Input Visualization:
Collecting and Modifying Data with Visual Representations

Nathalie Bressa
i3 (UMR 9217), CNRS,
Télécom Paris, Institut Polytechnique de Paris
France
nathalie.bressa@telecom-paris.fr

Jordan Louis
i3 (UMR 9217), CNRS,
Télécom Paris, Institut Polytechnique de Paris
France
jordan.louis@telecom-paris.fr

Wesley Willett
University of Calgary
Canada
wesley.willett@ucalgary.ca

Samuel Huron
i3 (UMR 9217), CNRS,
Télécom Paris, Institut Polytechnique de Paris
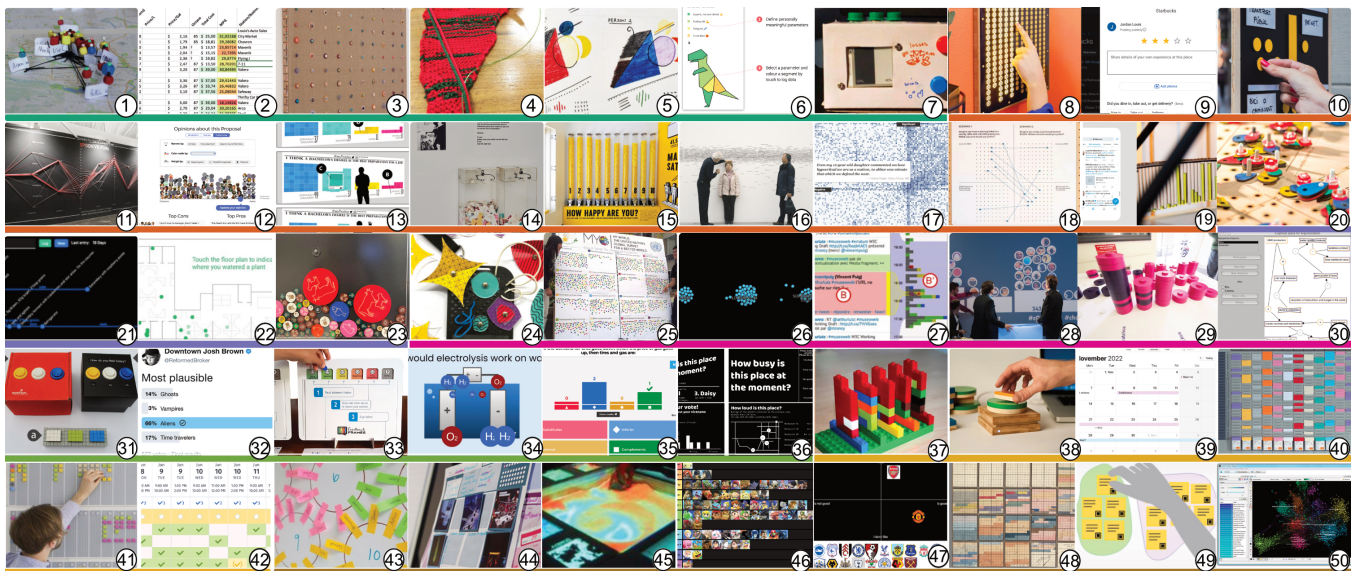France
samuel.huron@telecom-paris.fr

Figure 1: The 50 input visualizations in our collection, organized into seven purposes (■ Individual Reflection, ■ Public Group Reflection, ■ Public Activity Documentation, ■ Data Discussion, ■ Survey, ■ Planning, and ■ Organizing). Examples are numbered from ①—㊿. Browsable version: https://bit.ly/input-Vis. Supplemental material: https://osf.io/bw3gp.

ABSTRACT

We examine *input visualizations*, visual representations that are designed to collect (and represent) new data rather than encode preexisting datasets. Information visualization is commonly used to reveal insights and stories within existing data. As a result, most contemporary visualization approaches assume existing datasets as the starting point for design, through which that data is mapped to visual encodings. Meanwhile, the implications of visualizations as inputs and as data sources have received little attention—despite the existence of visual and physical examples stretching back centuries. In this paper, we present a design space of 50 input visualizations analyzing their visual representation, data, artifact, context, and input. Based on this, we identify input modalities, purposes of input visualizations, and a set of design considerations. Finally, we discuss the relationship between input visualization and traditional visualization design and suggest opportunities for future research to better understand these visual representations and their potential.

CCS CONCEPTS

• Human-centered computing → Information visualization.

KEYWORDS

input visualization, physicalization, data collection, interaction, visualization, participation, survey, public engagement, data discussion

1 INTRODUCTION

Information visualization is typically thought of as a set of methods and approaches for giving visual structure to existing datasets, leveraging visual perception to enhance the analysis and interpretation of data [18]. In most information visualization models, pipelines, and tools, data serves as the starting point for the design or analysis process, after which designers, developers, and analysts select visual mappings to make that data more legible and actionable [18, 22, 30]. Over the past 50 years, a large body of research has successfully focused on optimizing visual mappings and interactions, creating a diversity of different visualization genres tailored to unique data, tasks, audiences, and contexts. Most of these approaches implicitly assume that 1) the data (or its characteristics) are known in advance, 2) the principal goal of the visualization is to reveal trends and features in the underlying data, and 3) interactions with the visualizations (filtering, computing new values, etc.) do not alter the original underlying data.

Yet, a variety of visualization and visualization-like approaches exist which eschew the "data-first" orthodoxy of the academic information visualization community and instead use visualizations as mechanisms for data input. We define *input visualizations*—visual representations that are designed to collect and/or modify new data rather than encode pre-existing datasets. Notably, input visualizations are characterized not by their form (which includes a wide cross-section of common visualization types) or how they are created, but instead by their intended use. By emphasizing data input and modification rather than traditional visualization goals like perception, exploration, or communication, input visualizations force a rethinking of the relationship between data and visuals. Recognizing the potential for visualizations as data inputs also reveals unexamined use cases for visualization, suggests new interaction techniques, and highlights opportunities for visualization tools that support new ways of engaging with data.

Examples of input visualizations include common tools like Doodle [35], which uses a preference matrix as a data input and collection mechanism, as well as to represent the collected data. Numerous such examples of charts used as a data collection medium exist in different domains, such as data journalism, civic participation, time management, and education [35, 55, 66, 68]. In fact, almost all early examples of external visual representation of information (including tally marks, tables, and astronomical diagrams) are arguably input visualizations according to our definition. Everyday tools like physical and digital calendars also fit this definition—specifying a data schema by encoding periods of time using daily, weekly, and yearly grids, then allowing individuals to define new events using that structure.

Despite this, most design guidelines in information visualization recommend first considering the dataset, the user, and the task and then designing the appropriate visual representations. These guidelines do not focus on a dataset that remains to be collected or on a dataset that is collected through a visualization. Input visualizations invert traditional data encoding and design models, using visual structures to support the collection of data, the definition of new visual schemas, and the exploration of possible visual mappings. As a result, these approaches pose problems for classical information visualization reference models [18, 22], interaction taxonomies [61, 143], and design guidelines [30].

In this paper, we provide a first step toward conceptualizing and reframing discussions on this underconsidered area in information visualization to make it actionable for the information visualization community. We do this by both defining the concept and investigating the characteristics of existing input visualizations through a design space analysis. First, we describe and analyze four case studies to introduce the concept of input visualizations and motivate our broader analysis. Next, we describe 50 examples of input visualizations drawn from research, journalism, art, personal projects, and commercial products (examples are numbered from ①—㊿). We analyzed each example with regards to its visual representation, data, context, artifact, and input. By cross-analyzing these dimensions we identified seven input modalities and seven input visualization purposes. Drawing on this analysis, we present considerations for design that illustrate both the potential of input visualizations and challenges posed by them—discussing visibility of prior input, scalability, readability, and data processing. Finally, we highlight research opportunities and discuss the relationship between input visualizations and traditional visualization norms, the potential for future input visualizations as sensemaking tools, and new design methods. We make the following contributions:

- We offer a first definition of input visualization.
- We introduce a design space of input visualizations anchored in an analysis of 50 diverse examples. We also identify 6 input modalities and 7 purposes for input visualizations.
- Drawing from the analysis of the design space, we extract a set of design considerations and outline new opportunities for research.

2 RELATED WORK

Information visualizations created to support data input have appeared in multiple domains of human-computer interaction, including work on civic participation, community engagement, online debate, personal reflection, planning, polling, and affinity diagramming. Some examples, including tools like BitPlanner [127] and Thudt et al.'s physical self-reflection kits [132] (as well as a wide variety of other participatory physicalizations [37]) rely on physical construction. Meanwhile, others like Koeman et al.'s urban voting systems [81], Kriplean et al. [84] and Valkanova et al.'s [137] web-based polling tools, along with scheduling systems like Framadate [45], Doodle [35], and when2meet [140] all involve input via on-screen visualizations.

Information visualization research has also examined how visual marks can serve as interactive controls for interacting with the data thought the visualization, for editing existing data, and for authoring new visualizations. Some examples of using visual marks as interactive controls include approaches like DimpVis [83] and À Table [112] in which viewers can manipulate marks to navigate between visualization views and change timespans. "You Draw It" visualizations in which viewers articulate predictions by drawing on visualizations [79] have also used input as a way of drawing attention to data values and encouraging recall.

Similarly, a variety of creative visualization design tools such as Charticulator [115], Data Illustrator [91], Data Ink [142], and Lyra [122] support direct graphical input as a way of designing and placing visual marks to create new visualizations. Yet these tools generally use input only to author graphical encodings for existing tabular data, rather than adding new data points themselves. Approaches like *scented widgets* [141], meanwhile, place visualizations directly on top of interactive controls—often for visualizing distribution of values entered. However, these visualizations have generally been framed as a way of understanding and facilitating input, rather than a mechanism for data collection.

Other recent work has examined how direct manipulation interactions like changing or repositioning marks within a visualization [120] might support view transitions and visualization editing. Yet these approaches have mostly treated sketching and manipulation of visual marks as ways of interacting with datasets, not as mechanisms for data collection.

2.1 Data Input in InfoVis Reference Models

Within visualization, several alternative conceptual models have also hinted at the potential for visualizations as input mechanisms. Based on their examination of personal physicalization [132], Thudt et al. discuss opportunities for visualizations as a means of input—highlighting approaches that support qualitative data input via sketching or manual manipulation of attributes like the position, size, or color of visual marks. Meanwhile, Offenhuber's characterization of *autographic visualization* approaches [108] offers an alternative framework for considering visual representations that reflect environmental processes and typically lack explicit data structures or encoding pipelines. Offenhuber contrasts autographic approaches, which start with a phenomenon and then introduce physical interventions to reveal visual traces of it, against more traditional visualization pipelines, which first collect data from a phenomenon then render that data as visualizations. Like autographic examples, input visualizations can capture and visualize information despite the absence of explicit encodings or data structures, but are explicitly *designed as interfaces*, relying on human interaction rather than environmental processes.

To our knowledge, the only information visualization model to describe data input via visualizations is Jansen & Dragicevic's *interaction model for visualizations beyond the desktop* [74]. Their model differentiates *concrete* rendering pipelines (in which existing data is rendered as a visual or physical output) from *conceptual* pipelines (which describe data and encodings implicit in the visualization but not implemented by a rendering process). They use this model to describe two physical input visualizations—DailyStack [131] and

Hunger's LEGO time trackers [101]—highlighting how interaction with these visualizations can manifest both physical and virtual instantiations of new data. Hunger's process and visual mapping are also detailed in Huron et al.'s exploration of *constructive visualization* [69], a paradigm in which visual representations are constructed by assembling elements that represent data.

Most other information visualization taxonomies (including those from Amar et al. [1], Brehmer & Munzner [15], Chi & Riedl [22], Rubab et al. [118], and Yi et al. [143]) do not cover input at the level of the data, or do so only tangentially. Interestingly, a few specifically mention input actions in the context of *metadata*—including marking data points [143], creating, deleting, and editing notes associated with them [54], or annotating visualizations [15]. Dimara & Perin [31] while examining interaction for data visualization, come perhaps the closest, mentioning input data action as a way to "operate on raw data values" including adding data points, correcting data points, and adding metadata. More recently, Dimara et al. [32] have highlighted the need for "flexible data input" in visualization tools for decision makers to perform direct actions on the raw data including collecting, correcting, and annotating. Although all of these discussions indicate the potential for visualizations as input mechanisms, the implications and design possibilities of visualizations that use them remain largely unexamined.

3 CASE STUDIES

To illustrate the potential of input visualizations, we showcase four case studies (Figure 2) which highlight the breadth and diversity of existing designs. We examine their common elements and introduce the motivations and questions that guided our subsequent analyses.

