



Charagraph: Interactive Generation of Charts for Realtime Annotation of Data-Rich Paragraphs

Damien Masson, Sylvain Malacria, Géry Casiez, Daniel Vogel

► To cite this version:

Damien Masson, Sylvain Malacria, Géry Casiez, Daniel Vogel. Charagraph: Interactive Generation of Charts for Realtime Annotation of Data-Rich Paragraphs. CHI 2023 - ACM Conference on Human Factors in Computing Systems (CHI 2023), ACM, Apr 2023, Hamburg, Germany. 10.1145/3544548.3581091 . hal-04017629

HAL Id: hal-04017629

<https://hal.science/hal-04017629v1>

Submitted on 7 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Charagraph: Interactive Generation of Charts for Realtime Annotation of Data-Rich Paragraphs

Damien Masson

Cheriton School of Computer Science, University of Waterloo
Waterloo, Canada
dmasson@uwaterloo.ca

Géry Casiez^{†‡}

Univ. Lille, CNRS, Inria, Centrale Lille, UMR 9189 CRIStAL
Lille, France
gery.casiez@univ-lille.fr

Sylvain Malacria*

Univ. Lille, Inria, CNRS, Centrale Lille, UMR 9189 CRIStAL
Lille, France
sylvain.malacria@inria.fr

Daniel Vogel

Cheriton School of Computer Science, University of Waterloo
Waterloo, Canada
dvogel@uwaterloo.ca

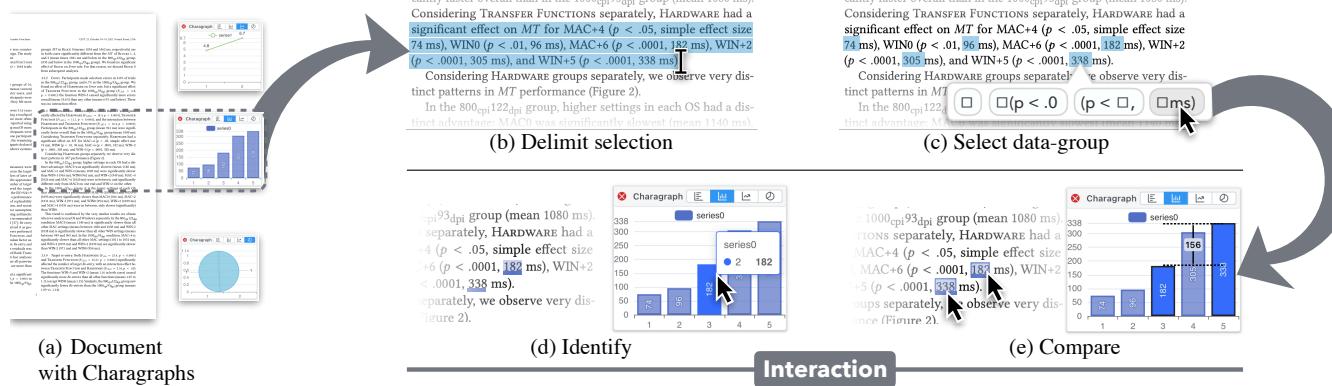


Figure 1: (a) Charagraphs are in-situ visualizations of numeric data included within text that are dynamically generated (b) by delimiting a selection and (c) selecting a data group. Charagraphs support common data exploration tasks through interactive features such as (d) identifying and (e) comparing values.

ABSTRACT

Documents often have paragraphs packed with numbers that are difficult to extract, compare, and interpret. To help readers make sense of data in text, we introduce the concept of Charagraphs: dynamically generated interactive charts and annotations for in-situ visualization, comparison, and manipulation of numeric data included

within text. Three Charagraph characteristics are defined: leveraging related textual information about data; integrating textual and graphical representations; and interacting at different contexts. We contribute a document viewer to select in-text data; generate and customize Charagraphs; merge and refine a Charagraph using other in-text data; and identify, filter, compare, and sort data synchronized between text and visualization. Results of a study show participants can easily create Charagraphs for diverse examples of data-rich text, and when answering questions about data in text, participants were more correct compared to only reading text.

*Also with University of Waterloo.

†Also with Institut Universitaire de France.

‡Also with University of Waterloo.

Author's version.

Paper published in CHI '23, April 23–28, 2023, Hamburg, Germany

Damien Masson, Sylvain Malacria, Géry Casiez, Daniel Vogel. Charagraph: Interactive Generation of Charts for Realtime Annotation of Data-Rich Paragraphs.

<https://doi.org/10.1145/3544548.3581091>

CCS CONCEPTS

• Human-centered computing → Visualization systems and tools; Interactive systems and tools.

KEYWORDS

reading interfaces, active reading, annotation, interactive documents, visualization, in-text data, data-rich documents

ACM Reference Format:

Damien Masson, Sylvain Malacria, Géry Casiez, and Daniel Vogel. 2023. Charagraph: Interactive Generation of Charts for Realtime Annotation of Data-Rich Paragraphs. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 18 pages. <https://doi.org/10.1145/3544548.3581091>

1 INTRODUCTION

Documents and scientific reports are often filled with numerical values. These regularly interrupt the flow of sentences with quantities like means, medians, confidence intervals, and p-values. While these are important to support claims, they can make it hard to compare and interpret patterns in numeric data due to the linear structure of sentences.

In fact, the use of sentences to present numbers has long been criticized [37, 72, 153]. Edward Tufte [26] notes that: “*The conventional sentence is a poor way to show more than two numbers because it prevents comparisons within the data*”.

While experts almost always recommend visualizations such as charts to present numeric data [3, 26, 44, 67, 153], they may not always be included or useful. Authors may omit a visualization because it takes time to create [66], occupies space in the document [3, 53, 64, 67], requires substantial effort and skill [23, 43, 66], or seems unnecessary to present a few numbers [34, 44, 67]. Worse, even when authors include visualizations, readers might ignore them if they are poorly designed or not aligned with the reader’s goal [101], overblown [149], or not accessible [69, 156]. And readers might struggle to connect text and figures because the layout of the document often separates them [6, 26].

One response to issues when data is included in text has been to provide interactive connections between text and existing visualizations in the document. Because archival documents are not easily modified once disseminated [94], most proposed solutions use some form of automatic linking [62, 75, 76, 98, 113] or an authoring system to create dynamic interactive documents from the beginning [74, 79–81, 144]. However, these approaches remain dependent on the author’s willingness to embed figures. Another strategy proposed by Elastic Documents [7] is to preprocess a document to generate new visualizations from the tables and then include ways to filter them while reading. But this requires inferring what readers need a priori and assumes that the data is presented in structured tables. What if, as is often the case, the data exists only in sentences?

