complexity—the resulting axes of this plot are now esoteric and difficult to interpret [51]: complexity has not been removed, but shifted to a complexity in *interpretation*.

Where and how complexity is shifted during visualization design should be the result of a careful balancing act and negotiation between designers, researchers, audiences, and many other stakeholders, and not the result of standardized complexity reduction and control. We hope that considering the movement and translation of complexity along our proposed axes might reveal places where novel design interventions or design goals could be realized, and places where we lack current guidelines or empirical guidance from the visualization literature.

The rest of the paper is organized as follows. We first employ on thinking about complexity both within and without the field of visualization to lay out existing notions of complexity (§2). We aim to integrate, enrich, and contextualize these existing notions with a process-oriented model of complexity along seven axes (§3). We conclude with a call to action for the visualization community to adopt a more nuanced and multivalent view of complexity in visualization design, and in particular to draw more attention to the forms of complexity beyond the visual that are often not well-discussed in existing visualization empiricism or pedagogy (§4).

2 RELATED WORK

In this section, we briefly overview complexity as a general concept and then provide an overview of existing notions of complexity in visualization.

2.1 Complexity in General

“Complexity” as a concept is used to characterize a wide range of topics and phenomena in a large number of knowledge domains, including physics [13], biology [89], psychology [136], sociology [26], economics [7], management [103], technology [70], and in cross-domain discourses such as systems and complexity science [45, 107]. In common parlance, the adjective “complex” is closely tied to—and often mostly synonymous with—“complicated” or “intricate” (although see Norman [100] for a theoretical distinction between “complex” as a property of user interpretation and “complicated” as a property of the exterior world). In contrast, many scholars have argued to reserve the use of “complex” to phenomena with a very large numbers of **differentiated elements and interactions**, frequently including increased levels of **descriptive “unknowns” and “unknowables”**, which leads

to a substantial **decrease of external comprehensibility and controllability** [52, 74, 108], and the need for phenomena to **“self-organize” and build up order internally**, from structural hierarchies to adaptive behavioral patterns.

For the rest of this paper, we use the term “complexity” with this whole spectrum of meaning in mind—from a general concept to system-specific definitions, rather than focus on one particular definition divorced from a context of use. Defining complexity in a rigorous fashion *aside from certain fields of study* has been called an “intractable problem” [128, p.52]. This has sparked controversies about the scientific value of the concept in general [129], but its ever-increasing use across all fields of societal discourse rather proves the value of the term as a transdisciplinary “problem marker”.

In approximation to information-theoretical definitions, this problem marker often refers to things or processes, which evade simple descriptions and attempts of control—relative to an established level of descriptive and cybernetic (i.e. management or control-oriented) resources. More specifically, “complex” is used to flag subject matters, which consist of a *large number of differentiated elements*, which interact in varied and often unpredictable ways, due to *numerous and diverse intra-system and system-environment relations and interactions*. In addition, a diachronic perspective (i.e. observation over time) often describes *non-trivial and unpredictable system behaviors*. Systems-theoretical perspectives often reflect on the “emergent” qualities of complex systems, which are said to create novel, “super-summative” characteristics and behaviors, which are non-reducible to an atomistic understanding of their parts, as complexity often leads to self-organizing and adaptive behavior [60, 86, 120].

According to information-theoretical definitions, complexity largely corresponds to the length of system descriptions needed to successfully model (i.e., describe, calculate, explain, compute, comprehend, communicate) and govern a system and its behavior [55]. Thinking and learning about complex things (i.e. building up viable mental models) and training how to handle them takes time—a burden which can be shared by dividing knowledge and labor. Many systems scientists diagnose a general trend in the evolution of societies and other systems towards an increase of internal complexity [72, 91]. Given complex and adversarial environments, systems with more knowledge and response options gain a competitive evolutionary edge. The “*law of requisite variety*” [75] states that “The larger the variety of actions available to a control system, the larger the variety of perturbations it is able to compensate” or (only) “complexity absorbs complexity”. By raising their socio-technical complexity (e.g., by dividing roles, knowledge and labor between individuals and organizations as task-specific problem-solvers) cultures and systems in general are able to address increasing amounts of problems [113, 114, 133]. However, this collective strategy also increases the complexity challenges for individual observers, whose natural perceptual and cognitive resources remain limited.

2.2 Complexity in Visualization

The outlined complexity of natural, cultural, technological, and semi-otic environments provides the standard argument for the *cognition-enhancing technologies* of modern times, including the tools of visualization [59, 85]. Given the raw number of complex subject matters—and their amassed descriptions in libraries, archives, and interlinking data collections, “second order technologies” become key. These technologies augment and amplify human cognition in face of complex (i.e. “massive, dynamic, ambiguous, and often conflicting” [130, p.10]) data.

Visualization thus aims to translate data complexity into forms that reduce the cognitive load of sensemaking [105], mediating between topic and data complexity on the one hand and limited cognitive bandwidth on the other [6]. At first glance, this objective favors aesthetics of simplicity, efficiency and reduction to the essence. Just as a map cannot and should not mirror the territory, neither can or should a diagram represent the full complexity of the underlying topic. Thus, a *complexity-reductionist agenda* undergirds many visualization design guidelines: visualizations should move us from data to insights, help to separate the signal from the noise, and put related evidence into a proper visual perspective. Above all else, visualizations should show



Fig. 2: Recurring definition criteria of “complex” subject matters (left) and application of this definition to different aspects of visualization practice.

the data, steer clear from all other distractions, and generally generate more (insights) with less (ink/pixels) [132].

According to these tenets, visualization pipelines serve as funnels for data complexity which minimize the “data-to-cognitive-work ratio”: The time and effort needed to process a given dataset alphanumerically is significantly reduced by visual encodings (if in line with rules of “graphical excellence”)—so that cognitive activities (including visual perception, mental modelling, reasoning and communication) can happen for a fraction of the original processing costs. If individual views cannot exhaust a dataset, facets of data complexity will not be visually superimposed, but distributed over *multiple* minimalist views. Tellingly, this should happen in line with the *rule of parsimony* that recommends to keep the number of views to a minimum [8, p.115].

While these minimalist design aesthetics have a pervasive and longitudinal impact on the visualization field, there are empirical and theoretical objections to the orthodox application of minimalism in visualization. Concepts such as “data-ink-ratio” or “chart junk” have attracted the scrutiny of visualization researchers, who have collected both empirical evidence against an overly-narrow application of minimalist strategies in visualization design as well as counter-examples or counter-framings [1, 2, 11, 12, 27, 61, 62, 67, 69, 83, 106].

Arguably, this debate around minimalism (see also §3.4) has largely dominated the discussion of complexity in the visualization field. However, the state of the debate is still quite confrontational and centers on visual complexity, while the multi-faceted concept of complexity we propose offers the chance to synthesize discussions about consequential translations and trade-offs across multiple dimensions and steps of the design process, and to do so in a more balanced and systemic fashion. Thought of as design material, complexity can be understood as a resource, rather than a problem marker only, that can be used to achieve different ends. Figure 2 (left) draws together definitional facets of complexity which we find in the literature (i.e., many elements or options, many relations or interactions, diversity of elements and relations, relevance of time) and illustrates how they can help to characterize the distributed complexity scenarios in visualization (right). As a design material, complexity can be used in many places: sparingly used, say, in visual design, but lavished on interaction design or in data preparation. In the following sections, we collect and connect related observations and arguments on either side of the focus area of *visual* complexity, and thus connect to other major decision scenarios for the design and reading of visualizations—what we term “axes” of complexity along the visualization design pipeline.

3 AXES OF COMPLEXITY

In this section, we focus on areas where complexity can arise in the process of visualization, with a particular focus on facets other than the complexity of a visualization *per se*. We structure these complexities along a set of axes capturing key decision points in the visualization design process, from selecting topics and hypotheses of interest to the communication of any derived insights to stakeholders. (The font color of the subsections correspond to the axes introduced in Figure 1.)