3.1 The Death of a Terrorist: A Turning Point?

In 2011, after the killing of Osama Bin Laden, *The New York Times* published an interactive visualization titled *The Death of a Terrorist: A Turning Point?* [66]. The piece was anchored in an interactive two-dimensional scatterplot with its y-axis ranging from significant (top) to insignificant (bottom) and x-axis ranging from negative (left) to positive (right). The story invited viewers to discuss the importance of the event by clicking a point in this two-dimensional space and then authoring a comment. Individual cells in the scatterplot were then colored based on the number of responses. Subsequent visitors could hover over these cells to read the comments. During its initial run, the visualization collected 13,864 comments—all data points that did not exist when the story launched.

3.2 Doodle

Doodle is an online scheduling tool designed for facilitating meeting coordination. It allows multiple participants to collect and visualize schedule availability. The organizer invites people to indicate and visually compare preferred meeting times in a visual matrix. Each cell represents one participant's availability in a given time slot, color-coded to show blocks where they are (or are not) available. Users can examine prior respondents' preferences while entering their own, helping identify times that work for all attendees.
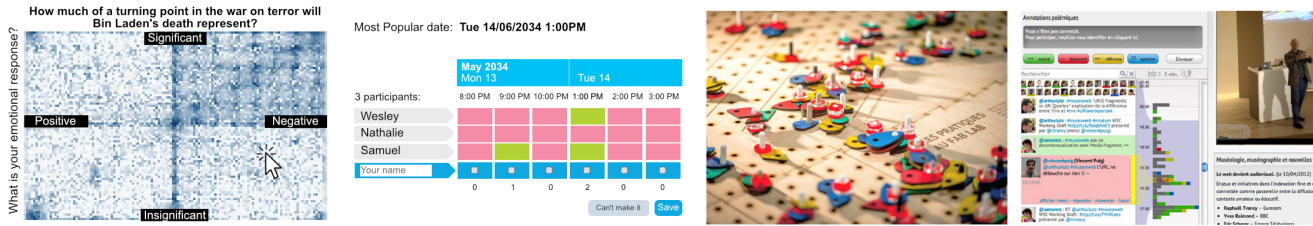
Figure 2: The four case studies: *The Death of a Terrorist: A Turning Point?* by *The New York Times* ⑰ solicited readers' reflections on the killing of Osama Bin Laden via interactions with a 2-dimensional scatterplot [66]. *Doodle* ㊷ polls allow groups to visually negotiate schedules [35]. The *Cairn* ⑳ tabletop used composable physical tokens to collect information about projects created in a shared makerspace [55]. *Polemic Tweet* ㉗ centered around a stacked bar chart which visualized the labels entered in participants' comments [68]. CC BY figure at https://osf.io/bw3gp.

3.3 Cairn

Gourlet & Dassè's *Cairn* was a tangible tabletop that enabled data collection, visualization, and analysis of activity in a shared makerspace [55]. However, in comparison to most other participatory cases, *Cairn* leveraged a more complex visual encoding schema. Using a variety of composable physical tokens and a more complex layout, this physicalization allowed makers to record detailed information about their work in the shared space—documenting the type, duration, and form of their projects, as well as qualitative information about the skills and techniques they learned. This more complex encoding captured a considerable amount of information in each visual mark, while simultaneously giving contributors opportunities for creative expression as they constructed small "cairns" out of multiple tokens.

3.4 Polemic Tweet

Similarly, the visual backchannel [36] tool *Polemic Tweet* [68] used visualizations to engage users in an evolving discussion around conference presentations. The *Polemic Tweet* interface included a Twitter client augmented with a vertical stacked bar chart. Conference participants were invited to tweet using a specific grammar ("++" for agreement, "−−" for disagreement, "==" for reference, "??" for questions). These tweets then appeared in a vertical list below the input box and in a stacked bar chart, colored according to the tags they included. The vertical axis of the visualization corresponded to a time window, and bar heights showed the number of tweets emitted at that particular time slot. Unlike *The Death of a Terrorist*, *Polemic Tweet*'s input mechanisms were not spatially overlaid with the visualization. However, the two were tightly integrated and all data visualized in the *Polemic Tweet* interface was generated in the same immediate context.

3.5 Case Study Analysis

These four case studies showcase a diversity of physical and virtual input visualization approaches—including both straightforward examples and ones that challenge visualization norms. Below, we deconstruct their common components and describe how they motivated our subsequent design space analysis. All four case studies combine VISUAL STRUCTURES, and an ASSEMBLY MODEL (the rules that govern the visualization's construction) with compatible INPUT MECHANISMS in service of a set of HIGH-LEVEL TASKS.

Visual Structure as Data Schema. The visual structure of an input visualization not only determines its appearance, but typically also defines the kinds of data that the system can collect. In contrast to most traditional visualizations, their data schema is often established using visual variables like space, color, and shape (rather than the other way around). These choices dictate the complexity of the data that viewers can input, ranging from simple choices in a grid (*Doodle*), to continuous and connected inputs (*The Death of a Terrorist*), category labels (*Polemic Tweet*), and complex multivariate data (*Cairn*). This relationship raises a variety of questions, including—which visual structures lend themselves well to input, and what types of data can they support? In response, we examine common visual idioms for input visualizations (Section 5.1) and characterize the kinds of data (Section 5.2), artifacts (Section 5.3), and contexts (Section 5.4) with which they align.

Input Mechanisms. Our case studies illustrate a variety of different input approaches. These include direct manipulation [126] interactions with on-screen visualizations (*The Death of a Terrorist, Doodle*), indirect input via interface elements closely associated with charts (*Polemic Tweet*), and physical assembly (*Cairn*). All these case studies embody an assembly model—defined by Huron et al. as "the internal model of how the constructing and deconstructing of the visual representation is carried out." [69]. In (*Doodle*), the color of the cells is updated according to user's input and in *Cairn* the rule is to assemble tokens on a stick to create a cairn that can be positioned on the table. For *The Death of a Terrorist*, the cell's hue changes according to the numbers of inputs and in *Polemic Tweet* every tweet will generate a token that will be stacked on the other tokens of the same time span. Given this diversity, we ask—what other kinds of artifacts and input modalities might lend themselves well to input visualizations? With this in mind, we characterize properties of input (Section 5.5), identifying three low-level tasks, and six relevant input modalities.

Why Input Visualizations? All four case studies use input visualization approaches to solicit opinions, sentiments, and other data from viewers. In some cases, the representation is used like a survey tool (*The Death of a Terrorist*), while in others it supports back-channel discussion (*Polemic Tweet*), schedule coordination

(*Doodle*), or documentation of fabrication projects (*Cairn*). But what other high-level tasks and social contexts typically drive people to create input visualizations? To address this, we set out to examine the settings and tasks that existing input visualizations support and highlight an initial set of seven purposes for input visualizations (Section 5.6).

4 METHOD

Despite the abundant literature on designing and analyzing visualizations themselves (including widely used textbooks [106]), the visualization community has no formally accepted methodologies for exploring and constructing design spaces [71]. Multiple definitions of the concept of a design space have been provided in HCI based on design rationales [96], constraints [14], or as a conceptual exploration of ideas [12, 60]. As mentioned by Beaudouin-Lafon [11] design spaces can be *descriptive*, *generative*, and *evaluative*. As we aimed to characterize an uncharted phenomenon, we focused on the descriptive power of design spaces. Depending on the research questions and the types of designs contained in a design space, the dimensions of a design space can be generic and broad [6, 19], more specific and narrow [51], or a mix of both [85, 125]. In particular, we were interested in understanding the extent to which this phenomenon manifests across various visualization and data types. Thus, we focused on common foundational dimensions for this domain, allowing us to compare and contrast our results against existing visualization norms and framings.

Because input visualization is a new and emerging research area without a clear definition or distinct boundaries (but for which huge numbers of examples exist), we focused on curating a scoped set of examples that showcase diversity rather than providing an exhaustive sample. Similar approaches have been used before in design space analyses of narrative visualization [125], casual information visualization [114], and anthropographics [104].

Over the course of our exploration, we collected a set of 50 examples [2, 4, 7, 13, 16, 23, 24, 27, 28, 34, 35, 40, 41, 44, 46, 47, 49, 50, 53, 55, 58, 66–68, 70, 73, 75, 80, 81, 84, 95, 103, 107, 109, 111, 119, 123, 127, 129–135, 137, 139] (Figure 1, Table 1) all of which support data input or modification via interactions with a visual representation. Our corpus includes both visualizations and physicalizations, and draws from academic research, commercial products and services, design projects, and journalism. When an academic paper presented a collection of very similar examples [16, 132], we selected the one example we found to be the most distinct from those already in our corpus—and excluded the rest to avoid over-sampling from the same source. Similarly, when faced with many functionally similar examples (like the vast number of input-oriented spreadsheets with visual conditional formatting) we included a single examplar with which we were intimately familiar.

4.1 Concept and Corpus Selection

As we collected examples for our corpus, we encountered a range of competing definitions of "input" posed by different subsets of the visualization research community—most of which treat input as a narrower set of operations that take place on or around visualizations of existing data. We characterize a few of these perspectives (and offer our own) below:

Input as Annotation. Input can mean adding metadata to visualizations through graphical and text annotations [15, 116], marking data items of interest [143], and adding notes [54].

Input as Data Editing. Another view on input is editing existing data points or attributes in a visualization for error correction. Related approaches include interactive data editing via graph manipulation [9], tools for creating what-if analyses [63], and interactive editing of node and edge attributes in node-link diagrams [43].

Input as Collecting and/or Modifying **New** *Data.* In contrast, our focus is on the collection and modification of new data points or dimensions via visualizations that are specifically designed to support these input actions. We focus on active data collection that requires deliberate input by a person (excluding examples such as activity tracking on smartwatches) and in which the visualization is visible to the person providing the input (excluding examples such as customer satisfaction polling systems).

4.2 Collection and Analysis

The four authors (Nathalie, Jordan, Wesley, and Samuel) iteratively collected, coded, discussed, and curated our corpus of examples over a two-year period, spanning seven different analysis cycles and a variety of tools (many of which are arguably input visualizations themselves). Inspired by Meyer & Dykes [100], we provide a comprehensive description of our collection and analysis process (Figure 3), with additional details and artifacts included in supplemental material.

Cycle 1: Definition. Wesley and Samuel created a Miro mood board [93] to identify and collect an initial set of 42 possible examples. We then narrowed this corpus to a subset of 16 unique examples in a spreadsheet (Google Sheets) and analyzed them using an open coding approach [87], formalizing our first set of categories and design dimensions in an initial codebook. This cycle allowed us to formalize a definition of input visualization and write a first workshop paper [72].

Cycle 2: Building corpus and first coding. Jordan joined the project and expanded the corpus, incorporating community feedback from our initial publication. He transferred the existing corpus to a new Airtable database (which allowed us to embed snapshots of each example in the table), adding new visualizations and removing redundant ones, coding based on the previous codebook, and identifying additional design dimensions.

Cycle 3: Defining final design dimensions. Nathalie joined the project and exported the entire collection of examples as a set of cards which allowed us to physically sort and group them. Nathalie, Wesley, and Samuel used versions of these cards to create affinity diagrams [90, 94, 113], explore new clusterings, and refine this set of examples and design space dimensions.

Cycle 4: Coding. Nathalie and Samuel used Airtable to code the dataset together over the course of several days, iterating until they reached consensus across all examples and dimensions.

Cycle 5: Dimensionality reduction and data export. Wesley then extracted the data from Airtable and analyzed it using dimensionality reduction and projection (in Google Colab) to identify clusters of related examples with similar characteristics. The resulting series of 2D plots inspired us to search the corpus for designs
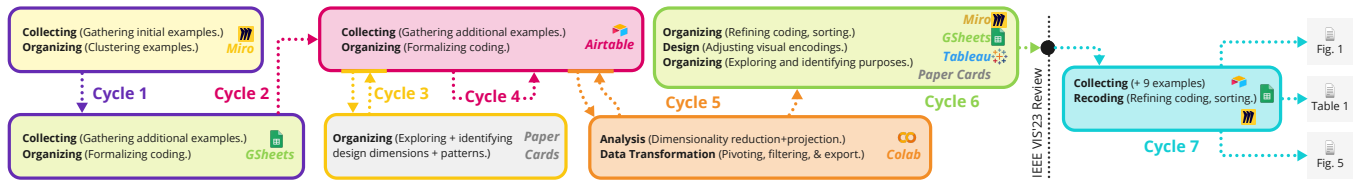
Figure 3: Overview of our collection and analysis process, including the tools we used during each cycle. See supplemental material for additional process documentation. CC BY figure at https://osf.io/bw3gp.

with related purposes. Wesley also transformed and pivoted the data, creating the basis for our final design space (Table 1).