We introduce the concept of *Charagraphs*, which are dynamically generated interactive charts and annotations for in-situ visualization, comparison, and manipulation of numeric data included within paragraphs of text. We explore the design of Charagraphs and contribute a system that allows readers to quickly generate them from existing documents. Instead of expecting authors to embed well-designed visualizations, our system empowers readers to quickly craft their own visualizations, customized to their needs, from the data presented in texts. Although our system can also extract data from document tables, we focus on the problem of extracting data from sentences. The created Charagraphs use data visualizations to offload cognitive work to the perceptual system [78, 101] alongside interactive features to compare and manipulate the data. This

can help when data is presented only in-text, or when the existing static visualizations are ineffective and readers would prefer a different representation. In a user study, participants created their own Charagraphs to annotate documents and answer questions. All participants found the system easy to use and were able to create the visualization they had in mind. Furthermore, all participants preferred to have Charagraphs available; had lower mental demand, effort, and frustration; and were more correct in completing comprehension, comparison, and interpretation tasks when they had access to Charagraphs (98.61%) compared to text alone (92.59%).

2 BACKGROUND

We first examine what numeric data can be expected in text and then review the different ways of presenting data and their advantages and drawbacks depending on the context.

2.1 What Values Are Reported In-Text?

Most documents report numerical values in text to support claims because “*a stronger quantitative result is more informative than a weaker, qualitative result that subsumes it*” [39]. For example, statistical results are often reported alongside means or mean differences [38, 39]. Practises vary across domains and publishers.

The American Psychological Association (APA) style instructs authors to “*emphasize particular data in the text when they aid in the interpretation of the findings*” and when reporting inferential statistics, the data in text “*should allow readers to confirm the basic reported analyses (e.g., cell means, standard deviations, sample sizes, correlations)*” [5]. In fact, the APA style normalizes this process by presenting a list of statistical abbreviations and symbols to be used in text (e.g., “M” for “mean”, “Mdn” for “median”).

Within the HCI community, recommendations for transparent statistics call for quantitative results to provide effect sizes (including the direction of the difference, the unit, and the type of estimate and uncertainty) [53] and interval estimates [30]. These values are easier to interpret and compare graphically [30, 39, 87], but textual reports are commonly chosen as an alternative because “*space may not always permit a graphical report*” [53].

As a result, scientific documents often report numerical data in a semi-structured way using specific symbols and patterns.

2.2 How to Best Present Numerical Values?

There are three ways to include numerical values in documents: inline with sentence text, in a table, and in some form of graphical visualization [26, 153]. It has long been suggested that these representations are not equivalent and should be chosen carefully [5, 37, 43, 64, 65, 78, 87, 92]. We review two relevant situations.

When there are few values: While visualization experts generally agree that “*under no circumstance should text be used by itself to convey important statistical data if more than a very few facts are to be presented*” [37] (also [26, 153]), they are divided on which representation to use when faced with small datasets. Some state that graphical visualizations with little content indicate a lack of purpose [34], take up a lot of space [3, 64, 67], add to readers’ cognitive load [5, 44], and are costly to create [66]. Yet, experimental evidence suggests a general advantage for visualizations such as bar charts, even when presenting only two values [31, 70, 138].

When presenting the same information: Two different representations of the same information are often beneficial [78, 131, 136]. For example, Ottley et al. found that visualizations make it easy to identify critical information, but the precise numerical value was easier to extract from text [108]. Early on, Tufte advocated for the combination of words, numbers, and graphics and deplored the segregation between figures and their associated text in scientific manuscripts [26]. Experimentally, the Multimedia Principle shows that people learn better when an image complements text [96] and strong evidence suggests that illustrated texts improve reading comprehension and recall [33, 45]. The general recommendation is that authors should use text reinforced by graphics [37].

2.3 Issues With In-Text Data and Rationale for Reader-Generated Interactive Charts

Data is usually reported in text, and at best presented visually in static charts, which remains far from ideal for readers.

First, *textual data is difficult to interpret*. Numeric values make it hard to perform analytic tasks such as identifying trends [99] and are presented linearly in sentences broken at arbitrary points depending on document layout [149]. Despite the advantages of data visualizations, a document might present data only in text or tables. In contrast, visualizations help understand the information [99, 139], discover facts and insights about the data [72, 101, 103] and perform general interpretive operations such as locating, deriving, interpolating, and comparing values [19, 29, 37, 99]. As such, allowing readers to generate visualizations would be beneficial.

Second, *data is segregated in documents* and it is often tedious for readers to combine and visualize the results reported in different documents or different parts of the same document [104], because documents segregate the information and the data, even when related [149]. If readers can generate charts from the text, then they could do it by combining two distinct parts of a document.

Third, *textual report of results might be misinterpreted*. Instead of using dichotomous statistical testing such as thresholds on p-values, it has been suggested that reporting interval estimates such as confidence intervals could reduce misinterpretations and improve understanding of results [30, 53]. Yet, these estimates are often best conveyed graphically [56, 87]. For example, readers will have difficulties deciding if two confidence intervals reported in text overlap, and by how much (e.g., how different is 95% CI [-0.08, 0.5] from 95% CI [0.1, 0.4]?).

We posit that the above mentioned issues could be alleviated by supporting the dynamic generation of interactive charts and annotations generated by the reader, in order to offer in-situ visualization, comparison, and manipulation of numeric data included within paragraphs of text. In the remainder of this document, we refer to such interactive charts generated by the reader as Charagraphs.

3 RELATED WORK

Reading is a multifaceted activity and skilled readers use a variety of strategies to achieve their goals [2, 104]. In this section, we review previous types of digital augmentations to help people actively read and extract information from documents, and highlight how previous existing solutions differ from Charagraphs.

3.1 Supporting Active Readers

Active reading was first identified with physical papers as “*the combination of reading with critical thinking and learning, and involves not just reading per se, but also underlining, highlighting, and commenting.*” [2, 115]. Initial efforts focused on bringing the capabilities of pen and paper to digital environments [60, 86, 105, 112, 137] and proceeded to extend them through powerful annotation [126, 130], navigation [148], diagramming [143], and note-taking features [57]. Most readers engage in active reading to externalize thoughts [143], reduce cognitive load [129], improve understanding [124], emphasize key concepts [2], help memorization [100], aid in later retrieval [130] and share marks with others [147].

Since then, Bret Victor proposed a new definition by introducing *Explorable Explanations* that “*encourage truly active reading*” and turn text into “*an environment to think in*” instead of “*information to be consumed*” [17]. Victor’s definition of active reading focuses on interactive documents to explore, build intuition, make discoveries, and reinforce or challenge knowledge. This idea has been applied to different contexts, such as transparent statistical reports [32] and visual exploration [46] (see Hohman et al. [59] for a review).

Both definitions express similar objectives and are used interchangeably in the literature, but differ in the role of the reader. In Victor’s definition, readers engage with interactive elements. In the original definition, readers create their own marks and this act may be as important, if not more [68], than the marks left on the document [2, 125, 143, 157].

These active reading ideas inspired our approach. In the same way that readers highlight sentences and write in the margin [2, 115], Charagraphs can serve as a new kind of document annotation. The act of creating a Charagraph prompts the reader to think critically and learn, following the original definition of active reading. And the resulting charts are compact representations of data that can be quickly understood without more context [70, 153], and serve as visual landmarks while browsing a document [85]. But Charagraphs also support Victor’s definition: once created, a Charagraph lets readers explore, make discoveries, and reinforce their knowledge about the data presented in the text.