The concept of “axes” serves as a heuristic device to mark decision points of complexity regulation or transformation along the sensemaking pipeline (e.g., [54, 92]). While we discuss our axes in a separated fashion (acting like sliders of a mixer console, see Figure 1), in practice, these processes have blurred areas of responsibility, overlap each other, or are tightly coupled (e.g. interaction may lead to direct data transformations and encoding changes). Rather than present a complete and compartmentalized view of sensemaking using visualization, it is the general aim of the following sections to highlight sources of complexity that might otherwise be missed in the familiar discussion of *visualization complexity* (focusing on the visual) and to bring connected consequences of any complexity decision into a sharper focus.

3.1 Initiation

Initiating visualization projects entails numerous decisions about the overall complexity of the ensuing setup, including deliberations on all the following phases and axes. At the outset, there is a *topic* of interest (or one is selected), for which knowledge has to be built up, or around which collaborators with extensive prior domain knowledge exist [122]. This meeting point between domain knowledge and visualization process is also a site of friction and heightened complexity. At this frontier, complexity stems from two complementary aspects:

(1a) **Topic complexity:** No matter the discipline, the original phenomena, topics, or study materials can be of varying levels of complexity depending on the nature of the material as well as the goals addressed. As an example, we look at the phenomenon of analyzing and communicating climate change data to policy makers or the wider public [94, 121, 134], which might involve forecasting of many interconnected factors like temperature, sea level rise, precipitation, etc. The studied phenomenon in this case rests on a multivariate interaction of several parameters over time, supported by complex mathematical and statistical analysis. Narrowing the sorts of questions involved (say,

from climate change as a gestalt phenomena to merely one aspect, like sea level rise) may still rest, at heart, on complex and interconnected *phenomena*, but reduces the scope of the topic at hand and has a large impact on the downstream complexity of any visualization project.

(1b) **Descriptive complexity:** Topic complexity arguably translates into complexity of topic descriptions and representations—whether of texts, equations, models, or algorithms, which can be approximated with measures such as linguistic or algorithmic complexity, and secondarily with regard to depth, variety and history of related knowledge, discourse, and controversies (e.g., for climate change, see [53, 109]). As the complexity of topics arguably cannot be measured directly, we treat descriptive complexity as a proxy for topic complexity.

(2) **Epistemic differences** between the domain discipline and visualization research: The distance in epistemic frameworks between the application domain and visualization is a compounding factor for topic complexity. In computational disciplines, the complexity of the discipline is largely due to what we describe above as inherent to the data or tasks. It can then be relatively smoothly translated into data, and transformed into a design and interaction design task. Rather than a straightforward translation of phenomenon or topic into data, many disciplinary collaborations require a broader gap to be bridged in order to align the studied objects with visualization-specific paradigms. In disciplines within the humanities, for example, scholarship is interpretative, qualitative, often focused on a small dataset or singular object of study. The inherent complexity of the task and theoretical framework is then supplemented by the additional effort to bridge the epistemic and methodological differences [20, 63].

Uses and abuses: These sources of complexity suggest that the interaction between domain phenomena and visualization inherently makes complexity a core element of the design process from the very start of the visualization process. The amount of attention that is initially given to a topic’s epistemic, descriptive and discursive complexity is a critical factor, deciding the success of a project (e.g., whether it will be considered to be a legitimate contribution, or rather a “trojan horse”, bringing undue computational assumptions and simplifications to the field [37]). Visualization design allows for the reduction of topic complexity for instance by abstracting or re-scoping tasks, focusing on a case study, or otherwise performing “data counseling” [43] during the process of operationalization. However, the insights and interpretations that can be taken from data and its visualization in relation to the real-world phenomenon depend on this translation. If the phenomenon becomes oversimplified through datafication or visualization, insights will remain limited based on the (semantic or causal) gap between what is represented and how it is represented [78, 102].

3.2 Datafication

Datafication is the process of transforming aspects of a real-world phenomenon or topic into data that can be processed computationally for different purposes, and is an active process of shaping rather than a mere passive collection (indeed, Drucker [38] uses the term “capta” over the usual “data” to precisely denote this active and teleological process). Datafication includes activities of a) native data creation and collection from observations or digital sensors, as well as b) digitization practices which convert analog content or representations of a subject matter—such as books, films, photographs—into digital information. In our context, datafication also includes c) the selection and compilation of given datasets about a topic, to either work with them directly or to prepare them for further transformation. While many of these processes are fairly standardized, they concatenate, aggregate and introduce a substantial number of (ontological, epistemological, methodological, practical) design choices and thus are acutely non-trivial procedures, which can be done more or less comprehensively [44, 123].

Datasets for any object of study thus can be more or less complex: Simply put, low complexity datasets appear as spreadsheets with small numbers of rows, columns, facets, or data types, while high complexity data sources aggregate multiple data sources into large numbers

of rows, columns, facets, data types, and relations—with fundamental implications for all (computational, analytical, or interpretational) downstream activities. Descriptions of datasets, such as datasheets [50], can therefore contain a myriad of factors, from sourcing to metadata to an assessment of potential biases and limitations of the data that have been collected.

When thinking about complexity as “design material”, datafication is a major decision point, even though it’s often overlooked for two reasons: (1) As visualization researchers often work with pre-existing data sets and sources, they tend to refer back to domain experts for decisions on both topic and datafication complexity, in order to jump straight to transformation activities, and (2) Visualization experts are used to starting from high data complexity as a *fait accompli*: data is mostly complex, that’s why the visualization design expertise is needed to distill it into simpler, more accessible and understandable formats. That is, there is a selection bias in visualization research, where we often focus on complex (but exotic) “zebra” problems while giving shorter shrift to more quotidian “horse” problems [28] that can be solved by existing methods and techniques.

Uses and abuses: Given this dominant notion of pre-existing data complexity, related design questions and complications have been discussed on various levels:

(1) Datasets can be complex, but full of errors, problems, and issues which leads to the next phase of *transformation* activities.

(2) Datasets can be correct on their faces, but represent just one perspective or descriptive option amongst many—especially for topics with a long and complex, controversial discursive history.

(3) Datasets can be constructed to deliberately hide, distort, mislead with respect to a topic. Even without intentional malicious decisions to bias, there is no objectively created dataset that represents a “view from nowhere” [57]: the situatedness that we bring to knowledge also is reflected in how we construct datasets.

(4) Even corrected and enriched by further transformations, datasets can be complex, but not complex enough to e.g. enable or support meaningful analysis and interpretation of a subject matter. This is a challenge extensively discussed in the humanities, whose language-based inquiries allow for all kinds of quantitative expressions, while largely focusing on qualitative assessments, interpretations and arguments. Datafication and computational methods with their focus on quantification thus significantly reduce and limit the existing means of “mixed methods” modeling, sensemaking, and reasoning [39].¹ Without due attention to these datafication-related questions, no downstream means or efforts (including visualization interpretation and communication) will be able to correct related design choices and challenges. However, critical transparency and design strategies [34] together with mixed-methods project designs can help to counterbalance detrimental datafication consequences.

3.3 Transformation

After a topic and data source has been selected, there is almost always still remaining work before visualization can begin. The process of acquiring, preparing, cleaning, federating, verifying, and otherwise building the data set to be visualized contains many complex steps involving diverse sets of expertise [3, 10, 71, 95] that, in repeated surveys of work practices of data scientists, dwarf the time and effort spent on actual visualization and analysis [4]. Yet, despite the importance and potential depth of these data transformation and preparation steps, relatively few tools exist for visualizing data provenance [18, 24, 32], or the impact of different potential analytical “paths” on the subsequent conclusions generated from the data [36, 111, 119].

The creation of models and other algorithmic ways of enriching data can also be thought of as a form of transformation, with all of the requisite complexities that come from modeling. We note that the complexity of the models *per se* is of interest for this particular axis, rather than the outputs of the models. For instance, dimensionality

¹See also “dataism” as modern-day ideology “rooted in a belief in the capacity of data to represent social life, sometimes better or more objectively than pre-digital (human) interpretations” [64].

reduction methods would, from certain perspectives, appear to be a reduction in data complexity (going from many dimensions down to only a few). However, the myriad techniques for dimensionality reduction have varying levels of sophistication, fidelity of the collected data from the datafication step, and many resulting practical interpretations [41]. We argue that eventually the price for this reduction is paid later on in the sensemaking process (for instance, in complexity of interpretation). Both modeling and more traditional data preparation steps can interact in unexpected ways that impact downstream complexity as well. For instance, Crisan & Correll [31] note that (often invisible or comparatively minor) decisions around data cleaning and text processing often had more significant impacts on resulting visualizations of topic models than the hyper-parameters of the actual topic modeling algorithms used.