Cycle 6: Identifying purposes. Nathalie, Wesley, and Samuel sorted and adjusted the visual encoding of the final table in a new Google Sheet, making final updates to improve the consistency of the design space coding. To identify purposes for each design Samuel created a fresh set of paper cards for each project and organized them on a physical whiteboard. Using these initial clusters as a starting point, Wesley plotted the same data digitally using Tableau, organizing the examples according to their synchronicity and number of users. Nathalie and Samuel then recreated this schema on Miro with digital cards, which they and Wesley used to finalize the coding of both high-level tasks and design purposes.

Cycle 7: Adding further examples. After submitting the paper for review to IEEE VIS'23 and receiving feedback, Nathalie and Samuel added new examples suggested by reviewers and coded an additional set of input visualizations bringing the final corpus to 50 unique examples. This allowed us to broaden our findings and include more diverse visualization idioms and tasks.

In summary, we constructed the design space via three main processes: collecting, coding, and organizing. We focused on *collecting* examples from a diversity of sources including personal sampling and curation of design examples [57] (from blog posts, newspapers, commercial products and services, social networks, and research papers), social elicitation (discussing the topic at conferences among our peers), and reviewer suggestions. We *coded* and *organized* the design space dimensions through an iterative process, adapting or using new tools when we felt limited by existing ones.

Our resulting corpus represents a descriptive and cohesive overview of the phenomenon of input visualization, rather than a comprehensive one. While diverse, this final collection represents only a fraction of the input visualizations that exist. Instead, our corpus and design space demonstrate the existence, richness, and potential of input visualizations across a range of domains.

5 DESIGN SPACE FINDINGS

We analyzed all 50 examples (Figure 1, Table 1) with respect to their visual representation, data, artifact, context, and input, while also grouping them to highlight purposes of designs that share similar characteristics. Additional information on all examples is available both in the supplemental material (https://osf.io/bw3gp) and on the paper's browsable companion website (https://bit.ly/input-Vis). All numbered references to examples (①—㊿) in the text are clickable links to those examples' pages on the companion site.

5.1 Visual Representation

The input visualizations in our corpus are not limited to one type of representation and span many visualization idioms (such as bar charts, scatter plots, and network diagrams). Multiple examples consist of a mix of idioms, emphasizing the degree of customization possible in the design of the input visualizations.

Our corpus includes both visualizations (23/50) and physicalizations (23/50) indicating that input visualization have been created in both domains. Physicalizations labeled with P* involve drawing on paper or other surfaces (including *Observe, Collect, Draw!* ⑤, *Measuring the Universe* ⑯ Fig. 4, and *The Polish System* ㊽) and thus share some characteristics of both virtual and physical examples. *Daily Stack* ㊳ Fig. 4, *Bit Planner* ㊶, *Sandscape* ㊺ and *Affinity Lens* ㊾ are hybrid systems that use both a digital and a physical data representation.

Visualization idioms [106] describe the type of visual representation(s) used in each example. The collection contains numerous examples of bar charts (21/50), matrices (11/50), and timelines (8/50). Other idioms included heatmaps (5/50), maps (4/50), bubble charts (4/50), dot plots (3/50), scatter plots (3/50), network diagrams (2/50), affinity diagrams (1/50), and parallel coordinates plots (1/50). We also categorized several (5/50) visual representations (including *Observe, Collect, Draw!* ⑤ and *Measuring the Universe* ⑯) as custom idioms when their visual mapping did not follow existing conventions.

A number of examples incorporated combinations of multiple idioms (16/50). For instance, *Doodle* ㊷ Fig. 4 consists of a timeline visualization in combination with a matrix. Several of these combinations also involved nested idioms that incorporate one
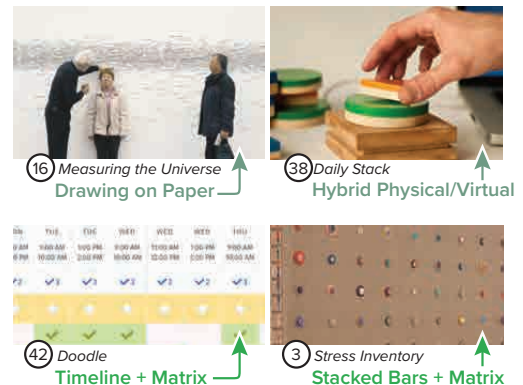

Figure 4: Examples—visual representations and idioms.

Table 1: Design space coding of our 50 examples organized by ■ Visual Representation, ■ Data, ■ Artifact, ■ Context, and ■ Input. Purposes in the first column (Individual Reflection, Public Group Reflection, Public Activity Documentation, Data Discussion, Survey, Planning, and Organizing) gather examples with similar high-level tasks, numbers of participants, synchronicity of input, data visibility at input time, and data types. A browsable index of all examples is available at: https://bit.ly/input-Vis
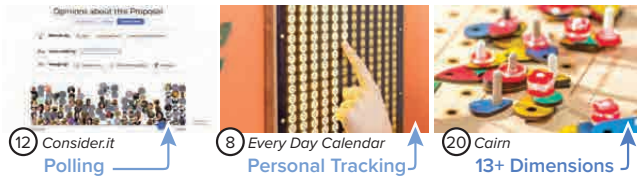
Figure 5: Examples—data semantics and dimensions.

Figure 6: Examples—artifact types and generalizability.

design within another. For example, *Stress Inventory* ③ ^{*Fig. 4*} and *Cairn* ⑳ both use stacked bar charts embedded within a matrix.

5.2 Data

We found examples of input visualization designs that support larger numbers of data dimensions (as many as 17) as well as large numbers of data points (up to hundreds of thousands), with some designs potentially supporting even more. However, most examples focused on a few dimensions and just tens or hundreds of data points. Different data types, including ordinal, categorical, quantitative, and text data, were all widespread. The input data semantics—which describe the genres of data that can be input via a visualization—were also diverse, although subjective judgments and activity tracking were especially common.

A large number of examples allowed participants to input subjective judgements (30/50) such as opinions, preferences, and feelings. These examples often involved polling (such as *Consider.It* ⑫ ^{*Fig. 5*}, *MoMA Poll* ⑭, and *Twitter Poll* ㉜) or personal tracking (such as *Stress Inventory* ③, and *Every Day Calendar* ⑧ ^{*Fig. 5*}). Other genres of data included time periods (13/50), which were often used in combination with activity (11/50) for time management purposes (as in *LEGO Time Tracking* ㊲, *Personal Apple Calendar* ㊴, and *Bit Planner* ㊶). Other categories included data about people (8/50) (like *Measuring the Universe* ⑯ and *Data Badges* ㉔) which included demographics or social relationships, and spatial data (6/50) which was often collected on maps (like *Self Reflection Physicalization* ① and *Plant Watering Tracker* ㉒).

The number of data dimensions that participants could input varied widely across our examples, with (11/50) supporting just one dimension and others supporting as many as 17. When data was input in sequence or in several rounds (as with *Polylog* ㉖ or *Kahoot* ㉟) we counted the number of dimensions possible in a single round. Examples marked with an asterisk—like the *Auto Mileage Spreadsheet* ② (17*) and *Cairn* ⑳ (13*) ^{*Fig. 5*}— could also be used to input even more dimensions. We observed no clear difference in the number or type of data dimensions between the physicalizations and visualizations.

Lastly, we examined the low level data types [106] that participants could input. The examples in our collection supported input of ordinal (38/50), categorical (35/50), quantitative (19/50), and text (18/50) data. While we coded the data attributes that could be input by participants, in some examples the input data differs from the output the system produces. For instance, timestamps are often automatically recorded and visualized but not explicitly input by the participant (as in *Tea Brewing Tracker* ㉑ and *Bubble TV* ㉘). Similarly, categorical and quantitative data is often aggregated before being visualized (as in *Visualizing Mill Road* ㉛ or *Twitter Poll* ㉜).
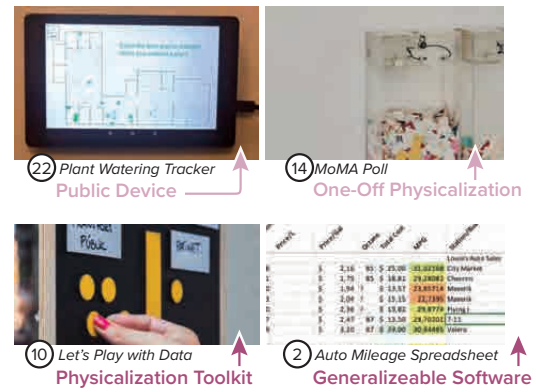
5.3 Artifact

The artifacts in our collection include visualizations on a variety of display types and physicalizations made out of a diverse set of building materials, as well as hybrid systems. While some are highly customized artifacts intended for a single purpose, a range of commercial products and toolkits also allow people to author their own input visualizations and physicalizations.

The visualizations spanned a mix of desktop devices (18/50), mobile devices (16/50), projectors (4/50), e-paper displays (2/50), and TVs (1/50). While mobile and desktop devices typically corresponded with individual use, this was not always the case. For instance, the *Plant Watering Tracker* ㉒ ^{*Fig. 6*}) was shown on a tablet, but was still used as a public display. The physicalizations, meanwhile, involved many different materials including paper, wood, plastic, gum, sand, fabric, chalk, magnets, beads, string, yarn, nails, liquid, glass, and LEGO bricks. Our collection also contains multiple instances of visualizations on paper, including examples where participants enter data by writing or drawing with pens (5/50) or by adding stickers (2/50).

Next, we coded whether each example allowed viewers or designers to create new input visualizations. One group of examples allowed *no authoring* (21/50) and were designed to be used for only one purpose (such as the *MoMA Poll* ⑭ ^{*Fig. 6*}). Another group were *parametric templates* (23/50), which allowed viewers to change the labels and dimensions (as with *Doodle* ㊷), but not the visual encoding or inputs. More general *toolkits* (6/50) provided building blocks for creating new input visualizations with custom displays and inputs. These included bespoke physicalization toolkits (like *Self Reflection Physicalization* ① and *Let's Play With Data* ⑩ ^{*Fig. 6*}) as well as widely accessible commercial tools like Microsoft Excel ② ^{*Fig. 6*}, where conditional formatting, scripting, and charting tools can be readily used to visualize live data as it is being entered. These more general tools provide opportunities for the creation of more kinds of visualizations adapted to specific uses.

5.4 Context

Our examples support various timelines of use, ranging from minutes to months (or longer). They also facilitate diverse numbers of participants, from individual users to hundreds or thousands of people. Many come from areas that are underrepresented in

Figure 7: Examples—contexts and time frames.

information visualization research, like personal tracking, time management, and civic participation.

Our collection includes (self-)tracking (17/50) and time management (6/50) tools for both individuals and groups. Other examples are designed to support civic participation (5/50), deliberation (5/50), and politics (5/50). Further clusters address education (8/50), art (6/50), and entertainment (5/50), as well as media (3/50), social networks (3/50), research (3/50), health (2/50), and urban planning (1/50).

Our examples also span both individual and group use cases, ranging from single user artifacts to visualizations intended to support thousand of participants (or more). Those with the largest numbers of participants included both physicalizations (like *MoMA Poll* ⑭ and *Measuring the Universe* ⑯, each with numbers in the thousands) and visualizations (like *The Death of a Terrorist* ⑰ with 13,864 participants or *Twitter Poll* ㉜ which can scale to millions).

Additionally, we categorized each example's input time frame—the duration over which inputs can be registered. While some instances handle inputs over just minutes (7/50) or hours (15/50), many can continue to accept data over the course of days (7/50), weeks (10/50), or months (9/50). Many of the examples with shorter time frames (like *Dot Voting* ㉕ and *Polylog* ㉖ *Fig. 7*) are intended to be used during rapid activities or live events. Those with timeframes in the range of weeks (like *The Death of a Terrorist* ⑰ and *Visualizing Mill Road* ㉛ *Fig. 7*) or months (like the *MoMA Poll* ⑭ and *The Happy Show* ⑮) were often associated with ongoing events like exhibitions or related to ongoing news. Other examples (especially planning and time management tools like *LEGO Time Tracking* ㊲, *Daily Stack* ㊳, and *Planning Fiche T* ㊵ *Fig. 7*) were typically designed to work indefinitely (8/50).

5.5 Input

We identified examples that supported a variety of high- and low-level tasks, but observed a tendency towards designs that collect data via additive interactions. Input modalities varied more widely, although physicalizations typically relied on embodied interactions with data elements while visualizations incorporated a wider range of mediated and indirect input approaches.