3.2 Augmenting Existing Documents

Data-rich documents are notoriously challenging to write [21] and visualization difficult to create [23, 43, 66]. Systems have been proposed to help authors write better documents [21] and create better data visualizations [28, 61, 89, 135, 140, 155]. However, the created documents remain static which limits the exploration possibilities for readers [17], and forces a split of attention between text and figures [6]. Other systems have targeted the design of interactive documents [24, 81, 144], but these require a substantial amount of work for authors and are often incompatible with PDF-based workflows such as academic publishing.

An alternate approach adopted by our work is to augment existing documents semi-automatically. Such systems have been used to direct readers’ attention to specific sentences [22, 40, 154, 158], provide definitions of unknown terms and symbols [54], integrate relevant information and commentary from follow-up work [120], and generate summaries [142]. Some systems help readers interpret and manipulate the results by enhancing the visualizations

already present in the document. For example, redesigning charts to be more useful [128], with added overlays [73], and interactive features [88, 95]. Or the text can be leveraged to automatically annotate existing charts [62, 76] and interactively connect text and charts [113]. However, all of these approaches assume the author originally included visualizations in the document.

Another idea explored by Elastic Documents is to generate new chart visualizations from tables in a preprocessing data extraction step performed on the entire document [7]. Data in tables are extracted and used to generate different static chart visualizations. Readers can filter and browse the generated charts or select a sentence in the document and let the system show the most relevant charts. The relevance is calculated by matching the keywords from the sentence and the table headers.

Charagraphs are similar to Elastic Documents in that they are new visualizations complementing existing documents. However, Charagraphs are annotations dynamically generated by readers and for data contained within text such as sentences. This makes a Charagraph highly customized by the reader for their needs and tightly coupled with the text through bidirectional interactive features to manipulate and compare the data.

3.3 Generating Visualizations from Text

Several tools propose methods to transform high-level textual descriptions into data visualizations [41, 90, 102, 133] and infographics [27, 117] (see Shen et al. [134] for a survey). However, these approaches differ from Charagraphs in that they are essentially visualization authoring tools with natural language as the input modality. Thus, they require a curated dataset provided by the author and use short queries specifically crafted to generate a visualization from the dataset by referencing metadata like column headers (e.g., “How much do various cars weigh?” or “draw a line chart of daily sales forecasts” [141]).

Closer to our approach are methods to augment documents with visualizations. Essentially the content of the document forms a query to generate a visualization using an associated dataset. For example, Kim et al. help readers relate to distances and areas mentioned in text by re-expressing them on a map and using familiar landmarks [71]. Hullman et al. generalized this idea to help readers understand physical measurements such as weights, lengths, and volumes with different strategies to re-express them in a relatable way such as comparing them with familiar objects [63]. An optimization method and a large dataset of familiar objects is used to generate re-expressions using both text and images. Similarly, Contextifier [62] and NewsViews [42] generate contextual visualizations when reading news articles. Both systems pre-process

the document text and search for relevant data in existing external databases. NewsViews generates interactive annotated maps using commonly available data like unemployment and education statistics. Contextifier generates stock charts annotated with news article content using commonly available stock data.

In contrast, Charagraphs are generated by readers using only data contained within the text of a document. Using an external database is not necessary, and may not even be possible since text data within scientific documents may not be easily accessed or even available in common databases. In addition, Charagraphs support many forms of chart visualizations, and are designed to be interactively customized with strong linkages with the text data.

4 CHARACTERIZATION OF CHARAGRAPH

A Charagraph is a visualization that augments in-text data: this visualization comprises text annotations overlaying the document and a chart. Both the text annotations and the chart are interactive and in-sync: modifications and interactions done on one representation are reflected on the other.

We characterize a Charagraph along three dimensions: 1) what in-text *information* can be leveraged; 2) how to *present* the textual and graphical representation; and 3) what *interactions* can be performed with both representations.

4.1 Information Leveraged

Designing an effective visualization often requires understanding the data, its provenance, and its meaning [101]. Fortunately, paragraphs are semantically rich. Beyond numbers, the structure and formatting of sentences, the words, the style, and the suffix or prefix used with numbers provide valuable information about the data. We consider the different information that can be extracted from the text and their impact on the graphical representation encodings (marks) [13], embellishment (visual channels) [9], and meta-data (titles, labels, and legend).

- **Data** corresponds to the values reported in text that are of interest to the reader (Figure 2a). These are often numerical, but could be categorical such as country names. Without additional information, meaning can be inferred. For example, numbers adding up to a 100 are likely proportions.
- **Statistical terms** explain *how* the data was obtained. For example, how the data was aggregated or what transformation was applied to it. These terms generally precede values (e.g., $M=12$, $SD=1.2$ or 95% CI [1, 2]), and can inform the choice of encodings. For example, error bars for standard deviation (Figure 2b).
- **Units** give general information on *what* the data is. They generally follow the values (e.g., 25%, 33 years, 8°C) and might inform

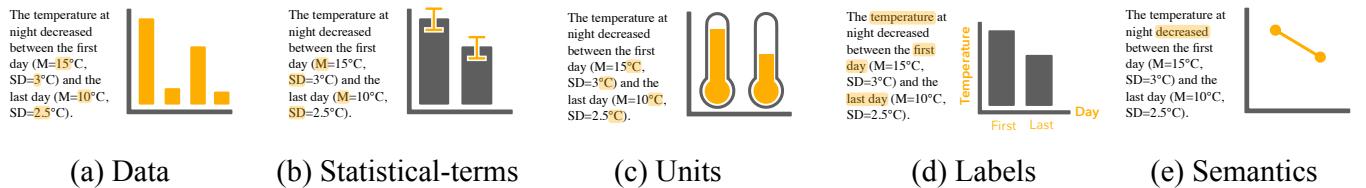


Figure 2: Information that can be leveraged from the text with examples of the design of the Charagraph.

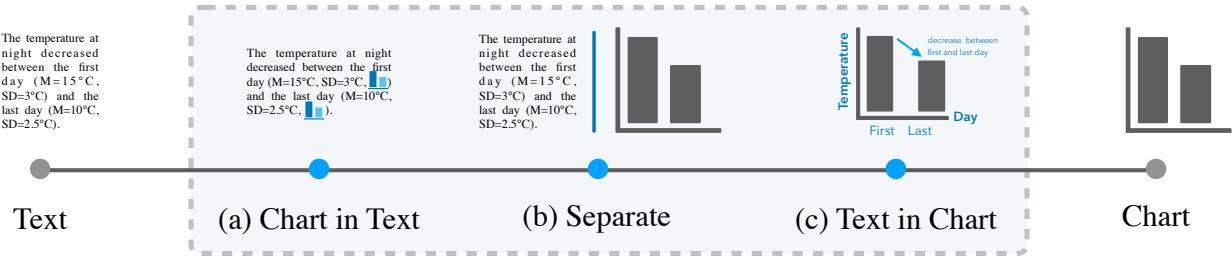


Figure 3: Level of integration of the chart and text representations. Dashed outline indicates levels specific to Charagraphs.

the choice of encodings. For example, a percent suggests a pie chart whereas a timeline may be best for years. Units can also suggest embellishment, such as representing temperature data as bars styled as thermometer (Figure 2c).