Uses and abuses: There are many places where surfacing the provenance of a visualization, or the analytical steps taken to arise at a final data set, can be beneficial, despite introducing considerable complexity in the design process. The transparency, reliability, and credibility of data-driven insights is dependent on truthfully communicating not just results but also methods. Prominent scandals in the “replication crisis” and, in some cases, explicit scientific fraud, have been driven by the fact that there are many possible ways to subtly or even blatantly manipulate data (and so conclusions) in ways that can be difficult to detect even with direct inspection [79]. Since data cleaning and transformation also involves many subjective and potentially tendentious choices, it can also be used to discount the credibility of otherwise authoritative datasets [81].

3.4 Visualization

Visualization complexity is arguably the core around which the complexity debate in visualization research has crystallized: If data complexity creates issues for sensemaking, visualizations provide a solution by distilling it into simpler, more accessible and understandable formats. But what is the right amount or level of visual simplicity?

Given a limited amount of effective visual variables and screen space on the one side, and a limited perceptual and interpretational bandwidth on the other side [105], “minimalist” guidelines aim for maximum reduction of structural and formal chart complexity, influenced by notions from Tufte such as “data-ink ratio” or “chartjunk” [69, 132]. The driving notion has been likened to the modernist pursuit of the “anti-sublime”: “If Romantic artists thought of certain phenomena and effects as un-representable, as something which goes beyond the limits of human senses and reason, data visualization artists aim at precisely the opposite: to map such phenomena into a representation whose scale is comparable to the scales of human perception and cognition.” [87]. If one visualization is not enough, representational complexity should be distributed across the smallest number of multiple views [8].

Similar to meta- or postmodernist reflections in art history, a growing body of work built on these minimalist principles but challenged the all too rigid and orthodox applications, emphasizing the many ‘paradoxical’ or plainly beneficial exceptions, where complexity-increasing elements such as adornment and ornamentation strengthen intended effects such as memorability [16], pedagogical efficiency, aesthetics, playfulness, joy of use, and a whole range of other desiderata [1, 2, 14, 61, 62, 67, 83, 106].

Like other forms of representation (e.g., oral vs. written language), the visual representation and remediation of data necessarily involves abstraction and reduction processes, be that of dimensions in the data or the level of properties, such as when we aggregate or summarize. The outlined discussion thus revolves around the omnipresent danger of eliminating useful complexity, such as the isolation of entities from context, the removal of relevant interdependencies, and the reduction of nuance. During each step of reducing or eliminating data complexity, assumptions are made to achieve a productive level of abstraction. These assumptions must be presented to the target audiences to ensure information is preserved and accurate, and semiotic complexity for the interpretation of a chart is maintained e.g. by annotations, legends, and captions. As for appropriate design choices, it is relevant to consider comprehensively whether visually preserving the phenomenon’s complexity is meaningful or beneficial, despite raising difficulties [67].

Uses and abuses: What does visual complexity do to visualization readers and users? While specific reactions depend on context of users and use, the general assumption is that it *raises costs associated with visualization interpretation and communication*—which can be detrimental or beneficial to certain ends. *Arguments against the use of visual complexity* (i.e. assumptions driving the minimalist mantra) comprise cognitive and emotional (side) effects such as the tendency of complex charts to bewilder, scare, repel, mislead, slow consumption or shut people down, to confuse and overwhelm them, or to hide relevant details and discourage exploration. In short, the guiding minimalist notion is that chart complexity provides no solution, as that’s what complex data already does to most of its observers.

Arguments for the use of visual complexity (i.e. assumptions driving objections to minimalism and to maintain or increase visual complexity) turn the table on many of the pertinent worries and focus on the potential of chart complexity to impress, attract, and to appeal aesthetically, to rise emotional impact, add nuance, allow for polyvocality, promote idiosyncratic interpretation and deeper reflection, to defamiliarize the object of study, promote memorability, convey authority, slow consumption down (as in “slow analytics” [21]), and also to encourage interaction and exploration. Complexity can also be used rhetorically to convince people that a problem really is complicated, because there is no justifiable way to visualize it by simple means [76].

3.5 Interaction

Interactivity is known to aid the comprehension of data across multiple levels of details and perspectives through multiple operations on data or their representation [96, 138]. Thus, it is the main technique to modulate and control upstream and downstream axes of complexity according to the users’ developing intent. By enabling the display of data facets across multiple views, interaction can help to reduce or increase complexity dynamically with regards to the visualization and the displayed data, by reducing clutter, or by focusing only on relevant aspects. Most notably, interaction is used to regulate the complexity described with the axes of transformation and visualization. However, despite its purported purpose and potential to effectively solve many complexity challenges, interaction introduces a complexity layer of its own on top of every representational system or device.

Part of this added complexity is commonly referred to as “interaction costs” [77]. Such interaction costs can result from the required additional cognitive work in the decision- and sensemaking cycle. For instance, an abundance of options to interact, together with multi-level interaction patterns demand more visual and mental effort for analytical decision-making in contrast to limited interaction options and single-level interfaces. Here, the complexity of the interaction is a trade-off between the exploratory freedom of a user and a lower cognitive load [77]. Similarly, motor effort (e.g., hitting small buttons with a pointer) and physical effort (e.g., distance of pointer movements and quantity of clicks) tend to increase the operational complexity [77]. With this, the complexity of interaction work varies between low efforts for simple interactions (e.g. scrolling) and the steeply increased effort for using advanced and combined interaction techniques (e.g. multi-touch gestures or combination of multiple consecutive interaction patterns). Another cost is, in environments where many visualizations are presumed to be static, signaling the interactive affordances of a given visualization [19], or providing alternate methods of communication for those who do not interact with a chart in sufficient detail.

A further layer introduced through interaction is the temporal dimension. While static data representations oftentimes suffer from spatial limitations of page or screen sizes, interaction is able to separate the representation into multiple views with the trade-off that it introduces time as another dimension [138]. Temporal separation also is relevant with regard to directness of an interaction. Keeping temporal, spatial and conceptual separation low can improve directness of interactions [131] and by this decrease cognitive complexity.

Also the familiarity and experience of users with an interaction techniques or patterns has influence on the perceived complexity. Training or experience may reduce the demanded mental effort and therefore the perceived complexity. In contrast, when using novel interaction

patterns, unexpected behaviour (e.g., non-standard zoom behaviour in maps) of an interface can lead to confusion or frustration for unfamiliar users [77].

Uses and abuses: In the end, interaction complexity is a balancing act between considerations of user freedom, expertise, familiarity, and the goals of an interactive visualization. For instance, scrollytelling pieces are capable of narrating complex topics with complex visual representations: A narrative introduction to machine learning [137] can be considered as complex regarding encoding and content, while the interaction itself is minimized to scrolling-based navigation of the information in an author-driven way.

In contrast, a t-SNE based world map, can arrange countries in form of circles based on similarity metrics [115]. The encoding of the visualization itself is rather simple, however, the interface offers many options and possibilities to adjust the representation. It allows selection of the countries for detailed information, their re-coloring and re-sizing, in joint with adjusting the weights of the similarity calculations with fine-tuned sliders for over 30 data dimensions. Users are not only provided with a lot of freedom for exploration, but also with an abundance of possibilities to choose. Despite rather simple interface elements for interactions (drop-down menus, sliders, check boxes), the multitude of options and possibilities increases the complexity of the overall visualization. Additionally, the abstractness of t-SNE calculations adds to the complexity, by making results of interactions and new layouts rather unpredictable.