We categorized all the examples into three non-exclusive high-level tasks (collecting, sorting, and planning) that describe the primary goal of the visualizations. Collecting (36/50) involves input visualizations that are designed to primarily track and collect new data points (as in *I/O Bits Streak Tracker* ⑦ or *Feedback Frames* ㉝). Planning (7/50) involves scheduling future events by inputting data about upcoming activities (as in *Personal Apple Calendar* ㊴ or
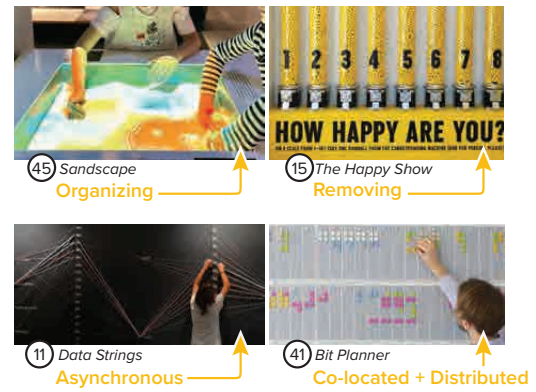

Figure 8: Examples—input tasks and activities.

Bit Planner ㊶). Organizing (8/50) involves categorizing or sorting existing content instead of collecting new data (as in *Sandscape* ㊺ *Fig. 8* or *Alignment Chart Maker* ㊼).

We also identified possible low-level input data actions [31, 74] that can be performed with each input visualization, including adding, modifying, and removing data records. We coded all of the actions that are explicitly supported by each tool. However, for many (particularly the physicalizations) additional operations may also be possible. All but one of the examples allow adding—the exception being *Sandscape* ㊺, which relies on modifying and re-configuring existing sand. A group of examples (19/50) exclusively support adding, while most of the remainder (25/50) support a variety of interactions including adding, modifying, and removing. That said, the unique designs of individual visualizations can substantially change the character of a given interaction—as seen in *The Happy Show* ⑮ *Fig. 8* where participants register new votes by subtractively *removing* gumballs from pre-filled tubes.

We also analyzed the locality and synchronization of input for visualizations that support multiple participants (but not for those intended for a single user). Synchronous data input (21/50) (in examples like *Kahoot* ㉟ or *Planning Fiche T* ㊵) allows multiple participants to input data and see results simultaneously. This contrasts with asynchronous input (13/50) in which participants' inputs are added in larger discrete chunks (*Doodle* ㊷) or where multiple simultaneous inputs may not be possible (*Data Strings* ⑪ *Fig. 8*).

Co-located data input (26/50) occurs at a shared physical location (as in *Participatory Matrix* ㉙ and *Edo* ㉓) whereas distributed data input (11/50) can occur across many different locations (as in *The Death of a Terrorist* ⑰ or *Twitter Poll* ㉜). Some examples, like the hybrid *Bit Planner* ㊶ system *Fig. 8* and the *Polemic Tweet* ㉗ platform (which supports both remote and in-person conference participants) can be used in both co-located and distributed modes.

Finally, most of our examples (44/50) keep any existing data visible during data input. However, a small number (6/50) of tools (like *Visualizing Mill Road* ㉛ and *Twitter Poll* ㉜) do not reveal any data until after a participant has successfully entered new data—typically to keep that information from influencing their responses.

5.5.1 Input Modalities. We identified six high-level input modalities used across the set of visualizations in our corpus. These modalities correspond to different high-level modes of interaction that lend themselves to different visual representations, types of data,
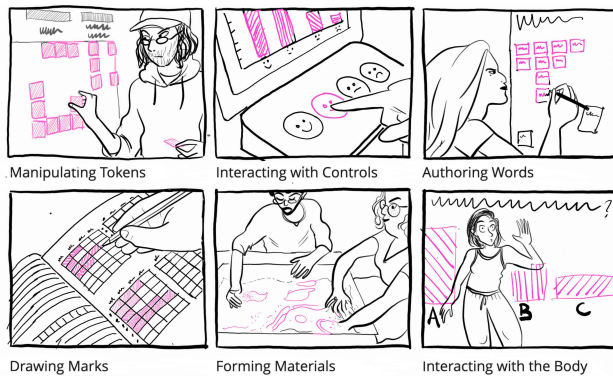
Figure 9: The six input modalities, with input data represented in pink. Virtual and physical variants of all input modalities are possible. CC BY figure at https://osf.io/bw3gp.

and input strategies. While the boundaries between these modalities are permeable (for example, versions of all of these modalities could be created using tokens), they highlight design opportunities as well as areas for future research.

Manipulating Tokens (24/50). These visualizations use tokens [69] each of which usually corresponds 1-to-1 with a row in the dataset (making most of them unit visualizations). Tokens can be digital (as in *Tier List Maker* ㊻) or physical (as in *LEGO Time Tracking* ㊲). Interactions with tokens are often additive (as in most of the voting systems in our collection), but can also be subtractive (as in examples like *The Happy Show* ⑮ where the tokens are taken away by participants as a souvenir), and can rely either on tokens created by the participants (as in *Stress Inventory* ③ where the token creation is part of the process) or tokens defined by the visualization's creators. Placing tokens provides a mechanism for inputting data using semantically-defined layouts, similar to token+constraint [136] systems in tangible interaction—where physical objects that represent digital information are placed in confining regions. Axis-based layouts involve positioning tokens on an axis to construct visualizations like bar charts (like *LEGO Time Tracking* ㊲ and *Daily Stack* ㊳) or scatter plots (like *Alignment Chart Maker* ㊼). Other token-based systems include matrix-based (like *Planning Fiche T* ㊵ and *Bit Planner* ㊶) or pile-based (like *Dot Voting* ㉕ or *Polylog* ㉖) variants where tokens are placed into piles that represent different categories. Stacked or linked tokens (like those in *LEGO Time Tracking* ㊲) are often more readily countable, while token systems that rely on containers that are filled up or emptied with tokens to create a visual representation (like *MoMA Poll* ⑭ and *The Happy Show* ⑮) are more difficult to count.

Interacting with Controls (20/50). Another large class of input visualizations in our sample support input by manipulating interface controls—typically either physical or digital instantiations of standard WIMP interface widgets like buttons, sliders, or menus. These include desktop and mobile applications that support voting and other inputs via virtual sliders (*Consider.It* ⑫) or buttons (*Bubble TV* ㉘ and *Kahoot* ㉟). Physical installations like *Visualizing Mill Road* ㉛, meanwhile, use tactile buttons and other inputs to support input in civic spaces. One common input pattern often

used in polling and rating systems (including *Polylog* ㉖, *Visualizing Mill Road* ㉛, *Twitter Poll* ㉜, and *Drip By Tweet* ⑲) involved selecting inputs from a set of pre-defined options displayed near the visualization—but without touching the visual marks directly.

Authoring Words (16/50). Many of the examples we examined allowed people to input text data, either by typing or writing. While text input is quite expressive, we found few examples that relied on it exclusively—probably because free-text input is less likely to result in structured tabular data that is readily visualized. Instead, examples in our collection typically relied on text entry for labeling data points (as with the *Tea Brewing Tracker* ㉑ or *Wedding Planner* ㊸) or introducing comments alongside other more structured quantitative data inputs (as in *Polemic Tweet* ㉗).

Drawing Marks (6/50). Another set of visualizations used drawing tools to support the creation of new expressive shapes and marks. These include visualization templates like those in *Observe, Collect, Draw!* ⑤ which encourage data collection via both structured and free-form paper sketching. This approach is often used to encode expressive, uncertain, or qualitative data using artistic media. Many of these examples use a "coloring book" metaphor common in bullet journaling [5], where viewers input data by adding color either manually (as in *Observe, Collect, Draw!* ⑤ and *The Polish System* ㊽) or digitally (as in *Trackly* ⑥) to predefined shapes. Other examples like the *Plant Watering Tracker* ㉒ rely on more limited forms of drawing, allowing people to make simple marks on different image substrates.

Forming Materials (4/50). Some examples also extend the notion of drawing into 3D space, using physical materials like sand (*Sandscape* ㊺) and string (*Knitting City Council* ④ and *Data Strings* ⑪) to support data input. In our corpus, all of the examples of formative input are physical. However, these manipulations are also possible in virtual spaces, where they might sit alongside existing 3D modeling approaches and input techniques for scientific and medical visualization.

Interacting with the Body (1/50). Finally, we encountered a single example (*MyPosition* ⑬) which takes the body position of viewers as input, allowing viewers to register votes by standing in front of a part of the visualization and changing their pose. While relatively unexamined thus far, the huge expressive range of human poses and movement (including hand, face, and body gestures) suggests this area is ripe for future experimentation.

5.6 Purposes of Input Visualizations

We identified seven purposes for input visualizations (Figure 10) based on common characteristics we observed across our examples. These purposes highlight the various objectives and scenarios that the input visualizations in our collection have been used for: Individual Reflection, Public Group Reflection, Public Activity Documentation, Data Discussion, Survey, Planning, and Organizing. We identified these purposes by clustering the examples based on 1) the high-level task they support, 2) the number of participants, 3) whether they make existing data visible during input, 4) if they operate synchronously or asynchronously, and 5) the types of data they incorporate. (For additional details see our supplemental material.)
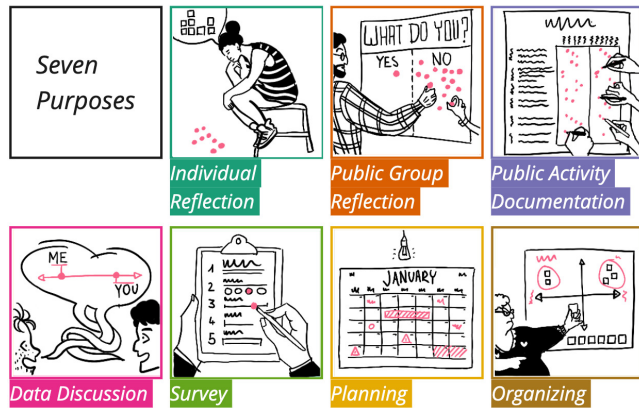
Figure 10: Seven Purposes of input visualizations: ■ Individual Reflection, ■ Public Group Reflection, ■ Public Activity Documentation, ■ Data Discussion, ■ Survey, ■ Planning, and ■ Organizing. CC BY figure at https://osf.io/bw3gp.

Individual Reflection focuses on personal data collection by an individual to support self-knowledge. These include tools for capturing subjective judgments (*Trackly* ⑥), activities (*I/O Bits Streak Tracker* ⑦), and data about people (*Observe, Collect, Draw!* ⑤).

Public Group Reflection examples like *Consider.It* ⑫ and *The Death of a Terrorist* ⑰ allow multiple people to collectively reflect on a topic. The data collection process is asynchronous and focuses on gathering subjective judgements like sentiments, opinions, or shared observations.

Public Activity Documentation examples like *Cairn* ⑳ or the *Plant Watering Tracker* ㉒ document ongoing group activities. The data collection process is synchronous and focuses on accumulating time and activity data.

Data Discussion examples such as *Dot Voting* ㉕ and *Polemic Tweet* ㉗ operate synchronously and allow ongoing collective discussions mediated by the input visualization.

Survey examples including *Visualizing Mill Road* ㉛ and *Feedback Frames* ㉝ aim to collect data from a group of people. The existing data is hidden during input and only shown afterwards, reducing the likelihood that prior entries will influence the data currently being collected.

Planning examples (including individual examples like *LEGO Time Tracking* ㊲ or *Personal Apple Calendar* ㊴ and group tools like *Bit Planner* ㊶ or *Doodle* ㊷) often bring an explicit focus on time and logistics—helping people make sense of ongoing time use, develop plans, and forecast future events.

Organizing examples focus on sorting and categorizing, often allowing viewers to change data points, attributes, or even data dimensions. For example, *Tier List Maker* ㊻ and *Alignment Chart Maker* ㊼ provide frameworks for sorting and ranking arbitrary sets of images or concepts, while *You Name It* ㊸ focuses on identifying design patterns by sorting cards generated from the data physicalization list [37].

While these purposes capture common themes and characteristics shared by sets of examples in our corpus, they are not mutually exclusive. In fact, individual artifacts might still share aspects of more than one purpose or shift between purposes over time depending on how they are used. For example, the *Tea Brewing Tracker* ㉑—which was used both individually and by a couple—illustrates a continuum between Individual Reflection and Public Activity Documentation. Similarly, while city councillor Sue Montgomery created *Knitting City Council* ④ as an Individual Reflection piece, it documents the collective activity of a larger group and was posted later publicly on social media, encouraging broader discussion and Public Group Reflection.

6 CONSIDERATIONS WHEN DESIGNING INPUT VISUALIZATIONS

Our design space and reflections on our corpus of examples suggest a variety of design considerations that creators of new input visualizations may need to consider, and that researchers may wish to further investigate.