- **Labels** describe the context of the data in the document (Figure 2d). Their placement relative to the data depends on the sentence structure (e.g., the first day of May [...] was 15°C). It can serve as meta-data to label data points, legends, and axes and also change embellishment. For example, colouring a series blue if nearby text refers to an ocean.
- **Semantics** correspond to *why* the data was included in the document (e.g., it dropped from [data] to [data]). Semantic-aware Charagraphs might change the encodings depending on the reason; a line chart better represents a sentence describing a trend, whereas a bar chart better represents a comparison (Figure 2d).

4.2 Level of Integration

The presentation of the textual and graphical forms offers one dimension to control the narration. Mandl and Levin note that “*the medium processed in the first place can elicit an attitude, which determines the processing of the other medium*” [93]. The prominence of each representation can be manipulated to orient readers toward a more controlled (text-driven or author-driven) or exploratory (visualization-driven or reader-driven) narrative [132]. Figure 3 lists the different levels of integration of the graphical and textual information, going from only text to only charts.

- **Chart in Text** puts the emphasis on the textual representation and follows the narration of the document. A Charagraph can achieve this by using word-scale visualizations [47, 49, 149] (also called micro visualizations [110] or word-sized graphics [12, 149])

that are known to help readers retain information while reading [48] and also support interaction [50, 51, 79] (Figure 3a).

- **Separate** lets readers decide which representation best suits their needs. A Charagraph achieves this by having both representations equally visible, shown side-by-side in their own separate views (Figure 3b).
- **Text in Chart** commonly referred to as an “annotated chart” puts the emphasis on the graphical representation. Readers are less guided than with a strict sentence order (Figure 3c). Annotations are a common device to support data-driven story-telling [123]. Similarly, a Charagraph achieves this by presenting the chart visualization with textual annotations to emphasize a few elements (see Ren et al. [122] for a review of annotation types).

4.3 Interaction Contexts

A Charagraph acts as a static visualization when not being interacted with. Otherwise, it has two interaction contexts based on the position of the pointer: in-text and in-chart. Previous work on reading behaviours with interactive documents [25, 46, 160] suggests that readers use in-text interactions to follow the document narrative and in-chart interactions to explore and engage with the data. To connect the representations, feedback is synchronized between text and visualizations.

- **In-Text** interaction is most likely during close-reading, when sentences are read carefully one after the other. As such, the interaction should be limited to connecting text to a visualization to support the narrative rather than distract from it. For example, the pointer position can serve as proxy for reading position so only values included in the current sentence under the pointer

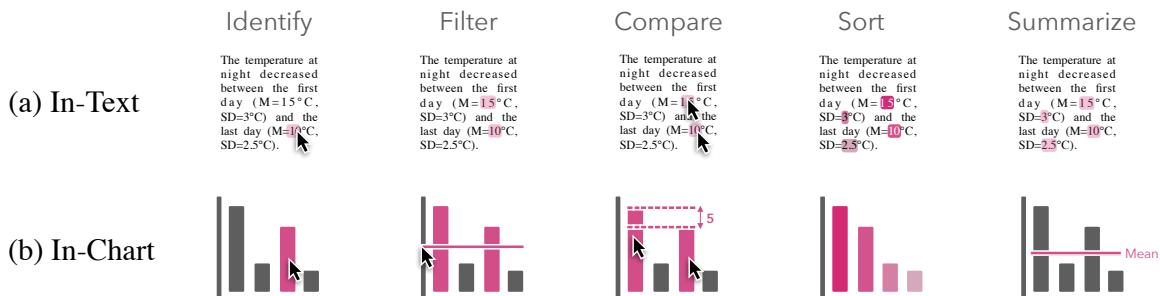


Figure 4: Interaction contexts for different data exploration tasks.

Day	Mean (°C)	SD
Sunday	22.1	4.2
Monday	23.9	3.9
Tuesday	21	4
Wednesday	22.5	2.3

(a) Delimit selection

From students' feedback in the 7-point Likert post-task questionnaire, we found that students did not feel nervous ($M=2.25$, $SD=0.14$), anxious ($M=2.26$, $SD=0.15$), uncomfortable ($M=2.1$, $SD=0.13$), or being distracted by the student end of Glancee ($M=2.3$, $SD=0.14$). And being monitored only made them somewhat engaged ($M=4.21$, $SD=0.21$). Overall, students were willing to use such a tool in the future ($M=3.36$, $SD=0.15$).

During the study, we found that no student participants turned on the camera (several students turned on the camera and then

From students' feedback in the 7-point Likert post-task questionnaire, we found that students did not feel nervous ($M=2.25$, $SD=0.14$), anxious ($M=2.26$, $SD=0.15$), uncomfortable ($M=2.1$, $SD=0.13$), or being distracted by the student end of Glancee ($M=2.3$, $SD=0.14$). And being monitored only made them somewhat engaged ($M=4.21$, $SD=0.21$). Overall, students were willing to use such a tool in the future ($M=3.36$, $SD=0.15$).

=□ (M=□, SD=□, , SD=□) □, SD=0.1 □, participants turned on the camera (several students turned on the camera and then

(b) Select a data-group

Figure 5: Selecting data to visualize: (a) delimiting using traditional text select or marquee selection; (b) selecting one of the data groups suggested by the system.

are highlighted with values before or after the pointer position de-emphasized or hidden. Figure 4a illustrates the minimal feedback in text as a result of the synchronization with the chart.

- **In-Chart** interaction indicates that readers are going beyond the document narrative. As such, the interaction should apply to all the data from the Charagraph, and can support common data exploration tasks such as comparing, filtering, sorting, and summarizing [4] (Figure 4b).

5 CHARAGRAPH DOCUMENT VIEWER

We built a system to generate Charagraphs inside a document viewer. We focus on giving readers tools to create and customize Charagraphs instead of fully automating their creation. We argue that, akin to active reading, the creation process is part of sense-making. By enabling readers to select and decide on the best visualization, they actively participate in understanding the text and the data. From our characterization, we focus on the “Separate” level of integration as it offers the maximum bandwidth of information for both representation and let readers decide on the representation to use. Similarly, the system only leverages information in “Data”, “Statistical-terms” and “Units” and gives the tools to customize the Charagraph further. A live version of our system is accessible online: <http://ns.inria.fr/loki/charagraph>

5.1 Select In-Text Data

A reader who wants to visualize part of data included in text first needs to delimit where it is located in the document and then specify the data of interest through a “data-group menu”.

- **Delimit a Selection.** There are two mechanisms to let readers delimit a selection before prompting the data-group menu.