3.6 Interpretation

A common framing of work from graphical perception that influences visualization design is that interpreting visualization is primarily a matter of *decoding*: that is, taking a data value that has been *encoded* via a visual variable (like height in bar charts) and converting it back into a value through visual estimation. Bertini et al. [14] challenge this conception. For one, the decoding of individual values might be just one of the steps involved in extracting the desired information from a visualization—other tasks might involve statistical judgements around aggregate values, or cross multiple levels of description and summarization [22] that involve not just decoding values but performing additional sensemaking. For another, mismatches in visual metaphors, genre, or visual literacies in the intended audience can produce errors or misinterpretations in visualizations even if lower-level perceptual skills are correctly applied [99].

We refer to interpretative complexity as the relative ease or difficulty with which a viewer successfully reads and interprets a visualization. This difficulty is based on several sub-components:

(1) *Visual literacy and chart familiarity.* Unfamiliar designs or genres of charts can be difficult to interpret, and can require active effort to understand [82], and occasionally even explicit designs to “onboard” [33] new audiences. Likewise, dis-congruent or unfamiliar visual metaphors can also impact task performance and chart fluency [139].

(2) *Task and analytical intent.* While “decoding” a single value (say, determining the height of a bar in a bar chart and then reading off the corresponding value on the y-axis) is straightforward, analytical goals and insights in charts are rarely at the level of reading off a single value, but can involve *aggregate* [127] tasks, or even statistical judgements and estimations (like visually estimating trends [30]). The notion of interpretational complexity is particularly apposite when considering *insights*. Per Shneiderman [66], “the purpose of visualization is insight, not pictures”: and the nature of these insights, even those derived from relatively simple displays or data, are inherently “complex” and “deep” [101].

(3) *Grounding and explanation.* Even if the proper value(s) have been successfully read from the chart, there is still the resulting step of translating those values within the context of use (i.e., what decoded visual features mean in actual interpretative terms), which can vary in complexity. To revisit an example from our introduction, the *meaning* of a difference between two points in a scatterplot is dependent on their axes. If the axes are “simple” (say, sales over time), then the interpretation of a point might also be simple (say, a point might be two months in the future, and have twice as many sales, as another). However, axes

generated by processes like multidimensional scaling or projection can be difficult to explain [42, 51] or conceptualize, involving combinations of multiple underlying data dimensions in occasionally non-linear geometric space. The common visual form of the scatterplot is identical between these two cases, but the complexities of interpretation are not. We also note that forms of hermeneutical inquiry like close reading [9] can involve considerable interpretative effort even if the underlying data are relatively “simple”: the process of *reading into* a work can introduce interpretative complexity.

Uses and abuses: As per Bertini et al. [14], designs that may be more “complex” from the perspective of decoding individual values may still have benefits either for supporting the intended tasks of the visualization, better aligning with existing or expected visual metaphors. There are other purported benefits to what would otherwise be seen as “complex” interpretative designs: “beneficial difficulties” [67] that can improve engagement and promote a “slow analytics” [21] or “slow reading” [46] with deeper and more thoughtful engagement with the material. We note, however, that techniques that superficially appear to simplify data (like projection, clustering, modeling, and sampling) can incur follow-up costs in terms of interpretative complexity. Machine learning might be a quintessential example: a predictive algorithm might condense a high-dimensional space down to a binary “yes/no” decision. This decision is trivial to visualize. However, good designs that afford a simple *interpretation* of this decision occupies the expansive field of explainable AI (XAI) research.

3.7 Communication

Another stage of complexity management comes into play if visualization interpretation should be optimized and actively supported for more than one user group. For one, visualizations do not reveal or disclose themselves, but depend on different types of mediation, onboarding, contextualization, presentation, argumentation, persuasion, or explanation strategies [125]. For another, depending on the aims and context of a visualization project, the addressed audience can be complex (heterogeneous, diverse, differentiated) in itself. Visualization designers thus are facing distinct user groups with different levels of prior knowledge, motivation, preferences, intentions, and tasks. As a consequence, the conditions of success multiply and require multiple efforts of communication design.

For visualization projects addressing more than one user group, designers have to consider developing either multiple user-specific i) onboarding and communication programs [33] (e.g. with varying depth for lay persons or experts) or ii) multi-user interface designs (i.e. with user-specific combinations of visualization and interaction design). Related design strategies have been discussed, amongst others, with focus on adaptive, personalized, or customizable interfaces [104], progressive disclosure [124], multi-channel approaches to data visualization [135], design for expert vs. casual users [110] and persona-driven [118], or accessible, inclusive, and universal design [88].

Connecting back to axis 1 (topic complexity, defined during a project’s *initiation* phase), the communication scenarios for some topics are known to be specifically complex, as they are widely present in contemporary culture, politics, and media, or even argued to be of planetary concern, but with different implications for different population subgroups (e.g., climate change). Related knowledge, discourse, and decision scenarios are commonly complex, and full of polarizing epistemological, political, ethical, and interpretative controversies. Under such circumstances, visualization-based communication arguably becomes hyper-complex, requiring designers not only to think about all available measures of inclusive, appealing, and engaging design, but also to coordinate their local project initiatives with the complex communication and engagement strategies of large (inter- and non-) governmental activity programs [121, 134].

Uses and abuses: The inherent strengths of visualizations to reduce and accessibly convey topic and data complexity for a wide range of audiences provides a main rationale for the whole field [98]. As such, its applied, communication-supporting branches weave through all fields of society, including science [47], business [73], journalism [48], and

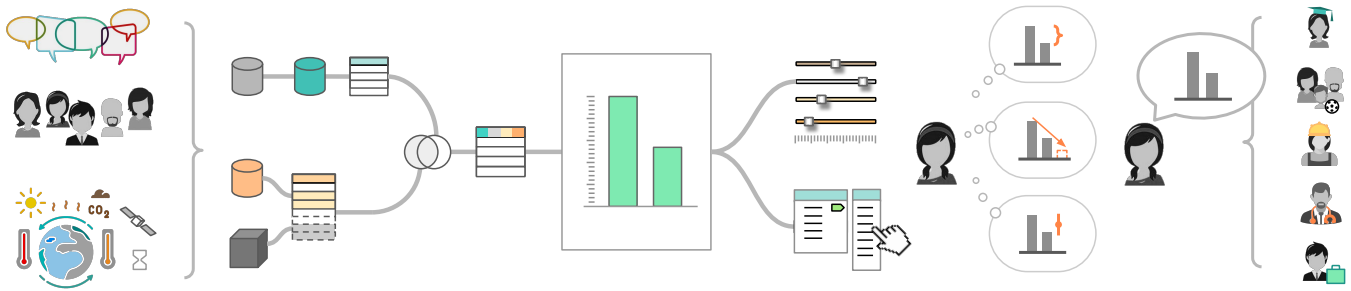


Fig. 3: Even a superficially “simple” visualization like a bar chart can exhibit entire chains of interconnected complexities, each with their own associated design decisions. These decisions about complexity are associated not only with the data that are eventually used to create the chart, but also how a viewer might interact with the chart, derive insights from the data contained within, and communicate those insights to diverse audiences.

politics [97]. Arguably, the discussion of visualization and interpretation complexity is of specific relevance for communication endeavors addressing heterogeneous audiences with a large share of non-expert or casual users, where prior knowledge, literacy and motivation to engage with parsimonious designs vary starkly [68, 110].

Strategies to misuse simplified or distorted representations of data complexity in this context are even-aged with the whole field [23, 84] and continuously updated [56]. But also the deceptive complexification of adversarial topics (see e.g., Figure 5 in [29]) is part of the deceptive playbook. As for corrective antidotes against these visualization and complexity “for villainy” [93], the expansion of data and visualization literacy initiatives [17, 25], visualization provenance standards [112], and critical design strategies [34, 40] are of the essence.