6.1 How to Deal With the Dynamic Nature of Input Visualizations?

All input visualizations are dynamic by default and need both an assembly model [69] and an approach to deal with dynamic data. This means that input visualizations can suffer from the same kinds of scaling pressures faced by other visualizations of dynamic or real-time data [26, 102]. The exact quantity, scope, and nature of the data that will be collected with an input visualization is often unknown at the time it is created. Both the **number of participants** and the **input time frame** (Sec. 5.4) can influence the amount of data that will be collected, with larger amounts of data causing cluttering, reducing readability, or exhausting the availability of tokens and other input elements. A variety of strategies can be called upon to deal with these challenges, including using dynamic visual mappings, aggregating or binning data, limiting the amount of data a visualization can show, or introducing explicit maintenance or cleanup mechanisms.

Binning and aggregation approaches involve summarizing and potentially separating the data into intervals. For example, *The Death of a Terrorist* ⑰ uses discrete bins in a grid to aggregate inputs, allowing the visualization to scale to many thousands of responses. Similarly, visualizations like *Twitter Poll* ㉜ and *Google Star Rating* ⑨ can aggregate data and scale their visual marks to accommodate practically unlimited numbers of inputs.

However, when it is important to see individual contributions, an aggregation approach may not be suitable. This can be the case for group Planning systems that require the visibility of individuals' availabilities, as well as for Data Discussion and Public Group Reflection tools that aim to foreground individual participants' opinions. Some input visualizations solve this problem by using dynamic visual mappings or adaptive scales that show a limited amount of data to reduce clutter while preserving detail. For instance, marks in the *Plant Watering Tracker* ㉒ (Public Activity Documentation) gradually fade and disappear, while tokens in *Bubble TV* ㉘ (Data Discussion) shrink over time using a metaphor of visual sedimentation [71]. The subtractive approach used in *The Happy Show* ⑮ (Public Group Reflection) has a similar effect, with participants removing physical tokens from the bars and reducing visual cluttering over time.

Scalability concerns are perhaps biggest for **physicalizations** (Sec. 5.1), where common **visualization** approaches like increasing the display size and resizing marks can be much more cumbersome. Physicalizations like *Mindworks* ⑱, *MoMA Poll* ⑭, *The Happy Show* ⑮, and *Cairn* ⑳ require routine maintenance to deal with dynamic data, including physical restocking of input materials and manual removal of elements when the view becomes too cluttered. Scaling issues can also take the reverse form when input visualizations collect less data than expected. For example in *Edo* ㉓, Sauvé et al. [123] reported that because the canvas of their physicalization was *larger* than the amount of data that was ultimately input, the resulting visuals were less striking.

Consideration: Keeping track of the data while reducing clutter can be a challenge, particularly for Data Discussion, Public Activity Documentation, and Public Group Reflection visualizations where data quantities and distributions may be hard to anticipate. Designers creating input visualizations with these properties need to think critically about both the cost and usability of their designs across a range of usage scenarios and consider strategies for dealing with too much data. Potential strategies include dynamic visual mapping, aggregation, binning, limiting the data shown, and incorporating explicit maintenance. Designers creating input visualizations need to weigh factors like the anticipated amount of data (**number of participants**/**input time frame**), the target **technology/material**, **visual representation**, and **purpose**.

6.2 How do Design Choices Influence the Data People will Input?

The experience of entering data into a visualization is heavily shaped by initial design choices, including the supported **input modalities**, whether **existing data is visible during input**, and the overall **visual representation**.

For example, whether or not **existing data is visible during input** (Sec. 5.5) can have an impact on subsequent data collection. Survey-style designs, where prior data is only revealed after an input is completed, reduce the possibility that existing data will influence the new values. These approaches have the potential to help mitigate bias, including anchoring [48] and conformity [3] effects, in both individual and group settings. Depending on the **input time frame** (Sec. 5.4) these reveals can happen immediately after the input (*Twitter Poll* ㉜), after a voting period (*Kahoot* ㉟), or after a longer data collection period—as in *Visualizing Mill Road* ㉛, where the results were painted onto the sidewalk the following day. Hiding data during input may be desirable in cases where bias due to the existing data is a source of concern, while making it visible may benefit social or reflective applications. The material constraints of physicalizations often make hiding data more challenging—however approaches like *Feedback Frames* ㉝ highlight the potential of simple hide-and-reveal interactions.

On the other hand, Public Group Reflection designs often explicitly choose to surface prior data up front to provide context and create a sense of social engagement or to support reasoning based on prior responses. However, this means that each data input reflects a different state of the visualization, and the values are not comparable in the same way. Instead, these visualizations tend to serve as collective social artifacts which evolve over time.

Similarly, the choice of **input modality** (Sec. 5.5.1) can influence the character of the data collected. Previous work on survey design has investigated how the visual design and layout of surveys can influence participants' answers [128]. For example, the presence of tick marks in sliders and visual analogue scales impacts the responses that people give [97]. Likewise, the data schema of an input visualization—including data attributes, category groups, scales, and bounds—will almost certainly prime participants, impacting the data that is collected. Setting these can be challenging, however, as the real distributions of values, outliers, and in some cases even data dimensions might be unknown until after the visualization has been deployed.

Consideration: The character of the data collected via an input visualization will differ depending on various aspects of its design. Identifying the **purpose** (Sec. 5.6) the input visualization will serve can be helpful when making decisions about data visibility, input modalities, and the intended character of the data to be collected.

6.3 How to Balance Readability and Freedom of Input?

An input visualization's choice of **visual representations** (Sec. 5.1) and **input modalities** (Sec. 5.5) can enforce constraints and provide viewers with different degrees of freedom to change the visual representation. Flexible approaches can allow greater expressivity—as in our *Dot Voting* ㉕ example (Public Group Reflection) where participants rate sustainable development goals by positioning dots on a matrix. Here, the open canvas allows participants to express enthusiasm, create groupings, and even split votes across categories, but makes extracting and comparing category counts more challenging. Other systems, such as *Feedback Frames* ㉝ constrain the position of the tokens, resulting in more countable and comparable results, increasing readability of the visual representation.

We find that Organizing input visualizations tend to provide more freedom to change the visual representation and influence the form of the data—often by allowing viewers to not only add new data points but also new data dimensions or attributes. For instance, the *Wedding Planner* ㊸ example allows participants to not only add or remove guests, but also add new dimensions in the dataset (for instance, by drawing new tables on the canvas or adding social interests to the tokens representing guests). Similarly, *Affinity Lens* ㊾ provides a great degree of freedom when organizing and grouping cards. Survey artifacts (like *Twitter Poll* ㉜, *Visualizing Mill Road* ㉛, or *Citizen Dialogue Kit* ㊱), on the other hand, tend to be more restricted—ensuring comparable results by asking viewers to choose from sets of pre-defined options.

While different degrees of freedom when inputting data are possible with all **input modalities** (Sec. 5.5), some approaches lend themselves more to either free or restricted input. In token-based visualizations, designers can set constraints on the input data by defining properties of the tokens (as in *Cairn* ⑳) or their position (as in *Participatory Matrix* ㉙). Similarly, incorporating interface controls like buttons, sliders, or menus can limit inputs to a predefined set of options (as in *Twitter Poll* ㉜) leading to more uniform datasets. On the other hand, authoring words, drawing marks, and forming materials can facilitate more freeform and expressive input (as in *Observe, Collect, Draw!* ⑤). Combinations of input modalities

can also be used to leverage the strengths of multiple input types. For instance, the *Tea Brewing Tracker* ㉑ combines interacting with controls to input time stamps for a timeline and authoring words to describe each of the events. Similarly, token-based systems that allow participants to draw on their tokens can systematize some input attributes while leaving participants free to define others.

Consideration: Constraining input mechanisms and visual encodings can improve the readability of the visual representation and the consistency of the data, but also limits participants' freedom to input more complex and expressive values. While having more freedom might increase the degree of expression at the data level, it also leaves space for interpretation or uncertainty.

6.4 How to Support Multiple Inputs?

When examining visualizations that allow individuals to input multiple data values, we observed two main strategies. Systems like *Polylog* ㉖ and *Kahoot* ㉟ collect responses sequentially by incorporating a sequence of questions or prompts that can be answered one-by-one to input data. Meanwhile, examples like *Cairn* ⑳ or *Alignment Chart Maker* ㊼ take a spatial approach, assigning each response to a new token or spatial position, making it possible for people to input multiple data values that are distributed across a space. While sequential approaches are typically realized through interacting with controls, spatial approaches more often utilize tokens as an **input modality** (Sec. 5.5.1). Interestingly, some physical examples (like *Data Strings* ⑪) also combine these patterns by leveraging material properties like string tension to create spatial layouts that need to be completed in a particular sequence.

Consideration: Spatial strategies can create opportunities for more complex representational systems and input mechanisms that let participants integrate multiple observations (for example, creating a cairn by assembling multiple tokens or weaving a string in a parallel coordinates plot). Meanwhile, sequential approaches can reduce cognitive load and provide a more directed experience.

6.5 How Hard is it to Get the Data?

Depending on an input visualization's design, acquiring a tabular or structured record of the data from an visualization can sometimes entail considerable additional effort—particularly for **physicalizations** (Sec. 5.1). Most physical examples we considered still required manual measurement or data transcription, and we counted only three hybrid systems (*Daily Stack* ㊳, *Bit Planner* ㊶, and *Affinity Lens* ㊾) capable of closing this loop in an automated way. When designing an input visualization, a designer needs to consider if having a structured record of the data is important. In some cases, especially for Public Group Reflection or Data Discussion, the goal of the input visualization might be primarily to initiate discussion rather than to obtain a data record or to enable further data processing. An example like *MoMA Poll* ⑭, for instance, relies on the visual impact of filling up a space with tokens, but counting the exact number of inputs is less important.

Consideration: Exporting data from input visualizations for processing, storage, or visualization in other tools is still typically easier for digital visualizations and can entail considerable effort for physicalizations or any design with particularly free-form input.

7 DISCUSSION

Input visualization is not a new phenomenon but a recognition of a set of existing patterns both within and adjacent to what we typically think of as data visualization. Yet considering these examples as visualizations raises deeper questions about the nature of data and the ways in which our community has traditionally delineated its boundaries, while suggesting opportunities for research and new kinds of data-driven thinking tools. With this in mind, we reflect on the nature and value of these representation and propose a set of future research opportunities (RO1-RO8) that highlight ways the visualization research community might embrace and expand upon the potential of input visualization.

7.1 Considering the Relationship between Visualization and Input Visualization

We found it interesting that almost all of the examples of contemporary input visualizations we identified come from outside of the information visualization community. This likely helps explain why little research has investigated the approach. However, the practical utility of input visualizations for a wide range of polling, planning, and thinking tasks suggests that they merit further study by visualization researchers. Alternatively, one could argue that while many of our examples may look like visualizations, they are not. After all, most of these input visualizations do not align neatly with garden-variety descriptions of visualization, which usually explain visualization approaches as "visually encoding data to make them easier to understand".

Some of the visualization research community's most beloved definitions of information visualization incorporate enough generality to include input visualization approaches, but others leave less room for them. For example, Card et al.'s definition of visualization as "the use of interactive visual representations of data to amplify cognition." [18], says nothing about where the data comes from and places the most weight on the more abstract goal of amplifying cognition. On the other hand, Keim et al.'s description of information visualization as "the communication of abstract data relevant in terms of action through the use of interactive visual interfaces" [78] places an emphasis on *communication* that seems to exclude most input visualizations. Data collection also does not fit neatly into any of the three major goals of visualization (presentation, confirmatory analysis, and exploratory analysis) that Keim et al. identify.

One might also take the position that many of these examples *are* visualizations, but that they are trivial ones—and that considering spreadsheets, calendars, or token voting systems as visualization tools is reductionist or simply not useful. Yet such an assertion runs counter to a variety of recent work highlighting the relevance of spreadsheets [8], bullet journals [5], and other representations *as* visualization tools. Moreover, doing so runs the risk of drawing a boundary around visualization that excludes most of the pre-industrial history of visualization, as well as a considerable slice of contemporary work on physicalization [38], infographics [17], and other visualization-adjacent topics.