- **Traditional Text Selection** use standard text segmentation methods such as dragging to have a selection snapping to letters and triple-clicking to select a whole sentence.
- **Rectangular Marquee Selection** allows readers to select all text contained within a rectangle formed by the selection. This tries to address the limitations of traditional text selection caused by small motor errors [11], snapping mechanisms [20], and unrelated elements that are selected in documents that do not preserve sentence-structures such as PDF. For example, selecting the column of a table in a PDF often selects other columns or entire rows because the selection is designed for sentences.

To distinguish between selection methods, a marquee selection must be started outside the text. In both cases, readers only need

to select a text portion containing the data of interest to guide the system and reduce the number of data-group suggestions.

- **Select a Data-Group.** A data-group is a group of values deemed semantically similar by the system. Because data included in text can take various forms and may be intertwined with other information, readers can refine their selections using system-generated suggestions of data-groups.

When a portion of text containing data is selected, a data-group menu is displayed below the selected portion (Figure 5b). The groups are sorted based on the number of values they contain. The first group always contains all the values in the selected portion, whereas the other groups are formed based on the text preceding and following the values. This allows to form groups based on the “statistical-terms” and “units” as per our characterization (Section 4) and potentially also works when the sentence use a similar structure to introduce numbers. For example, text reporting mean and standard deviation in APA-style will have at least two data groups for values preceded by “M=” and “SD=” (Figure 5). Similarly, text reporting the evolution of a statistic over time will often use prepositions such as “from” and “to” that would distinguish them.

Options in the menu are named after their preceding and following text, and the position of the value is represented by a square. Readers can preview the selection by hovering over each option. When the cursor is placed above an option, the corresponding numeric values in the text are highlighted (Figure 5b). Clicking an option immediately creates a Charagraph from the selected values.

- **Edit the Selection.** After a Charagraph is created, readers can add or remove values to further refine the selection. A right-click on one of the underlined values in the text opens a context menu with the option to remove the value. Right-clicking a data point in the chart opens the same context menu. A right-click on a value not currently included opens a context menu with the option to add the value to the current Charagraph.

5.2 Customize the Visualization

By default, newly created Charagraphs appear in the margin, show values with vertical bars, and have numbered tick and legend labels. Readers can change those defaults to customize the visualization. Similarly, all Charagraphs can be resized and moved using direct dragging manipulations.

- **Change Visualization.** Readers can choose visualizations more suited to their task [127] or the data [119]. The system supports

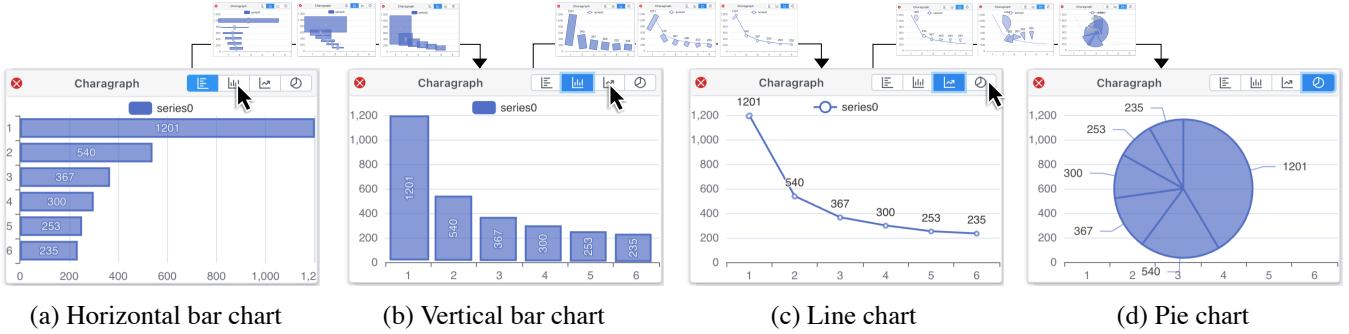


Figure 6: Changing visualizations. All transitions are animated to keep track of changes.

popular chart types [10] and related visualizations that can be generated from data typically found in-text: horizontal and vertical bar charts for aggregated quantities, such as means, and for comparison between data series; line charts to visualize the evolution of a measurement; and pie charts to show proportions (shown to have a slight advantage over bar charts for more complex data [139]). Readers may change the visualization using buttons above the Charagraph (Figure 6). To help readers track the data and understand the new visualization, the change is animated. The current representation “morphs” into the new one, following animated transition recommendations [55]. For example, each bar in a bar chart progressively morphs into each marker counterpart in a line chart (transition Figure 6b-c).

5.2.2 Rename Elements. By default, the axis ticks and legends are labelled using a whole number index based on their order of appearance in the text. The index numbers and any other text element in the Charagraph can be renamed by direct manipulation. Double clicking any text element enables readers to change the text associated with that element such as tick labels or legends.

5.3 Combine Information

Readers can combine data from different parts of a document by merging Charagraphs through a drag-and-drop interaction (Figure 7). Those compound Charagraphs support multiple series to compare different results and identify trends, and using other series as error bars to visualize the variability and uncertainty of the data.

5.3.1 Add Series. Readers can create a Charagraph with two or more series by creating a Charagraph for each series separately and then merging them through a drag and drop interaction. To reveal this feature to users, drop zones overlay all other Charagraphs whenever a Charagraph is dragged (Figure 7a). If the Charagraph is dropped over the “Add Series” zone of another Charagraph, it is modified to include the new series and the dragged Charagraph is removed. If the Charagraph is a bar chart, it becomes a grouped bar chart (Figure 7b).

5.3.2 Add Error Bars. Adding error bars follows the same interaction as adding series, except the Charagraph with the error bar values needs to be dropped in the zone called “Add Error Bars” (Figure 7a). The dragged Charagraph is turned into error bars that overlay the Charagraph (Figure 7c).

5.4 Interact with Charagraphs

Charagraphs support interactivity to perform data exploration tasks [159], facilitate the connection between text and chart [76, 113, 160], and boost engagement [59]. The goal is to enable the most common visual data exploration tasks where simple visual inspection is not enough. Following Amar et al.’s taxonomy of analytic tasks, this includes filtering to find data cases matching certain conditions, computing derived values such as means or counts, and sorting to rank data cases [4]. Annotations in the text are synchronized with the corresponding chart and both text and chart representations are interactive, following recommendations from our characterization (Section 4).

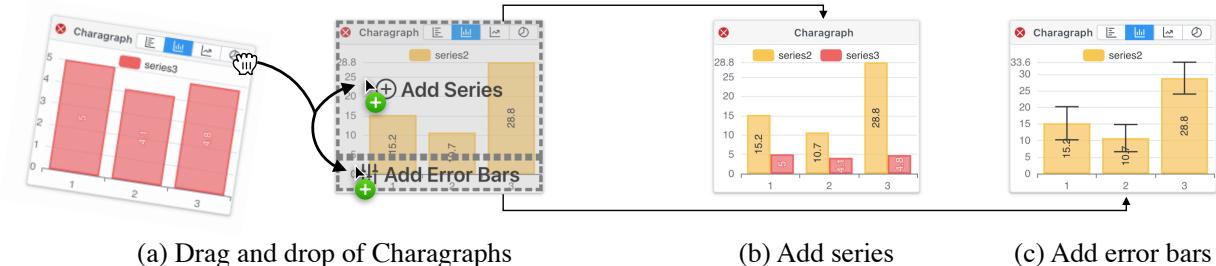


Figure 7: Combining information in two Charagraphs: (a) dragging a Charagraph onto a drop zone in another Charagraph combines data in different ways; (b) adding a new series; or (c) adding error bars.