4 DISCUSSION

We present the discussions of complexity in the previous sections of the paper not to present a complete exploration of the concept of complexity as it might relate to visualization—as previously mentioned, given the polysemic and occasionally ambiguous nature of complexity as a concept, any attempt to be complete is likely doomed to failure. Rather, we present our axes of complexity to highlight two important aspects overlooked in existing debates around complexity:

(1) To highlight that **complexity is not an exclusive property of the visual design of a chart**. Even an almost prototypical “simple” visualization like a bar chart can be the product of (and result in) a complex web of interconnected complexities (Figure 3), akin to the “referential” [80] and “tentacular” [58] entanglements traced by science-and-technology studies: upstream complexities arising from the long chains of provenance and preparation of data, as well as downstream complexities, resulting from analytical and interpretative tasks, and even from strategies for communication.

(2) To show how these various kinds of **complexity are contingent and interconnected**. Complexity that is seemingly reduced at one stage of the process can pop up again, in another guise, in other sections of the pipeline. Decisions about complexity have interconnected upstream and downstream effects on the overall complexity in a visualization when perceived of as not just a visual design, but an entire system of sense-making. Apparent reductions in complexity could, therefore, be more like tradeoffs: reductions of complexity in one area of the sense-making process that incur corresponding costs in another area: complexity may not be truly reduced, but merely shifted around.

We therefore echo calls from other areas of HCI and design—to “recognize” [5], “live with” [100], or “dance with” [49] complexity in our own visualization design practices. We likewise echo the call from Akbaba et al. [2] to remove terms like “chartjunk” from our vocabulary as ultimately limiting and inflexible for describing when, how, and why to make use of complexity in designs. We believe, instead, that the concept of a *design material* captures the way that complexity can be *strategically* employed in the design process, just as other concepts like AI can likewise be thought of as design materials [35, 65] whose legibility and utility improve as a function of growing a designer’s familiarity

and expertise with their incorporation into the design process.

We point to two examples of the strategic use of complexity that arise from our consideration of the pipeline. The first is an extension of the previously mentioned concepts of “beneficial difficulties” for visualization [67], and related “slow analytics” [21], the seemingly counter-intuitive notion that including additional visual complexity (for the former) or interaction complexity (for the later) can produce benefits in terms of engagement or understanding. If the goal of visualization is merely minimalist efficiency, then these design choices would seem to be nonsensical. But, just as speed bumps and traffic signals are ubiquitous elements of traffic flow design, so too are these strategic uses of complexity ultimately functional. The second example is that of the use of complexity to support interpretative goals beyond the efficient extraction of values. One motivating example is the Poemage project [90]: the designers of the visualization tool were uncomfortable presenting a visually complex and difficult to untangle visualization of connections in poetry. But the critics, poets, and close readers that used the tool saw value in the resulting “beautiful mess” that supported hermeneutical patterns of sensemaking that thrive off of hitherto unseen connections and enrichment rather than reduction.

There are consequences to the current visual-centric view of complexity in visualization. As we discuss in prior sections, many crucial aspects of a data visualization (like data provenance or intended audience) are often overlooked in the way that we present and evaluate visualizations, to our peril. For instance, the numerous “visualization mirages” [92] that can occur where the insight purportedly extracted from a chart does not survive scrutiny, or the ways that visualizations can be mis-applied or misused once they are appropriated by unintended audiences [81]. The job of a visualization designer is not finished merely because they have generated the visually simplest chart that contains all of the data of interest.

4.1 A Call to Action and Future Work

To revisit a metaphor from our introduction, just as making proper use of color as a design material in visualization requires knowledge of the human visual system, research on color perception, and potential assistive systems for, e.g., selecting color-blind safe but effective color palettes, so too do we believe that treating complexity as a design material seriously will require considerable future effort of researchers, designers, and practitioners of data visualization. We point to several areas where we see either a lack of existing research, or existing opportunities to translate or apply research from other fields to the specific case of complexity for visualization. In particular:

(1) New **metrics and measurements** for complexity, especially non-visual complexities. For visual complexity, in addition to contentious metrics like Tufte’s data-ink ratio [132], there are also metrics motivated by psychophysics for concepts like visual clutter [116]. What are equivalent metrics for our other axes of complexity? Are there ways of capturing, for instance, that a particular data set is reliant on a highly complex set of underlying data transformations? Or that the intended audience of a visualization might have a diverse and complex array of

incompatible analytical tasks or objectives?

(2) Empirical study of **intersections and inter-relations** between differing types of complexity. Our supposition that complexity is often not *eliminated* by actions taken to reduce it, but often merely *moved* to other facets of the sense-making process, is just that: a *supposition*, supported by adages from design like Tesler's law [117] and by point examples from the brainstorming sessions that formed the core of this paper. Further research is needed to establish the borders and constraints of such assertions in visualization design, and, particularly, to assess *where* complexity shifts when it is reduced with respect to one axis.

(3) New designs and techniques for **usefully surfacing** complexity in visualizations. We lack consistent and widely adopted techniques (or even strategies) for surfacing the "hidden" complexities in visualizations, or places for designers to signal where complexity was intentionally reduced (or increased).

(4) Lastly, new **pedagogies and frameworks** for teaching and conceptualizing visualization design. Somewhat ironically, we call for a more nuanced and ultimately *complex* view of complexity in the way that we think of visualization design.

ACKNOWLEDGMENTS

This work was initiated by the Dagstuhl Seminar 23381: "Visualization and the Humanities: Towards a Shared Research Agenda." We also thank Linda Freyberg and Bridget Moynihan for their help in the conceptualization of this paper, as well as Mathieu Jacomy, Fabian Offert, and Menna El-Assady for their contributions to the initial discussion that led to this work.