Similarly, what constitutes an *input* visualization is largely a matter of perspective and context. For example, simple visualizations like a scented slider widget [141] with bars encoding the values
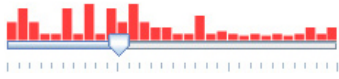
Figure 11: A scented slider showing values entered by prior users. (Reproduced from Willett et al. [141])

selected by prior users (Figure 11) could function as an input visualization if used as a collective voting or preference elicitation tool. However, when used as a passive indication of past activity in the context of a larger task, an input visualization framing may be less relevant. As a result, it is challenging to delineate the set of input visualizations and non-input visualizations based purely on their appearance or interactivity, as these distinctions depend primarily on the visualizations' context of use. However, the decision to use a visualization for data input brings a variety of unique design considerations—including questions about the influence of design choices on the input data, repeated inputs, and data access (as detailed in Sec. 6)—that have thus far been largely unconsidered in the visualization literature.

Part of the challenge of characterizing input visualizations arises from their position at the intersection of different research domains. On the one hand, input devices have classically been studied in HCI [10, 20], while visual structures and visual mapping that represent data are more an information visualization topic [56, 106, 138] while physicalization bridges these domains as well as others, like design and tangible interaction [6, 42, 64]. This intersection suggests a rich set of opportunities for further research including:

> **RO1** - *Understanding how the knowledge produced by the information visualization community to represent existing data is applicable to input visualization.*
>
> **RO2** - *Expanding our knowledge of the interactions between new and existing visualization types, input interactions, and the complexity of the collected data.*
>
> **RO3** - *Building a more cohesive and cross-disciplinary understanding of different facets of input visualization including input modalities, interactions, visual and physical mappings, and data representations.*

7.2 Input Visualizations and the Meaning(s) of Data

When do we consider data to *be* data? Does it need be recorded and encoded in a digital file or tabulated in a structured format? While some examples of input visualizations do indeed produce structured and easily-interpretable data, many others—including systems that rely on unstructured input, physical materials, or ambiguous encodings—may not. As Jansen and Dragicevic note, the lack of any underlying data structure can mean that the data are manifest *only* in the visual artifact [74], which may or may not be easily measurable or reproducible. A lack of formalized data schemas or visual encodings can also mean that critical aspects of the data may exist only in the relationships between visual marks or in other intangible aspects of the visual representation and thus resist precise quantification.

For example, the two axes used in the case study *The Death of a Terrorist* ⑰ (negative↔positive and not-significant↔significant)

lack absolute values or landmarks, and the significance of individual points is largely implied by their relationship to those around them. Physical installations like *Let's Play With Data* ⑩ or *Cairn* ⑳ introduce further ambiguity. For example, how should we interpret a mark that intersects both "yes" and "no"? What information, if any, does an elaborate and intentional token stack (as in Figure 2) communicate if ordering is not formalized in the instructions? Converting this kind of ambiguous and contextual data into a structured or tabular format may be challenging or even impossible without information loss. For example, Thudt et al. reported (when reflecting upon *Self Reflection Physicalization* ①) that "[transforming] an experience directly into a visual and physical manifestation makes it more difficult to create alternate representations later on" [132].

Input visualizations, as mechanisms for collecting and displaying new information, also collide with deeper epistemological discussions about the nature of "data" itself. Already, humanities researchers such as Latour [86] and Drucker [39] have criticized the implications of the term *data*, whose very etymology—from the Latin *datum* "(thing) given"—implies that information is somehow objective in nature, and obscures the myriad biases, errors, and sources of uncertainty intrinsic to any attempt to observe or record external phenomena. Drucker advocates instead for the notion that all data is in fact "*capta*", which is actively "taken" from the world and reflects the unique tools, approaches, and biases implicit in each mode of knowledge production or inquiry. The notion of capta and considerations of the constructivist nature of data collection and visualization production are already important veins of discussion within the visualization community. However, they become even more salient in the context of input visualizations, which more explicitly surface the mechanisms of data collection.

Venturing even further down the epistemological rabbit hole, some pre-digital definitions of data, including from the Diderot and d'Alembert's 18th-century *Encyclopedia* [29], further differentiate data which are given (*data*) from "those which are unknown, and which one seeks" (*quæsita*). Given this perspective, one could argue that the information captured in an input visualization are only data (or capta) *after* they have been input. Up until that point, including during the design of the visual representation, these future pieces of information remain quæsita—*sought* but not yet obtained. In Sec. 6.2, we discussed how the visibility of previous data can lead to well-known social or cognitive biases, such as conformity, polarization, and anchoring, all of which can directly influence the collected data. However, visibility is just one facet of how the design of the input mechanism and process may affect the properties and qualities of the data. In reality, several aspects of input visualization design can influence group dynamics and the data collected. This suggests several research opportunities:

> **RO4** - *Developing new approaches for reducing information loss and supporting data retrieval from a variety of input visualizations, including physicalizations.*
>
> **RO5** - *Understanding and formalizing the different qualities of data collection processes and characterizing the relationships between different input visualization designs and the qualities of the data they produce.*
>
> **RO6** - *Developing new input mechanisms to mitigate social and cognitive biases during public data collection.*

7.3 New Application Areas: Input Visualizations as Sensemaking Tools

Moving forward, thinking of visualizations as input spaces opens up possibilities not only for data collection, but also for cleaning, synthesis, and other sensemaking tasks. After all, data wrangling—which constitutes a huge part of the analysis process—*almost always* involves changing data [76], and visual feedback in data entry interfaces has long been viewed as an important mechanism for improving input data quality [62]. Modern data cleaning and wrangling tools like Wrangler [77] and Tableau Prep [121], already incorporate visualizations alongside tabular representations of data to support data cleaning and transformation. Similarly, visualization approaches like conditional formatting and sparklines can be invaluable tools for identifying issues when inputting and updating data in spreadsheets [8] In fact, despite receiving little attention in visualization research, visually-augmented spreadsheets are very likely the most ubiquitous input visualizations in use today and represent a promising area for future research in their own right.

Thinking of visualizations as input spaces creates opportunities for bringing subjective observations, context, and expectations into the analysis process in more formal and operationalizable ways—building on recent concepts like implicit error [99] or data hunches [89]. This suggests untapped opportunities for treating input visualizations as a core component of data analysis and sensemaking cycles, particularly for tasks like thematic analysis or strategic planning where observations and decisions often depend on qualitative judgments. Consider our workflow for this paper, which contained many iterative rounds of data collection, coding, clustering, modeling, and refinement spread across a variety of virtual and physical platforms (and incorporating multiple encodings of the same data in both tablular and associative formats). Sensemaking workflows like these are suffused with subjective and collective decision-making tasks, which call for the ability to adjust, group, split, and modify data, as well as create new representations. There is also potential for input visualization in deliberation and civic participation contexts which involve decision-making processes between groups of people, as highlighted by Dimara et al. [32].

Yet current tools make it challenging to connect data entry (which typically takes place in structured spreadsheets or databases) and analysis (in output-only visualization tools) with more subjective exploratory approaches like virtual or physical card-sorting. New input visualization systems for externalized sensemaking could close this gap by supporting more fluid transitions between interactive input visualizations and structured tabular data—permitting people to interactively modify data and schemas throughout their analytic workflows by interacting with visualizations:

> **RO7 -** *Developing new tools that integrate input visualizations into analytic workflows to support richer data-informed sensemaking, decision-making, and thinking.*

7.4 Design Methods for Input Visualization

Numerous design approaches exist in information visualization research, including design studies [124], the nested model [105], data-first approaches [110], action design research [98], design by immersion [59], and user-centered strategies [52, 82, 92]. These

approaches primarily assume that the data is known in advance. However, when designing an input visualization, the data, visual representation, and input mechanisms are interdependent and thus must be defined simultaneously. This implies that when creating an input visualization, designers may confront multiple challenges including 1) defining the data abstraction, 2) devising the input mechanism, and 3) creating a visual representation that supports both analytical purposes and data inputs. These challenges underscore the limitations of applying current design approaches to input visualization.

Further research can investigate design methods tailored to the distinct challenges presented by input visualizations. The design processes for examples in our corpus (where documentation exists) encompass a range of methods, such as autobiographical design and co-design (*PlantWatering*, *TeaBrewing* [16], and *SelfReflection* [132]), user-centered design (*LetsPlaywithData* [40]), and author-centered design (*DataStrings* [27]). Other fields, such as personal informatics and self-tracking [65, 88, 117] for personal data collection, decision support systems [21], and crowdsourcing for collectively structured data collection [25, 33], offer insights into data collection methods that can complement existing methods and inform the design of input visualizations. This highlights an additional opportunity:

> **RO8 -** *Developing design methods for input visualization that take into account the unique challenges they pose.*

8 CONCLUSION

For now, input visualizations remain a niche and underconsidered corner of the visualization universe, but one that we suspect is full of untapped potential. Our work represents just a first step towards mapping the broader design space of input visualizations by investigating their visual representations, data, artifacts, contexts, input techniques, and purposes. Based on our initial investigation, we introduce a set of design considerations and highlight new research opportunities for examining the relationship between input visualization and visualization, exploring the use of input visualizations as sensemaking tools, and developing new design methods for input visualization. With this in mind, we encourage the visualization community to further examine this space—building an understanding of the potential of this approach, one input at a time.

REFERENCES

[1] Robert Amar, James Eagan, and John Stasko. 2005. Low-level components of analytic activity in information visualization. In *IEEE Symposium on Information Visualization, 2005. INFOVIS 2005*. IEEE, 111–117.

[2] Amazon. 2023. Nobo 7 Panel Weekly Business T Card Organiser Midi Kit, 480 x 480 mm - Assorted. https://www.amazon.com/Nobo-Planner-colonnes-supports-aluminium/dp/B000SHVX4M/r. Last accessed 2023-03-08.

[3] Solomon E Asch. 1951. Effects of group pressure upon the modification and distortion of judgments. *Groups, Leadership, and Men* (1951), 177–190.

[4] Amid Ayobi, Paul Marshall, and Anna L Cox. 2020. Trackly: A Customisable and Pictorial Self-Tracking App to Support Agency in Multiple Sclerosis Self-Care. In *Proc. of the 2020 ACM CHI*. 1–15.

[5] Amid Ayobi, Tobias Sonne, Paul Marshall, and Anna L Cox. 2015. Flexible and mindful self-tracking: Design implications from paper bullet journals. In *Proc. ACM CHI*.

[6] S Sandra Bae, Clement Zheng, Mary Etta West, Ellen Yi-Luen Do, Samuel Huron, and Danielle Albers Szafir. 2022. Making data tangible: A cross-disciplinary design space for data physicalization. In *Proc. of ACM CHI*. 1–18.

[7] Michael J. Baker, Matthieu Quignard, Kristine Lund, and Arnauld Séjourné. 2003. *Computer-Supported Collaborative Learning in the Space of Debate*. Springer Netherlands, Dordrecht, 11–20. https://doi.org/10.1007/978-94-017-0195-2_4

[8] Lyn Bartram, Michael Correll, and Melanie Tory. 2021. Untidy Data: The Unreasonable Effectiveness of Tables. *IEEE TVCG* (2021).

[9] Thomas Baudel. 2006. From information visualization to direct manipulation: extending a generic visualization framework for the interactive editing of large datasets. In *Proc. ACM UIST*. 67–76.

[10] Michel Beaudouin-Lafon. 2000. Instrumental interaction: an interaction model for designing post-WIMP user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 446–453.

[11] Michel Beaudouin-Lafon. 2004. Designing interaction, not interfaces. In *Proceedings of the working conference on Advanced visual interfaces*. 15–22.

[12] Michel Beaudouin-Lafon and Wendy E Mackay. 2007. Prototyping tools and techniques. In *The human-computer interaction handbook*. CRC Press, 1043–1066.

[13] Anaelle Beignon. 2013. You Name It! http://www.anaellebeignon.fr/you-name-it.html. Last accessed 2023-03-08.

[14] Michael Mose Biskjaer, Peter Dalsgaard, and Kim Halskov. 2014. A constraint-based understanding of design spaces. In *Proceedings of the 2014 conference on Designing interactive systems*. 453–462.

[15] Matthew Brehmer and Tamara Munzner. 2013. A multi-level typology of abstract visualization tasks. *IEEE TVCG* 19, 12 (2013), 2376–2385.

[16] Nathalie Bressa, Jo Vermeulen, and Wesley Willett. 2022. Data Every Day: Designing and Living with Personal Situated Visualizations. In *Proc. of the 2022 ACM CHI*. 1–18.

[17] Lydia Byrne, Daniel Angus, and Janet Wiles. 2015. Acquired codes of meaning in data visualization and infographics: Beyond perceptual primitives. *IEEE TVCG* 22, 1 (2015), 509–518.

[18] Stuart Card, Jock Mackinlay, and Ben Shneiderman. 1999. *Readings in information visualization: using vision to think*. Morgan Kaufmann.

[19] Stuart K Card and Jock Mackinlay. 1997. The structure of the information visualization design space. In *Proceedings of VIZ'97: Visualization Conference, Information Visualization Symposium and Parallel Rendering Symposium*. IEEE, 92–99.