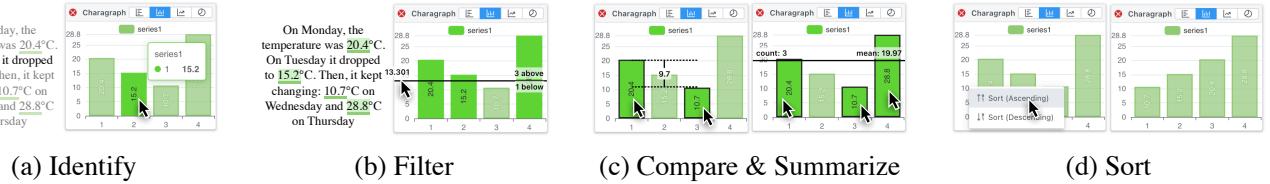


Figure 8: Text and chart interactions: (a) hovering above numbers in text or chart elements identifies connections; (b) moving the cursor on the Y axis controls a threshold to filter numbers; (c) clicking two values compare them and clicking more than 2 values calculates the mean; (d) right-clicking opens a contextual menu to sort.

5.4.1 Identify. Readers can hover over a value in the text or a data point in the chart to see the value highlighted in the text and the chart. The entire document is dimmed except for the sentence containing the value to highlight the context in which the value was reported (Figure 8a). This highlighting interaction helps connect the values across the text and chart [80]. For example, even with unlabelled charts, the reader can determine the highest value by locating the highest bar in the chart, then pointing at it to locate it in text and find the corresponding label (Figure 8a).

5.4.2 Filter. Finding values above or below a threshold is done in the chart by moving the cursor over the Y-axis (X-axis for horizontal bars). A horizontal line is shown at the location of the cursor indicating how many values are above or below. Values above the threshold are highlighted in both the text and the chart (Figure 8b).

5.4.3 Compare and Summarize. Values in the text and in the chart can be clicked to remain selected. Derived values are calculated from the selection: if two values are selected, their difference is shown in the chart (Figure 8c left); if three or more values are selected, the mean is calculated with a bar displayed at the corresponding position (Figure 8c right).

5.4.4 Sort. Values follow the order of appearance in the text by default, but can be sorted by ascending and descending order by right-clicking the chart and selecting the corresponding option in a context menu (Figure 8d). The sorting operation is animated to help readers understand and locate elements [55].

5.5 Implementation

We implemented the system using TypeScript and React [121] for the interface, PrimeReact [116] for the library of graphical components, and Apache ECharts [35, 84] for the visualizations. Everything is implemented inside the PDF.js [111] PDF viewer, the default PDF reader used by Mozilla Firefox. Our system works with all PDF documents and can be used as a general PDF viewer. The full source code is hosted online: <http://ns.inria.fr/loki/charagraph>.

In this section, we provide details for the two main components required to create a Charagraph: the selection of text and the generation and extraction of data from text.

5.5.1 Structural Text Selection with PDFs. Because the PDF format is focused on visual display and small file size, the format stores text at a character-level layout with no formal representation of sentences and paragraphs¹. PDF viewers traditionally reconstruct

sentences based on the order of characters and their position on the page: two characters may be merged based on arbitrary distance thresholds. In PDF.js, the reconstructed sentences are invisibly overlaid on the PDF using a collection of `` tags. The web browser then handles text selection using the invisible text layer. This approach has two main drawbacks: (1) Sentence reconstruction is imperfect and words may be improperly merged while others may be merged when they should not; and (2) the web browser maps user input to a list of `span` tags with no structural knowledge of the PDF. This causes unexpected and spurious selections.

Text selection was re-implemented inside the system to be consistent and support marquee selection. The algorithm relies on a geometric and semantic representation of the document: The geometric representation is obtained by modifying PDF.js to return the precise location and size of the text on the page. The semantic representation is generated from the text reconstructed by PDF.js. The representations are aligned so that a shape maps to a precise index in the text, and vice versa. This allows the marquee selection to function because individual characters are precisely located and selected only if included within a rectangle. The traditional text selection is based on the semantic representation and reproduces typical behaviour.

Algorithm 1: Generation of data-group suggestions based on prefixes

```

Input: A text portion  $t$  and the maximum length  $m$  of the prefixes
Output: A list of data-group suggestions
numbers  $\leftarrow$  substrings within  $t$  matching
 $/[-+]?[0-9]+[.]?[0-9]^*([eE][-+]?[0-9]+)?/$ 
groups  $\leftarrow$  [group formed from numbers]
foreach  $n \in \text{numbers}$  do
    for  $l \leftarrow 0$  to  $m$  do
         $pg \leftarrow \{\}$ 
         $p \leftarrow$  substring of length  $l$  preceding  $d$ 
        if  $p$  does not match  $\backslash s^*[\backslash .\backslash (:)\backslash s]\backslash s^*$  then
            foreach  $c \in \text{numbers}$  do
                if  $c$  is prefixed by  $p$  then  $pg.add(c)$ 
                if  $pg \notin \text{groups}$  then  $\text{groups}.add(pg)$ 
    return  $\text{groups}$  sorted by size

```

¹PDF can technically be more structured using tags, often for accessibility purposes (see ISO 32000-1:2008§14.8). However, very few PDFs contain those tags in practise.

5.5.2 Data-Group Suggestion. The suggestion algorithm works by identifying numbers in text with a regular expression and then grouping these values based on their suffix and prefix. Because the algorithm forms all possible groups, the list is pruned. Groups that are identical in the values that they select are merged. And groups with prefix and suffix that are empty spaces or punctuation marks are excluded. The resulting list is sorted from largest to smallest group and used as suggestions. The pseudocode of the algorithm to form groups based on prefixes is provided in Algorithm 1 (the algorithm for suffixes is identical except it matches the text following the numbers). In the system, we set the maximum length of prefixes and suffixes to 20.

6 USER STUDY

It is generally accepted that charts are better at communicating trends [99], proportions [139], uncertainty, and effect sizes [30, 39, 53] all while making it easier to interpolate [19], predict [99], and perform general interpretive operations such as locating, deriving, or comparing values [29, 37]. In contrast, text is preferred for compact representations and to extract exact numerical values.

However, these findings were obtained with charts carefully designed. It is unclear whether these findings apply to interactive charts generated from and read alongside data-rich texts in the wild. And it is unclear whether the benefits of showing the same information in two different forms outweigh the cognitive costs.

We conducted a two-part user study: part one evaluates the *usability* of the system and its interaction model (*Can people use our system to create Charagraphs?*); part two evaluates the *utility* of Charagraphs (*Do people benefit from having Charagraphs compared to only text?*). Both parts isolate the aspect being evaluated. For the sake of clarity, the two parts are presented in different sections, although participants did both in the same session.