REFERENCES

- [1] K. Ajani, E. Lee, C. Xiong, C. N. Knaffic, W. Kemper, and S. Franconeri. Declutter and focus: Empirically evaluating design guidelines for effective data communication. *IEEE Transactions on Visualization and Computer Graphics*, 28(10):3351–3364, 2021. 3, 5
- [2] D. Akbaba, J. Wilburn, M. T. Nance, and M. Meyer. Manifesto for putting 'chartjunk' in the trash 2021! *arXiv preprint arXiv:2109.10132*, 2021. 3, 5, 7
- [3] S. Alspaugh, N. Zokaei, A. Liu, C. Jin, and M. A. Hearst. Futzing and moseying: Interviews with professional data analysts on exploration practices. *IEEE transactions on visualization and computer graphics*, 25(1):22–31, 2018. 4
- [4] 2022 state of data science. <https://www.anaconda.com/resources/whitepapers/state-of-data-science-report-2022>, 2020. 4
- [5] A. W. Andersson, A. Jansson, B. Sandblad, and S. Tschirner. Recognizing complexity: Visualization for skilled professionals in complex work situations. In *Building Bridges: HCI, Visualization, and Non-formal Modeling*, pp. 47–66. Springer, 2014. 7
- [6] R. Arias-Hernandez, T. M. Green, and B. Fisher. From cognitive amplifiers to cognitive prostheses: Understandings of the material basis of cognition in visual analytics. *Interdisciplinary science reviews*, 37(1):4–18, 2012. 2
- [7] W. B. Arthur. Foundations of complexity economics. *Nature Reviews Physics*, 3(2):136–145, 2021. 2
- [8] M. Q. W. Baldonado, A. Woodruff, A. Kuchinsky, et al. Guidelines for using multiple views in information visualization. In *Advanced Visual Interfaces*, vol. 10, pp. 345513–345271, 2000. 3, 5
- [9] A. Bares, D. F. Keefe, and F. Samsel. Close reading for visualization evaluation. *IEEE Computer Graphics and Applications*, 40(4):84–95, 2020. 6
- [10] L. Bartram, M. Correll, and M. Tory. Untidy data: The unreasonable effectiveness of tables. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):686–696, 2021. 4
- [11] S. Bateman, R. L. Mandryk, C. Gutwin, A. Genest, D. McDine, and C. Brooks. Useful junk? The effects of visual embellishment on comprehension and memorability of charts. In *Proc. of the SIGCHI conference on human factors in computing systems*, pp. 2573–2582, 2010. 3
- [12] M. Behrisch, M. Blumenschein, N. W. Kim, L. Shao, M. El-Assady, J. Fuchs, D. Seebacher, A. Diehl, U. Brandes, H. Pfister, et al. Quality metrics for information visualization. In *Computer Graphics Forum*, vol. 37, pp. 625–662. Wiley Online Library, 2018. 3
- [13] C. H. Bennett. *How to define complexity in physics, and why*. na, 1990. 2
- [14] E. Bertini, M. Correll, and S. Franconeri. Why shouldn't all charts be scatter plots? beyond precision-driven visualizations. In *2020 IEEE Visualization Conference (VIS)*, pp. 206–210. IEEE, 2020. 1, 5, 6
- [15] M. Boisoit and B. McKelvey. Complexity and organization-environment relations: Revisiting ashby's law of requisite variety. *The Sage handbook of complexity and management*, pp. 279–298, 2011. 1
- [16] M. A. Borkin, A. A. Vo, Z. Bylinskii, P. Isola, S. Sunkavalli, A. Oliva, and H. Pfister. What makes a visualization memorable? *IEEE transactions on visualization and computer graphics*, 19(12):2306–2315, 2013. 1, 5
- [17] K. Börner, A. Bueckle, and M. Ginda. Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. *Proceedings of the National Academy of Sciences*, 116(6):1857–1864, 2019. 7
- [18] C. Bors, T. Gschwandtner, and S. Miksch. Capturing and visualizing provenance from data wrangling. *IEEE computer graphics and applications*, 39(6):61–75, 2019. 4
- [19] J. Boy, L. Eveillard, F. Detienne, and J.-D. Fekete. Suggested interactivity: Seeking perceived affordances for information visualization. *IEEE transactions on visualization and computer graphics*, 22(1):639–648, 2015. 5
- [20] A. J. Bradley, M. El-Assady, K. Coles, E. Alexander, M. Chen, C. Collins, S. Jänicke, and D. J. Wrisley. Visualization and the digital humanities. *IEEE computer graphics and applications*, 38(6):26–38, 2018. 4
- [21] A. J. Bradley, H. Mehta, M. Hancock, and C. Collins. Visualization, digital humanities, and the problem of instrumentalism. In *Proc. Workshop Vis. Digit. Hum.*, 2016. 5, 6, 7
- [22] M. Brehmer and T. Munzner. A multi-level typology of abstract visualization tasks. *IEEE transactions on visualization and computer graphics*, 19(12):2376–2385, 2013. 6
- [23] A. Cairo. *How charts lie: Getting smarter about visual information*. WW Norton & Company, 2019. 7
- [24] S. P. Callahan, J. Freire, E. Santos, C. E. Scheidegger, C. T. Silva, and H. T. Vo. Vistrails: visualization meets data management. In *Proceedings of the 2006 ACM SIGMOD international conference on Management of data*, pp. 745–747, 2006. 4
- [25] J. D. Camba, P. Company, and V. L. Byrd. Identifying deception as a critical component of visualization literacy. *IEEE Computer Graphics and Applications*, 42(1):116–122, 2022. 7
- [26] B. Castellani and F. W. Hafferty. *Sociology and complexity science: A new field of inquiry*. Springer Science & Business Media, 2009. 2
- [27] M. Chen and H. Jänicke. An information-theoretic framework for visualization. *IEEE transactions on visualization and computer graphics*, 16(6):1206–1215, 2010. 3
- [28] M. Correll. What do we actually learn from evaluations in the "heroic era" of visualization?: Position paper. In *2020 IEEE Workshop on Evaluation and Beyond-Methodological Approaches to Visualization (BELIV)*, pp. 48–54. IEEE, 2020. 4
- [29] M. Correll and J. Heer. Black hat visualization. In *Workshop on Dealing with Cognitive Biases in Visualisations (DECISIVE)*, *IEEE VIS*, vol. 1, p. 10, 2017. 7
- [30] M. Correll and J. Heer. Regression by eye: Estimating trends in bivariate visualizations. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 1387–1396, 2017. 6
- [31] A. Crisan and M. Correll. User ex machina: Simulation as a design probe in human-in-the-loop text analytics. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–16, 2021. 5
- [32] Z. Cutler, K. Gadhav, and A. Lex. Trrack: A library for provenance-tracking in web-based visualizations. In *2020 IEEE Visualization Conference (VIS)*, pp. 116–120. IEEE, 2020. 4
- [33] V. Dhanoa, C. Walchshofer, A. Hinterreiter, H. Stitz, E. Groeller, and M. Streit. A process model for dashboard onboarding. In *Computer Graphics Forum*, vol. 41, pp. 501–513. Wiley Online Library, 2022. 6
- [34] M. Dörk, P. Feng, C. Collins, and S. Carpendale. Critical infovis: exploring the politics of visualization. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pp. 2189–2198, 2013. 4, 7
- [35] G. Dove, K. Halskov, J. Forlizzi, and J. Zimmerman. Ux design innovation: Challenges for working with machine learning as a design material. In *Proceedings of the 2017 chi conference on human factors in computing systems*, pp. 278–288, 2017. 7
- [36] P. Dragicevic, Y. Jansen, A. Sarma, M. Kay, and F. Chevalier. Increasing the transparency of research papers with explorable multiverse analyses. In *proceedings of the 2019 chi conference on human factors in computing systems*, pp. 1–15, 2019. 4
- [37] J. Drucker. Humanities approaches to graphical display. *Digital Humanities*