[20] Stuart K Card, Jock D Mackinlay, and George G Robertson. 1990. The design space of input devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 117–124.

[21] Eric D Carlson. 1978. An approach for designing decision support systems. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems* 10, 3 (1978), 3–15.

[22] Ed Chi and John T Riedl. 1998. An operator interaction framework for visualization systems. In *Proc. of IEEE InfoVis*. IEEE, 63–70.

[23] John Joon Young Chung, Hijung Valentina Shin, Haijun Xia, Li-yi Wei, and Rubaiat Habib Kazi. 2021. Beyond show of hands: Engaging viewers via expressive and scalable visual communication in live streaming. In *Proc. of the 2021 ACM CHI*. 1–14.

[24] Jorgos Coenen, Eslam Nofal, and Andrew Vande Moere. 2019. How the Arrangement of Content and Location Impact the Use of Multiple Distributed Public Displays. In *Proceedings of the 2019 on Designing Interactive Systems Conference* (San Diego, CA, USA) (*DIS '19*). Association for Computing Machinery, New York, NY, USA, 1415–1426. https://doi.org/10.1145/3322276.3322294

[25] Consider.it. 2023. Consider.it. https://consider.it/. Last accessed 2023-03-08.

[26] Joseph A Cottam, Andrew Lumsdaine, and Chris Weaver. 2012. Watch this: A taxonomy for dynamic data visualization. In *2012 IEEE Conference on Visual Analytics Science and Technology (VAST)*. IEEE, 193–202.

[27] Alexandra de Requesens Daniel Pearson, Pau Garcia. 2022. Data Strings Domestic Data Streamers. In *Making with Data Physical Design and Craft in a Data-Driven World*. A K Peters/CRC Press, 14–18.

[28] Jason Diceman. 2019. How to Use Dot Voting Effectively? https://dotmocracy.org/. Last accessed 2023-03-08.

[29] Denis Diderot and Jean Le Rond d'Alembert. [n. d.]. *Encyclopédie, ou, Dictionnaire raisonné des sciences, des arts et des métiers*. University of Chicago: ARTFL Encyclopédie Project (Spring 2021 Edition).

[30] Alexandra Diehl, Alfie Abdul-Rahman, Mennatallah El-Assady, Benjamin Bach, Daniel A Keim, and Min Chen. 2018. VisGuides: A Forum for Discussing Visualization Guidelines.. In *EuroVis (Short Papers)*. 61–65.

[31] Evanthia Dimara and Charles Perin. 2019. What is interaction for data visualization? *IEEE TVCG* 26, 1 (2019), 119–129.

[32] Evanthia Dimara, Harry Zhang, Melanie Tory, and Steven Franconeri. 2022. The Unmet Data Visualization Needs of Decision Makers Within Organizations. *IEEE TVCG* 28, 12 (2022), 4101–4112. https://doi.org/10.1109/TVCG.2021.3074023

[33] Anhai Doan, Raghu Ramakrishnan, and Alon Y Halevy. 2011. Crowdsourcing systems on the world-wide web. *Commun. ACM* 54, 4 (2011), 86–96.

[34] Daniel Goddemeyer Dominikus Baur. 2017. DATA FUTURES. http://do.minik.us/projects/ten-questions. Last accessed 2023-03-08.

[35] Doodle. 2023. Doodle. https://doodle.com/. Last accessed 2023-03-08.

[36] Marian Dörk, Daniel Gruen, Carey Williamson, and Sheelagh Carpendale. 2010. A visual backchannel for large-scale events. *IEEE TVCG* 16, 6 (2010), 1129–1138.

[37] Pierre Dragicevic and Yvonne Jansen. 2012. List of Physical Visualizations. dataphys.org/list/tags/participatory. Last accessed 2021-07-18.

[38] Pierre Dragicevic, Yvonne Jansen, and Andrew Vande Moere. 2021. Data Physicalization. In *Springer Handbook of Human Computer Interaction*, Jean Vanderdonckt (Ed.). Springer. https://hal.inria.fr/hal-02113248

[39] Johanna Drucker. 2011. Humanities approaches to graphical display. *Digital Humanities Quarterly* 5, 1 (2011), 1–21.

[40] Easy Dataviz Duarte Jose. 2022. Let's Play with Data. In *Making with Data Physical Design and Craft in a Data-Driven World*. A K Peters/CRC Press, 14–18.

[41] Hunter Dukes and Adam Green. 2021. Visualizing History: The Polish System. https://publicdomainreview.org/collection/visualizing-history-the-polish-system. Last accessed 2023-03-08.

[42] Žarko Dumičić, Katja Thoring, Hermann W Klöckner, and Gesche Joost. 2022. Design elements in data physicalization: A systematic literature review. (2022). https://doi.org/10.21606/drs.2022.660

[43] Christian Eichner, Stefan Gladisch, Heidrun Schumann, and Christian Tominski. 2016. Direct Visual Editing of Node Attributes in Graphs. *Informatics* 3, 4 (2016). https://doi.org/10.3390/informatics3040017

[44] Kahoot! Enterprise. 2013. Kahoot! https://kahoot.it/. Last accessed 2023-03-08.

[45] Framacloud. 2012. Framadate. framadate. Last accessed 2021-07-18.

[46] Feedback Frames. 2023. Let's See What We Think! https://feedbackframes.com/. Last accessed 2023-03-08.

[47] Laurie Frick. 2012. Stress Inventory. https://www.lauriefrick.com/stress-inventory. Last accessed 2023-03-08.

[48] Adrian Furnham and Hua Chu Boo. 2011. A literature review of the anchoring effect. *The Journal of Socio-Economics* 40, 1 (2011), 35–42.

[49] Gephi. 2020. Gephi. . Last accessed 2023-03-08.

[50] Simone Giertz. 2020. Every Day Calendar. . Last accessed 2023-03-08.

[51] Pascal Goffin, Wesley Willett, Jean-Daniel Fekete, and Petra Isenberg. 2014. Exploring the placement and design of word-scale visualizations. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (2014), 2291–2300.

[52] Sarah Goodwin, Jason Dykes, Sara Jones, Iain Dillingham, Graham Dove, Alison Duffy, Alexander Kachkaev, Aidan Slingsby, and Jo Wood. 2013. Creative user-centered visualization design for energy analysts and modelers. *IEEE transactions on visualization and computer graphics* 19, 12 (2013), 2516–2525.

[53] Google. 2023. Get Google Review. http://bit.ly/3Kni5p3. Last accessed 2023-03-08.

[54] David Gotz and Michelle X Zhou. 2009. Characterizing users' visual analytic activity for insight provenance. *Information Visualization* 8, 1 (2009), 42–55.

[55] Pauline Gourlet and Thierry Dassé. 2017. Cairn: A tangible apparatus for situated data collection, visualization and analysis. In *Proc. of the 2017 Conference on Designing Interactive Systems*. 247–258.

[56] Lars Grammel, Melanie Tory, and Margaret-Anne Storey. 2010. How information visualization novices construct visualizations. *IEEE transactions on visualization and computer graphics* 16, 6 (2010), 943–952.

[57] Saul Greenberg, Sheelagh Carpendale, Nicolai Marquardt, and Bill Buxton. 2011. *Sketching user experiences: The workbook*. Elsevier.

[58] Hans Haacke. 1970. Moma Poll. https://www.wikiart.org/en/hans-haacke/moma-poll-1970. Last accessed 2023-03-08.

[59] Kyle Wm Hall, Adam J Bradley, Uta Hinrichs, Samuel Huron, Jo Wood, Christopher Collins, and Sheelagh Carpendale. 2019. Design by immersion: A trans-disciplinary approach to problem-driven visualizations. *IEEE transactions on visualization and computer graphics* 26, 1 (2019), 109–118.

[60] Chris Heape. 2007. The Design Space: the design process as the construction, exploration and expansion of a conceptual space. (2007).

[61] Jeffrey Heer and Ben Shneiderman. 2012. Interactive dynamics for visual analysis. *Commun. ACM* 55, 4 (2012), 45–54.

[62] Joseph M Hellerstein. 2008. Quantitative data cleaning for large databases. *United Nations Economic Commission for Europe (UNECE)* 25 (2008).

[63] Tom Horak, Philip Berger, Heidrun Schumann, Raimund Dachselt, and Christian Tominski. 2020. Responsive matrix cells: A focus+ context approach for exploring and editing multivariate graphs. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2020), 1644–1654.

[64] Eva Hornecker, Trevor Hogan, Uta Hinrichs, and Rosa van Koningsbruggen. 2023. A Design Vocabulary for Data Physicalization. *ACM Trans. Comput.-Hum. Interact.* (aug 2023). https://doi.org/10.1145/3617366 Just Accepted.

[65] Dandan Huang, Melanie Tory, Bon Adriel Aseniero, Lyn Bartram, Scott Bateman, Sheelagh Carpendale, Anthony Tang, and Robert Woodbury. 2014. Personal visualization and personal visual analytics. *IEEE Transactions on Visualization and Computer Graphics* 21, 3 (2014), 420–433.

[66] Jon Huang and Aron Pilhofer. 2011. The Death of a Terrorist: A Turning Point? *New York Times* (2011). https://archive.nytimes.com/www.nytimes.com/interactive/2011/05/03/us/20110503-osama-response.html Last accessed 2023-03-08.

[67] Michael Hunger. 2008. On LEGO Powered Time-Tracking; My Daily Column. https://web.archive.org/web/20131207122823/http://jexp.de/blog/2008/08/on-lego-powered-time-tracking-my-daily-column/. Last accessed 2023-03-08.

[68] Samuel Huron, Petra Isenberg, and Jean Daniel Fekete. 2013. PolemicTweet: Video annotation and analysis through tagged tweets. In *IFIP Conference on Human-Computer Interaction*. Springer, 135–152.

[69] Samuel Huron, Yvonne Jansen, and Sheelagh Carpendale. 2014. Constructing visual representations: Investigating the use of tangible tokens. *IEEE TVCG* 20, 12 (2014), 2102–2111.

[70] Samuel Huron, Romain Vuillemot, and Jean-Daniel Fekete. 2013. Bubble-TV: Live visual feedback for social tv broadcast. In *ACM CHI 2013 Workshop: Exploring and enhancing the user experience for television*.

[71] Samuel Huron, Romain Vuillemot, and Jean-Daniel Fekete. 2013. Visual sedimentation. *IEEE TVCG* 19, 12 (2013), 2446–2455.

[72] Samuel Huron and Wesley Willett. 2021. Visualizations as Data Input? *alt.VIS workshop co-located with IEEE VIS* (2021).

[73] Hiroshi Ishii, Carlo Ratti, Ben Piper, Yao Wang, Assaf Biderman, and Eran Ben-Joseph. 2004. Bringing clay and sand into digital design—continuous tangible user interfaces. *BT technology journal* 22, 4 (2004), 287–299.

[74] Yvonne Jansen and Pierre Dragicevic. 2013. An interaction model for visualizations beyond the desktop. *IEEE TVCG* 19, 12 (2013), 2396–2405.

[75] Kendra Hagerman Jen Causey-Kidder, Lauren Koster. 2013. DIY | easy seating chart. https://somethingturquoise.com/2011/09/02/diy-easy-seating-chart/. Last accessed 2023-03-08.

[76] Sean Kandel, Jeffrey Heer, Catherine Plaisant, Jessie Kennedy, Frank Van Ham, Nathalie Henry Riche, Chris Weaver, Bongshin Lee, Dominique Brodbeck, and Paolo Buono. 2011. Research directions in data wrangling: Visualizations and transformations for usable and credible data. *Information Visualization* 10, 4 (2011), 271–288.

[77] Sean Kandel, Andreas Paepcke, Joseph Hellerstein, and Jeffrey Heer. 2011. Wrangler: Interactive visual specification of data transformation scripts. In *Proc. of the sigACM CHI*. 3363–3372.

[78] Daniel A Keim, Florian Mansmann, Jörn Schneidewind, and Hartmut Ziegler. 2006. Challenges in visual data analysis. In *International Conference on Information Visualisation (IV'06)*. IEEE, 9–16.

[79] Yea-Seul Kim, Katharina Reinecke, and Jessica Hullman. 2017. Explaining the gap: Visualizing one's predictions improves recall and comprehension of data. In *Proc. of ACM CHI*. 1375–1386.

[80] KnowAndBe.live. 2020. Participatory Matrix. . Last accessed 2023-03-08.

[81] Lisa Koeman, Vaiva Kalnikaité, and Yvonne Rogers. 2015. "Everyone Is Talking about It!" A Distributed Approach to Urban Voting Technology and Visualisations. In *Proc. of the 33rd annual ACM conference on human factors in computing systems*. 3127–3136.