6.1 Participants

We recruited 12 participants (23 to 36 age range, mean = 28.9, 7 identified as female, 5 as male) from a local university. All participants were graduate students or researchers from the fields of Computer Science or Medical Science. On a 5-point scale, they all reported frequently reading documents that include in-text numbers such as scientific or news articles (Mdn=5, SD=0.8) and that they were familiar with common descriptive statistics such as standard deviations and p-values (Mdn=4, SD=0.6). In appreciation for their time, participants received a compensation of \$15 CAD.

6.2 Apparatus

Participants took part in the study remotely from their personal computer. The experiment was hosted online and participants shared their screen with the experimenter. The interaction of the participants with the website (e.g., clicks, movements, time, answers to questions), a screen capture, and the microphone were recorded.

7 STUDY PART 1: USABILITY

Participants were asked to create Charagraphs to help answer four questions about data presented in a text excerpt. Since we are examining usability of the Charagraph system, the questions serve as

goals to motivate the use Charagraphs. The four questions represented two common tasks (two questions per task) based on taxonomies of tasks related to information visualizations [4, 16]. These questions can be solved easily and quickly, but require an understanding of the visualization. Each question is multiple-choice with up to eight choices. For example, a ‘Filter’ task was expressed in a question like “*How many are above/below [number]?*” and a ‘Find Extremum’ task expressed like “*What is the maximum/minimum*” or “*What is the [number]th largest/smallest?*”.

7.1 Procedure

Introduction and Tutorial (10 mins) – After completing a demographic questionnaire, participants were informed that they would “*use a system to visualize the data presented in-text*”. Participants went through a fragmented tutorial in which they reviewed a video-tutorial and then immediately practised what they just saw, but on a different text. There were 4 steps to the tutorial: 1) selecting and extracting numbers from the text to create a Charagraph; 2) adding error bars; 3) adding series; and 4) interacting with the Charagraph. The text used in the video was created for the study and presented the life expectancy for different continents in 2000 (in a paragraph) and in 2015 (in a table). Participants practised on a similar fictitious paragraph and table presenting the evolution of the temperature at two different periods. The experimenter could intervene if needed.

Usability Task (15 mins) – Participants were presented a text excerpt and asked to “*Create a visualization to help you answer questions about [some data presented in the text]*”. The experimenter did not intervene and participants were free to create any Charagraph. The questions were presented after the participants created at least one Charagraph and pressed “Next”. Both the paragraph and the created Charagraphs remained visible during question-answering. This phase was repeated for the three different text excerpts, always in the same order. At the end, participants completed a System Usability Scale (SUS) [18] and rated statements on a 5-point scale.

7.2 Text Excerpts

We extracted three text excerpts from different sources. All excerpts did not have an associated visualization in their original document and are therefore concrete use cases for Charagraphs. The excerpts were given to participants unchanged and also included other irrelevant data to the task. It is the participants’ responsibility to extract only the information of interest to create a Charagraph.

- The **SURVEY** excerpt is a paragraph extracted from a report published by the United Nations Educational, Scientific and Cultural Organization (UNESCO) about the most dominant languages in films produced in 2009 [1]. It is comprised of 97 words and reports 13 different numbers, of which only 5 (the number of films for different languages) are of interest to complete the task.
- The **SCIENTIFIC** excerpt is a paragraph extracted from a paper published at the ACM CHI 2022 conference [83] reporting the response to a Likert questionnaire. Both means and standard deviations are reported in APA style. The paragraph is 74 words long and contains 12 numbers, of which 10 are of interest (5 means and their corresponding standard deviations).
- The **WEEKLY REPORTS** excerpt was formed by merging two consecutive weekly reports about the number of cases of COVID

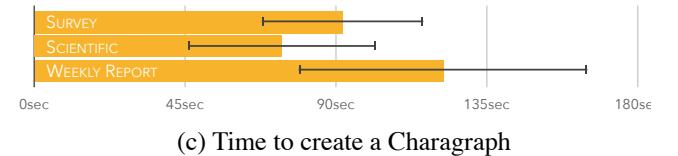
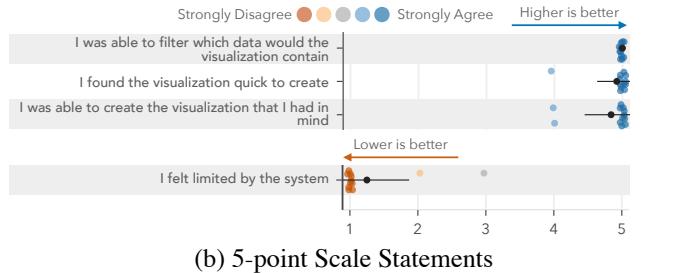
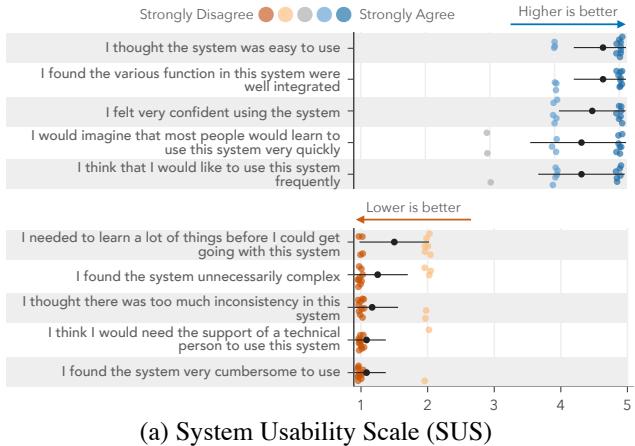


Figure 9: Result summary of the usability part of the study. Mean and standard deviation of participants' answers to the SUS and the statements on a 5-point scale. Average time to create a Charagraph.

published by the World Health Organization [106, 107]. The excerpt consists of 218 words and 31 numbers. Only six numbers are relevant to the task (the number of cases for three countries in the first week, and their corresponding number of cases in the following week).

The three excerpts were chosen to cover potential use cases for Charagraphs: show a distribution and compare different values (SURVEY), visualize means and uncertainty or variability (SCIENTIFIC), and visualize the evolution or trend (WEEKLY REPORTS).

7.3 Results

All participants were able to use our system to create Charagraphs. Additionally, they correctly answered all the stimuli questions about the text excerpts. Figure 9 summarizes the results.

7.3.1 Time. The time it took participants to read the text and create Charagraphs ranged from 36s to 3min 13s ($M=1\text{min } 36\text{s}$, $SD=38\text{s}$). Participants were the fastest with SCIENTIFIC ($M=1\text{min } 14\text{s}$, $SD=27\text{s}$), followed by SURVEY ($M=1\text{min } 32\text{s}$, $SD=26\text{s}$) and WEEKLY REPORTS ($M=2\text{min } 2\text{s}$, $SD=42\text{s}$).