- ties Quarterly*, 5(1):1–21, 2011. 4
- [38] J. Drucker. Graphical approaches to the digital humanities. *A new companion to digital humanities*, pp. 238–250, 2015. 4
- [39] J. Drucker. *Visualization and interpretation: Humanistic approaches to display*. MIT Press, 2020. 4
- [40] C. d’Ignazio and L. F. Klein. Feminist data visualization. Workshop on Visualization for the Digital Humanities (VIS4DH), Baltimore. IEEE., 2016. 7
- [41] M. Espadoto, R. M. Martins, A. Kerren, N. S. Hirata, and A. C. Telea. Toward a quantitative survey of dimension reduction techniques. *IEEE transactions on visualization and computer graphics*, 27(3):2153–2173, 2019. 5
- [42] R. Faust, D. Glickenstein, and C. Scheidegger. Dimreader: Axis lines that explain non-linear projections. *IEEE transactions on visualization and computer graphics*, 25(1):481–490, 2018. 6
- [43] D. Fisher and M. Meyer. *Making data visual: a practical guide to using visualization for insight*. " O’Reilly Media, Inc.", 2017. 4
- [44] S. Flensburg and S. Lomborg. Datafication research: Mapping the field for a future agenda. *New Media & Society*, 25(6):1451–1469, 2023. 4
- [45] R. L. Flood and E. R. Carson. *Dealing with complexity: an introduction to the theory and application of systems science*. Springer Science & Business Media, 2013. 2
- [46] F. Fragapane. Creating infographics and engaging data visualizations. Online video of a conference talk, 2022. Talk presented at the Adobe Max 2022 Conference. 6
- [47] S. L. Franconeri, L. M. Padilla, P. Shah, J. M. Zacks, and J. Hullman. The science of visual data communication: What works. *Psychological Science in the public interest*, 22(3):110–161, 2021. 6
- [48] Y. Fu and J. Stasko. More than data stories: Broadening the role of visualization in contemporary journalism. *IEEE Transactions on Visualization and Computer Graphics*, 2023. 6
- [49] O. Garcia de la Cerda and M. Soledad Saavedra Ulloa. Enactive management: dancing with uncertainty and complexity. *Kybernetes*, 43(8):1237–1247, 2014. 2, 7
- [50] T. Gebru, J. Morgenstern, B. Vecchione, J. W. Vaughan, H. Wallach, H. D. Iii, and K. Crawford. Datasheets for datasets. *Communications of the ACM*, 64(12):86–92, 2021. 4
- [51] M. Gleicher. Explainers: Expert explorations with crafted projections. *IEEE transactions on visualization and computer graphics*, 19(12):2042–2051, 2013. 2, 6
- [52] S. Glouberman, B. Zimmerman, et al. Complicated and complex systems: what would successful reform of medicare look like? 2002. 2
- [53] R. Grundmann. Climate change and knowledge politics. *Environmental politics*, 16(3):414–432, 2007. 4
- [54] P. Haghighatkah, M. El-Assady, J.-D. Fekete, N. Mahyar, C. Paradis, V. Simaki, and B. Speckmann. Characterizing uncertainty in the visual text analysis pipeline. In *2022 IEEE 7th Workshop on Visualization for the Digital Humanities (VIS4DH)*, pp. 25–30. IEEE, 2022. 3
- [55] J. Hale. Information-theoretical complexity metrics. *Language and Linguistics Compass*, 10(9):397–412, 2016. 2
- [56] M. N. Hannah. A conspiracy of data: Qanon, social media, and information visualization. *Social Media+ Society*, 7(3):20563051211036064, 2021. 7
- [57] D. Haraway. Situated knowledges: The science question in feminism and the privilege of partial perspective 1. In *Women, science, and technology*, pp. 455–472. Routledge, 2013. 4
- [58] D. Haraway. Staying with the trouble for multispecies environmental justice. *Dialogues in Human Geography*, 8(1):102–105, 2018. 7
- [59] N. K. Hayles. *How we think: Digital media and contemporary technogenesis*. University of Chicago Press, 2012. 2
- [60] F. Heylighen et al. The science of self-organization and adaptivity. *The encyclopedia of life support systems*, 5(3):253–280, 2001. 2
- [61] R. L. Hill, H. Kennedy, and Y. Gerrard. Visualizing junk: Big data visualizations and the need for feminist data studies. *Journal of Communication Inquiry*, 40(4):331–350, 2016. 3, 5
- [62] S. Hill, B. Wray, and C. Sibona. Minimalism in data visualization: Perceptions of beauty, clarity, effectiveness, and simplicity. *Journal of Information Systems Applied Research*, 11(1):34, 2018. 3, 5
- [63] U. Hinrichs, M. El-Assady, A. J. Bradely, S. Forlini, and C. Collins. Risk the drift! stretching disciplinary boundaries through critical collaborations between the humanities and visualization. 2017. 4
- [64] A. Hintz, L. Dencik, and K. Wahl-Jorgensen. *Digital citizenship in a datafied society*. John Wiley & Sons, 2018. 4
- [65] L. E. Holmquist. Intelligence on tap: artificial intelligence as a new design material. *interactions*, 24(4):28–33, 2017. 7
- [66] J. Hullman. The purpose of visualization is insight, not pictures: An interview with ben shneiderman. *ACM Interactions*, 2019. 6
- [67] J. Hullman, E. Adar, and P. Shah. Benefitting infovis with visual difficulties. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2213–2222, 2011. 1, 3, 5, 6, 7
- [68] Y.-H. Hung and P. Parsons. Affective engagement for communicative visualization: quick and easy evaluation using survey instruments. In *conference: visualization for communication (VisComm): an IEEE VIS WorkshopAt: Berlin, Germany*, 2018. 7
- [69] O. Inbar, N. Tractinsky, and J. Meyer. Minimalism in information visualization: attitudes towards maximizing the data-ink ratio. In *Proceedings of the 14th European conference on Cognitive ergonomics: invent! explore!*, pp. 185–188, 2007. 1, 3, 5
- [70] J. Kallinikos. The order of technology: Complexity and control in a connected world. *Information and organization*, 15(3):185–202, 2005. 2
- [71] S. Kandel, J. Heer, C. Plaisant, J. Kennedy, F. Van Ham, N. H. Riche, C. Weaver, B. Lee, D. Brodbeck, and P. Buono. Research directions in data wrangling: Visualizations and transformations for usable and credible data. *Information Visualization*, 10(4):271–288, 2011. 4
- [72] S. A. Kauffman. *The origins of order: Self-organization and selection in evolution*. Oxford University Press, USA, 1993. 2
- [73] S. Kernbach, M. J. Eppler, and S. Bresciani. The use of visualization in the communication of business strategies: An experimental evaluation. *Int Journal of Business Communication*, 52(2):164–187, 2015. 6
- [74] T. Kinni. The critical difference between complex and complicated. *MIT Sloan Management Review*, 21, 2017. 2
- [75] G. J. Klir and W. R. Ashby. Requisite variety and its implications for the control of complex systems. *Facets of systems science*, pp. 405–417, 1991. 2
- [76] M. Koerth, L. Bronner, and J. Mithani. Why it’s so freaking hard to make a good covid-19 model. <https://fivethirtyeight.com/features/why-its-so-freaking-hard-to-make-a-good-covid-19-model/>, 2020. Accessed: 09.07.2024. 5
- [77] H. Lam. A framework of interaction costs in information visualization. *IEEE transactions on visualization and computer graphics*, 14(6):1149–1156, 2008. 5, 6
- [78] H. Lamqaddam, A. V. Moere, V. V. Abeele, K. Brosens, and K. Verbert. Introducing layers of meaning (lom): A framework to reduce semantic distance of visualization in humanistic research. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1084–1094, 2020. 4
- [79] D. Lange, S. Sahai, J. M. Phillips, and A. Lex. Ferret: Reviewing tabular datasets for manipulation. In *Computer Graphics Forum*, vol. 42, pp. 187–198. Wiley Online Library, 2023. 5
- [80] B. Latour. Circulating reference: Sampling the soil in the amazon forest. *Pandora’s hope: Essays on the reality of science studies*, 24:36, 1999. 7
- [81] C. Lee, T. Yang, G. D. Inchoco, G. M. Jones, and A. Satyanarayan. Viral visualizations: How coronavirus skeptics use orthodox data practices to promote unorthodox science online. In *Proceedings of the 2021 CHI conference on human factors in computing systems*, pp. 1–18, 2021. 5, 7
- [82] S. Lee, S.-H. Kim, Y.-H. Hung, H. Lam, Y.-a. Kang, and J. S. Yi. How do people make sense of unfamiliar visualizations?: A grounded model of novice’s information visualization sensemaking. *IEEE transactions on visualization and computer graphics*, 22(1):499–508, 2015. 6
- [83] H. Li and N. Moacdieh. Is “chart junk” useful? an extended examination of visual embellishment. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 58, pp. 1516–1520. Sage Publications Sage CA: Los Angeles, CA, 2014. 3, 5
- [84] M. Lisnic, C. Polychronis, A. Lex, and M. Kogan. Misleading beyond visual tricks: How people actually lie with charts. In *Proceedings of the 2023 CHI conference on human factors in computing systems*, pp. 1–21, 2023. 7
- [85] Z. Liu, N. Nersessian, and J. Stasko. Distributed cognition as a theoretical framework for information visualization. *IEEE transactions on visualization and computer graphics*, 14(6):1173–1180, 2008. 2
- [86] S. Mann. Systems of creation: the emergence of life from nonliving matter. *Accounts of chemical research*, 45(12):2131–2141, 2012. 2
- [87] L. Manovich. Data visualization as new abstraction and anti-sublime. *Small tech: The culture of digital tools*, 3(9), 2002. 5
- [88] K. Marriott, B. Lee, M. Butler, E. Cutrell, K. Ellis, C. Goncu, M. Hearst, K. McCoy, and D. A. Szafir. Inclusive data visualization for people with