[82] Lian Chee Koh, Aidan Slingsby, Jason Dykes, and Tin Seong Kam. 2011. Developing and applying a user-centered model for the design and implementation of information visualization tools. In *2011 15th International Conference on Information Visualisation*. IEEE, 90–95.

[83] Brittany Kondo and Christopher Collins. 2014. Dimpvis: Exploring time-varying information visualizations by direct manipulation. *IEEE TVCG* 20, 12 (2014), 2003–2012.

[84] Travis Kriplean, Jonathan Morgan, Deen Freelon, Alan Borning, and Lance Bennett. 2012. Supporting reflective public thought with considerit. In *Proc. ACM CSCW*. 265–274.

[85] Xingyu Lan, Yanqiu Wu, and Nan Cao. 2023. Affective Visualization Design: Leveraging the Emotional Impact of Data. *IEEE Transactions on Visualization and Computer Graphics* (2023).

[86] Bruno Latour. 2007. *Pensée retenue, pensée distribuée*. Albin Michel.

[87] Sarah Lewis. 2015. Qualitative inquiry and research design: Choosing among five approaches. *Health promotion practice* 16, 4 (2015), 473–475.

[88] Ian Li, Anind Dey, and Jodi Forlizzi. 2010. A Stage-Based Model of Personal Informatics Systems. In *Proc. of CHI '10*. ACM, 557–566. https://doi.org/10.1145/

1753326.1753409

[89] Haihan Lin, Derya Akbaba, Miriah Meyer, and Alexander Lex. 2022. Data hunches: Incorporating personal knowledge into visualizations. *IEEE TVCG* 29, 1 (2022), 504–514.

[90] Jiali Liu and James Eagan. 2021. ADQDA: a cross-device affinity diagramming tool for fluid and holistic qualitative data analysis. *Proc. of the ACM on Human-Computer Interaction* 5, ISS (2021), 1–19.

[91] Zhicheng Liu, John Thompson, Alan Wilson, Mira Dontcheva, James Delorey, Sam Grigg, Bernard Kerr, and John Stasko. 2018. Data illustrator: Augmenting vector design tools with lazy data binding for expressive visualization authoring. In *Proc. of the 2018 ACM CHI*. 1–13.

[92] David Lloyd and Jason Dykes. 2011. Human-centered approaches in geovisualization design: Investigating multiple methods through a long-term case study. *IEEE transactions on visualization and computer graphics* 17, 12 (2011), 2498–2507.

[93] Andrés Lucero. 2012. Framing, aligning, paradoxing, abstracting, and directing: how design mood boards work. In *Proc. of ACM DIS*. 438–447.

[94] Andrés Lucero. 2015. Using affinity diagrams to evaluate interactive prototypes. In *Human-Computer Interaction–INTERACT 2015: 15th IFIP TC 13 International Conference, Bamberg, Germany, September 14-18, 2015, Proc., Part II 15*. Springer, 231–248.

[95] Giorgia Lupi and Stefanie Posavec. 2018. *Observe, collect, draw!: A visual journal*. Princeton Architectural Press.

[96] Allan Maclean and Diane McKerlie. 1995. Design space analysis and use-representations. *Scenario-based design: envisioning work and technology in system development* (1995), 183–207.

[97] Justin Matejka, Michael Glueck, Tovi Grossman, and George Fitzmaurice. 2016. The Effect of Visual Appearance on the Performance of Continuous Sliders and Visual Analogue Scales. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (*CHI '16*). Association for Computing Machinery, New York, NY, USA, 5421–5432. https://doi.org/10.1145/2858036.2858063

[98] Nina McCurdy, Jason Dykes, and Miriah Meyer. 2016. Action design research and visualization design. In *Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization*. 10–18.

[99] Nina McCurdy, Julie Gerdes, and Miriah Meyer. 2018. A framework for externalizing implicit error using visualization. *IEEE TVCG* 25, 1 (2018), 925–935.

[100] Miriah Meyer and Jason Dykes. 2019. Criteria for rigor in visualization design study. *IEEE TVCG* 26, 1 (2019), 87–97.

[101] Hunger Michael. 2008. On LEGO powered time-tracking; my daily column. web.archive.org/web/20170421133649/http://jexp.de/blog/2008/08/on-lego-powered-time-tracking-my-daily-column/. Last accessed 2017-04-21.

[102] Andrew Vande Moere. 2004. Time-varying data visualization using information flocking boids. In *IEEE Symposium on Information Visualization*. IEEE, 97–104.

[103] Sue Montgomery. 2019. Knitting Time of Men and Women Talking. https://twitter.com/MontgomerySue/status/1128092422314123265. Last accessed 2023-03-08.

[104] Luiz Morais, Yvonne Jansen, Nazareno Andrade, and Pierre Dragicevic. 2020. Showing data about people: A design space of anthropographics. *IEEE Transactions on Visualization and Computer Graphics* 28, 3 (2020), 1661–1679.

[105] Tamara Munzner. 2009. A nested model for visualization design and validation. *IEEE transactions on visualization and computer graphics* 15, 6 (2009), 921–928.

[106] Tamara Munzner. 2014. *Visualization analysis and design*. CRC press.

[107] University of Chicago Booth School of Business. 2020. Mindworks. . Last accessed 2023-03-08.

[108] Dietmar Offenhuber. 2019. Data by proxy—Material traces as autographic visualizations. *IEEE TVCG* 26, 1 (2019), 98–108.

[109] Roman Ondák, Bernhart Schwenk, and Magali Arriola. 2008. *Roman Ondák: measuring the universe*. Vol. 9. Jrp Ringier.

[110] Michael Oppermann and Tamara Munzner. 2020. Data-first visualization design studies. In *2020 IEEE Workshop on Evaluation and Beyond-Methodological Approaches to Visualization (BELIV)*. IEEE, 74–80.

[111] Georgia Panagiotidou, Sinem Gorucu, and Andrew Vande Moere. 2020. Data Badges: Making an Academic Profile Through a DIY Wearable Physicalization. *IEEE computer graphics and applications* PP (09 2020). https://doi.org/10.1109/MCG.2020.3025504

[112] Charles Perin, Romain Vuillemot, and Jean-Daniel Fekete. 2014. À Table! improving temporal navigation in soccer ranking tables. In *Proc. of ACM CHI*. 887–896.

[113] Craig Plain. 2007. Build an affinity for KJ method. *Quality Progress* 40, 3 (2007), 88.

[114] Zachary Pousman, John Stasko, and Michael Mateas. 2007. Casual information visualization: Depictions of data in everyday life. *IEEE TVCG* 13, 6 (2007), 1145–1152.

[115] Donghao Ren, Bongshin Lee, and Matthew Brehmer. 2018. Charticulator: Interactive construction of bespoke chart layouts. *IEEE TVCG* 25, 1 (2018), 789–799.

[116] Hugo Romat, Nathalie Henry Riche, Ken Hinckley, Bongshin Lee, Caroline Appert, Emmanuel Pietriga, and Christopher Collins. 2019. ActiveInk: (Th)

Inking with Data. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.

[117] John Rooksby, Mattias Rost, Alistair Morrison, and Matthew Chalmers. 2014. Personal Tracking as Lived Informatics. In *Proc. of CHI '14*. ACM, 1163–1172. https://doi.org/10.1145/2556288.2557039

[118] Sadia Rubab, Junxiu Tang, and Yingcai Wu. 2021. Examining interaction techniques in data visualization authoring tools from the perspective of goals and human cognition: a survey. *Journal of Visualization* 24 (2021), 397–418.

[119] Stefan Sagmeister. 2023. The Happy Show. https://sagmeister.com/work/the-happy-show/. Last accessed 2023-03-08.

[120] Bahador Saket, Samuel Huron, Charles Perin, and Alex Endert. 2019. Investigating direct manipulation of graphical encodings as a method for user interaction. *IEEE TVCG* 26, 1 (2019), 482–491.

[121] Salesforce. 2023. Tableau Prep. https://feedbackframes.com/. Last accessed 2023-03-30.

[122] Arvind Satyanarayan and Jeffrey Heer. 2014. Lyra: An interactive visualization design environment. *Computer graphics forum* 33, 3 (2014), 351–360.

[123] Kim Sauvé, Pierre Dragicevic, and Yvonne Jansen. 2023. Edo: A Participatory Data Physicalization on the Climate Impact of Dietary Choices. In *Proc. of ACM TEI*. ACM.

[124] Michael Sedlmair, Miriah Meyer, and Tamara Munzner. 2012. Design study methodology: Reflections from the trenches and the stacks. *IEEE transactions on visualization and computer graphics* 18, 12 (2012), 2431–2440.

[125] Edward Segel and Jeffrey Heer. 2010. Narrative visualization: Telling stories with data. *IEEE transactions on visualization and computer graphics* 16, 6 (2010), 1139–1148.

[126] Ben Shneiderman and Pattie Maes. 1997. Direct manipulation vs. interface agents. *interactions* 4, 6 (1997), 42–61.

[127] specialprojects.studio. 2013. Bit Planner. https://specialprojects.studio/project/bit-planner/. Last accessed 2023-03-08.

[128] Michael Stern, Don A. Dillman, and Jolene D. Smyth. 2007. Visual Design, Order Effects, and Respondent Characteristics in a Self-Administered Survey. *Survey Research Methods* 1, 3 (Dec. 2007), 121–138. https://doi.org/10.18148/srm/2007.v1i3.600

[129] Domestic Data Streamers. 2020. Drip by Tweet. . Last accessed 2023-03-08.

[130] Hariharan Subramonyam, Steven M. Drucker, and Eytan Adar. 2019. Affinity Lens: Data-Assisted Affinity Diagramming with Augmented Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) (CHI '19). Association for Computing Machinery, New

York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300628

[131] Sebastian Rønde Thielke and Anders Højmose. 2008. Daily Stack. ciid.dk/education/portfolio/idp09/courses/tangible-user-interface/projects/daily-stack/. Last accessed 2017-04-21.

[132] Alice Thudt, Uta Hinrichs, Samuel Huron, and Sheelagh Carpendale. 2018. Self-Reflection and Personal Physicalization Construction. In *Proc. of the 2018 ACM CHI* (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173728

[133] tiermaker.com. 2023. Alignment Chart. https://tiermaker.com/categories/dnd/alignment-chart-rankings-364890. Last accessed 2023-03-08.

[134] tiermaker.com. 2023. Tier List Community Rankings. https://tiermaker.com/categories/dnd/alignment-chart-rankings-364890. Last accessed 2023-03-08.

[135] Twitter. 2023. About Twitter Polls. https://help.twitter.com/en/using-twitter/twitter-polls. Last accessed 2023-03-08.

[136] Brygg Ullmer, Hiroshi Ishii, and Robert J. K. Jacob. 2005. Token+constraint Systems for Tangible Interaction with Digital Information. *ACM Trans. Comput.-Hum. Interact.* 12, 1 (mar 2005), 81–118. https://doi.org/10.1145/1057237.1057242

[137] Nina Valkanova, Robert Walter, Andrew Vande Moere, and Jörg Müller. 2014. MyPosition: sparking civic discourse by a public interactive poll visualization. In *Proc. ACM CSCW*. 1323–1332.

[138] Jagoda Walny, Samuel Huron, and Sheelagh Carpendale. 2015. An exploratory study of data sketching for visual representation. In *Computer Graphics Forum*, Vol. 34. Wiley Online Library, 231–240.

[139] Kendra A. Wannamaker, Sandeep Zechariah George Kollannur, Marian Dörk, and Wesley Willett. 2021. I/O Bits: User-Driven, Situated, and Dedicated Self-Tracking. In *Proceedings of the 2021 ACM Designing Interactive Systems Conference* (Virtual Event, USA) (DIS '21). Association for Computing Machinery, New York, NY, USA, 523–537. https://doi.org/10.1145/3461778.3462138

[140] When2meet. 2012. When2meet. When2meet. Last accessed 2023-03-08.

[141] Wesley Willett, Jeffrey Heer, and Maneesh Agrawala. 2007. Scented widgets: Improving navigation cues with embedded visualizations. *IEEE TVCG* 13, 6 (2007), 1129–1136.

[142] Haijun Xia, Nathalie Henry Riche, Fanny Chevalier, Bruno De Araujo, and Daniel Wigdor. 2018. DataInk: Direct and creative data-oriented drawing. In *Proc. of the 2018 ACM CHI*. 1–13.

[143] Ji Soo Yi, Youn ah Kang, John Stasko, and Julie A Jacko. 2007. Toward a deeper understanding of the role of interaction in information visualization. *IEEE TVCG* 13, 6 (2007), 1224–1231.