- disabilities: a call to action. *Interactions*, 28(3):47–51, 2021. 6
- [89] F. Mazzocchi. Complexity in biology: Exceeding the limits of reductionism and determinism using complexity theory. *EMBO reports*, 9(1):10–14, 2008. 2
- [90] 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, 2015. 7
- [91] W. H. McNeill. The human web. In *Teaching World History in the Twenty-first Century: A Resource Book*, pp. 23–25. Routledge, 2015. 2
- [92] A. McNutt, G. Kindlmann, and M. Correll. Surfacing visualization mirages. In *Proceedings of the 2020 CHI Conference on human factors in computing systems*, pp. 1–16, 2020. 3, 7
- [93] A. M. McNutt, L. Huang, and K. Koenig. Visualization for villainy. *arXiv preprint arXiv:2109.06007*, 2021. 7
- [94] T. Metze. Visualization in environmental policy and planning: A systematic review and research agenda. *Journal of Environmental Policy & Planning*, 22(5):745–760, 2020. 3
- [95] M. Muller, I. Lange, D. Wang, D. Piorkowski, J. Tsay, Q. V. Liao, C. Dugan, and T. Erickson. How data science workers work with data: Discovery, capture, curation, design, creation. In *Proc. of the 2019 CHI conference on human factors in computing systems*, pp. 1–15, 2019. 4
- [96] T. Munzner. *Visualization analysis and design*. CRC press, 2014. 5
- [97] T. U. Nærland. The political significance of data visualization: Four key perspectives. *Data visualization in society*, pp. 63–74, 2020. 7
- [98] R. Niño Zambrano and Y. Engelhardt. Diagrams for the masses: Raising public awareness—from neurath to gapminder and google earth. In *Diagrammatic Representation and Inference: 5th International Conference, Diagrams 2008, Herrsching, Germany, Sep. 19-21, 2008.*, pp. 282–292. Springer, 2008. 6
- [99] C. Nobre, K. Zhu, E. Mörth, H. Pfister, and J. Beyer. Reading between the pixels: Investigating the barriers to visualization literacy. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp. 1–17, 2024. 6
- [100] D. A. Norman. *Living with complexity*. MIT press, 2016. 1, 2, 7
- [101] C. North. Toward measuring visualization insight. *IEEE computer graphics and applications*, 26(3):6–9, 2006. 6
- [102] D. Offenhuber. Data by proxy—material traces as autographic visualizations. *IEEE transactions on visualization and computer graphics*, 26(1):98–108, 2019. 4
- [103] E. E. Olson and G. H. Eoyang. *Facilitating organization change: Lessons from complexity science*, vol. 5. John Wiley & Sons, 2001. 2
- [104] A. Ottley. *Adaptive and personalized visualization*. Springer Nature, 2022. 6
- [105] F. Paas, T. Van Gog, and J. Sweller. Cognitive load theory: New conceptualizations, specifications, and integrated research perspectives. *Educational psychology review*, 22:115–121, 2010. 2, 5
- [106] P. Parsons and P. Shukla. Data visualization practitioners’ perspectives on chartjunk. In *2020 IEEE Visualization Conference (VIS)*, pp. 211–215. IEEE, 2020. 3, 5
- [107] S. E. Phelan. What is complexity science, really? *Emergence, A Journal of Complexity Issues in Organizations and Management*, 3(1):120–136, 2001. 2
- [108] R. Poli. A note on the difference between complicated and complex social systems. *Cadmus*, 2(1):142, 2013. 2
- [109] A. J. Porter, T. R. Kuhn, and B. Nerlich. Organizing authority in the climate change debate: Ippc controversies and the management of dialectical tensions. *Organization Studies*, 39(7):873–898, 2018. 4
- [110] Z. Pousman, J. 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. 6, 7
- [111] X. Pu and M. Kay. The garden of forking paths in visualization: A design space for reliable exploratory visual analytics: Position paper. In *2018 IEEE evaluation and beyond-methodological approaches for visualization (beliv)*, pp. 37–45. IEEE, 2018. 4
- [112] E. D. Ragan, A. Endert, J. Sanyal, and J. Chen. Characterizing provenance in visualization and data analysis: an organizational framework of provenance types and purposes. *IEEE transactions on visualization and computer graphics*, 22(1):31–40, 2015. 7
- [113] N. Rescher. *Complexity: A philosophical overview*. Routledge, 2020. 2
- [114] P. J. Richerson and R. Boyd. *Not by genes alone: How culture transformed human evolution*. University of Chicago press, 2008. 2
- [115] N. Rokotyan, O. Stukova, and D. Kolmakova. An alternative data-driven country map. <https://projects.interacta.io/country-tsne/>, 2019. Accessed: 14.07.2024. 6
- [116] R. Rosenholtz, Y. Li, and L. Nakano. Measuring visual clutter. *Journal of vision*, 7(2):17–17, 2007. 7
- [117] D. Saffer. *Designing for interaction: creating innovative applications and devices*. New Riders, 2010. 1, 8
- [118] J. Salminen, K. Wenyun Guan, S.-G. Jung, and B. Jansen. Use cases for design personas: A systematic review and new frontiers. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–21, 2022. 6
- [119] A. Sarma, K. Hwang, J. Hullman, and M. Kay. Milliways: Taming multiverses through principled evaluation of data analysis paths. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp. 1–15, 2024. 4
- [120] R. K. Sawyer. *Social emergence: Societies as complex systems*. Cambridge University Press, 2005. 2
- [121] R. Schuster, K. Gregory, T. Möller, and L. Koesten. “being simple on complex issues”—accounts on visual data communication about climate change. *IEEE Transactions on Visualization and Computer Graphics*, 2024. 3, 6
- [122] 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. 3
- [123] C. Southerton. Datafication. In *Encyclopedia of big data*, pp. 358–361. Springer, 2022. 4
- [124] A. Springer and S. Whittaker. Progressive disclosure: empirically motivated approaches to designing effective transparency. In *Proceedings of the 24th international conference on intelligent user interfaces*, pp. 107–120, 2019. 6
- [125] C. Stoiber, M. Wagner, F. Grassinger, M. Pohl, H. Stitz, M. Streit, B. Potzmann, and W. Aigner. Visualization onboarding grounded in educational theories. In *Visualization Psychology*, pp. 139–164. Springer, 2023. 6
- [126] M. Stone. Get it right in black and white. *Functional Color*, pp. 140–142, 2010. 2
- [127] D. A. Szafir, S. Haroz, M. Gleicher, and S. Franconeri. Four types of ensemble coding in data visualizations. *Journal of vision*, 16(5):11–11, 2016. 6
- [128] P. Taborsky. Is complexity a scientific concept? *Studies in History and Philosophy of Science Part A*, 47:51–59, 2014. 2
- [129] A. Teixeira de Melo, L. S. D. Caves, A. Dewitt, E. Clutton, R. Macpherson, and P. Garnett. Thinking (in) complexity:(in) definitions and (mis) conceptions. *Systems Research and Behavioral Science*, 37(1):154–169, 2020. 2
- [130] J. J. Thomas and K. A. Cook. A visual analytics agenda. *IEEE computer graphics and applications*, 26(1):10–13, 2006. 2
- [131] C. Tominski and H. Schumann. *Interactive visual data analysis*. AK Peters/CRC Press, 2020. 5
- [132] E. Tufte and P. Graves-Morris. The visual display of quantitative information.; 1983. *Diagrammatik-Reader. Grundlegende Texte aus Theorie und Geschichte. Berlin: De Gruyter*, pp. 219–230, 2014. 1, 3, 5, 7
- [133] R. J. Wenke. Explaining the evolution of cultural complexity: A review. *Advances in archaeological method and theory*, pp. 79–127, 1981. 2
- [134] F. Windhager, G. Schreder, and E. Mayr. On inconvenient images: Exploring the design space of engaging climate change visualizations for public audiences. In *EnvirVis@ EuroVis*, pp. 1–8, 2019. 3, 6
- [135] J. Wood, R. Beecham, and J. Dykes. Moving beyond sequential design: Reflections on a rich multi-channel approach to data visualization. *IEEE transactions on visualization and computer graphics*, 20(12):2171–2180, 2014. 6
- [136] D. D. Woods. Coping with complexity: the psychology of human behaviour in complex systems. In *Tasks, errors, and mental models*, pp. 128–148. 1988. 2
- [137] S. Yee and T. Chu. A visual introduction to machine learning—part ii: Model tuning and the bias-variance tradeoff. <http://www.r2d3.us/visual-intro-to-machine-learning-part-2/>, 2018. Accessed: 14.07.2024. 6
- [138] J. S. Yi, Y. ah Kang, J. Stasko, and J. A. Jacko. Toward a deeper understanding of the role of interaction in information visualization. *IEEE transactions on visualization and computer graphics*, 13(6):1224–1231, 2007. 5
- [139] C. Ziemkiewicz and R. Kosara. The shaping of information by visual metaphors. *IEEE transactions on visualization and computer graphics*, 14(6):1269–1276, 2008. 6