7.3.2 Questionnaire. The average SUS usability score was 81² ($Mdn=81.2$, $SD=5.2$) and the participants felt that they could use the system quickly and successfully: on a 5-point scale (1-strongly disagree, 5-strongly agree), they rated that they were able to create the visualization that they had in mind ($Mdn=5$, $SD=0.4$, $\square\square\square\square\square$), they did not feel limited by the system ($Mdn=1$, $SD=0.6$, $\square\square\square\square\square$), they were able to filter which data to extract and visualize ($Mdn=5$, $SD=0.4$, $\square\square\square\square\square$), and they found the visualization quick to create ($Mdn=5$, $SD=0.3$, $\square\square\square\square\square$).

7.3.3 How Charagraphs Were Created. Participants took 55s on average to read the text before initiating the creation of their first Charagraph. Regular text selection was used 72% of the time, and all participants relied on the automatic suggestions rather than selecting the values manually. Participants were also mostly right the first time: they deleted only 9% of the Charagraphs they created.

²A SUS score above 80 is considered 'Good' [8]

In terms of type of visualization selected, most participants stuck to the default bar chart, adding error bars for SURVEY (11/12) and another series for WEEKLY REPORTS (11/12). Only one participant used a line chart for WEEKLY REPORTS, and another participant used a pie chart for SURVEY although they switched to a bar chart after seeing the questions.

8 STUDY PART 2: UTILITY

In this second part of the study, we focused on the utility of Charagraph. Because, we are interested by the quality of the answers given by participants, we use a broader range of tasks and questions that are more challenging to answer than in the usability task. We eliminate the usability component to keep the study focused on utility. All results are compared against a baseline text-only condition. Note that this part immediately followed the usability part.

8.1 Text Excerpts

We extracted three types of excerpts covering typical use cases for Charagraphs. For each type, we extracted two excerpts from different sources resulting in six excerpts in total.

- **PROPORTION** excerpts were obtained from two sources: 1) a news website reporting on the proportion of White Americans eligible voters per state during the 2020 presidential election [52]; and 2) the U.S. Bureau of Labor Statistics reporting on the unemployment rate among the major worker groups [151]. Both excerpts report seven different percentages.
- **UNCERTAIN** paragraphs were obtained from two papers published at CHI [91, 145]. They report the results for 6 and 7 items to a questionnaire using strict APA and APA-like styles resulting in 12 and 14 numbers for both mean and standard deviation.
- **EVOLUTION** paragraphs were obtained from two articles published by the UNESCO. One reported the evolution of the adult literacy rate between 1990 and 2016 for 5 regions, resulting in 10 numbers [150]. The other reported on the evolution of the production of films between 2005 and 2009 for five countries, resulting in 10 numbers [1].

The excerpts were given to participants unaltered (retaining style, wording, and spacing).

8.2 Conditions

This part of the study had two conditions:

- **BASELINE:** the text excerpts are loaded in a traditional PDF reader with typical features such as text selection, zooming, and keyword search (CTRL-F).
- **CHARAGRAPH:** the condition is identical to BASELINE except the PDF reader allows the creation of Charagraphs (using the system described in Section 5). Because this part of the study does not evaluate the usability of our system, and in order to guarantee that the Charagraphs would be useful to answer questions, the Charagraphs were designed prior to the study and participants were instructed to replicate them. We still chose to make participants create the Charagraph instead of just automatically generating it; we argue that the act of *creating* the Charagraph is part of the understanding. Additionally, we make participants use Charagraphs without labels. Our motivation is twofold: first, we believe that labels are optional because the information can be obtained through the interactivity and participants explicitly selected the data to plot, thus they already know what it represents. Second, with labels, most of the questions can be answered solely using the chart. While this is one of the motivations for Charagraphs, several studies have already shown the superiority of graphical representations over text [3, 26, 44, 67, 153]. By removing labels, we ensure that participants will go back-and-forth between graphical and textual representation, which is a more challenging and realistic scenario.

8.3 Test Questions

For each text excerpt, we designed six questions that could be answered only by reading the text. The questions were chosen to represent the most common data analysis activities [4, 16, 119]:

- **RETRIEVE:** “What is the value for [item]?”
- **FILTER:** “How many [items] are above/below [threshold]?”
- **DERIVE:** “What is the difference between [item] and [item]?”
- **EXTREMUM:** “What is the [i]th largest/smallest?”
- Specific questions depending on the type of excerpts.
 - **SIMILARITY** questions for PROPORTION excerpts: “What are the two most similar/dissimilar [items]?”
 - **VARIABILITY** questions for UNCERTAIN excerpts: “What is the [item] with the largest/smallest variability?”.
 - **TREND** questions for EVOLUTION excerpts: “What [item] had the largest/smallest increase in [value] between [year] and [year]?”.

All questions were pilot-tested before running the experiment to make sure they were clear and understandable. For each question, the participants were shown all possible answers or a list of eight possible answers (for DERIVE and FILTER) in a random order.

8.4 Procedure

Utility Task (25 mins) – Participants were presented with a text excerpt and, after pressing “Next”, had to answer six questions about some of the data presented in-text. The questions were shown below the PDF reader and participants were given no time limit to read and answer the questions. Participants went through both

conditions and all six excerpts, although in different orders. After each condition, participants completed a raw NASA-TLX and rated statements on a 5-point scale (1-strongly disagree, 5-strongly agree).

- In the **BASELINE** condition, when the text was first presented (and before answering questions) participants were prompted to “*Read the paragraph and press ‘Next’*”. The text remained visible during the question-answering.
- In the **CHARAGRAPH** condition, participants had to follow instructions to create a specific Charagraph: The instructions were only shown when the text was first presented (before answering questions) and the experimenter made sure that the participant created the expected Charagraph. To help participants understand what the visualization represented, the instructions were preceded by a brief explanation such as “*Visualize the number of films produced in 2009*”. The Charagraph, along with the text, remained visible during the question-answering and participants were free to interact with the Charagraph or not.

Semi-Structured Interview (10 mins) – At the end of the session, the experimenter conducted a semi-structured interview. Participants were given the opportunity to comment on any aspect of the study or the system. The experimenter also initiated discussions about the participants’ preferred condition, the aspects of the system they found most useful, their strategies, if they would use such a system, in what context, and on what kinds of documents.

8.5 Ordering

Text excerpts and their associated questions were presented with an order following a balanced Latin square. We ensured that 1) each participant saw each text excerpt exactly once; 2) each condition had exactly one text excerpt of each type; and 3) a text excerpt appeared exactly once at every possible position every six participants. As such, half of the participants saw the same text excerpt in the **BASELINE** condition, while the other half saw it in the **CHARAGRAPH** condition. Additionally, we balanced the order of the conditions: half of the participants started with the **BASELINE** condition while the other half started with the **CHARAGRAPH** condition.

8.6 Results

We calculate p-values using Student’s t-test when the values are normally distributed (time) and a Wilcoxon Signed-Rank test otherwise (correctness, workload, and 5-point scales) using the stats module from `scipy` [152].

8.6.1 Correctness. Wilcoxon Signed-Rank found a significant effect of the condition on correctness ($p=.003$, $z=-2.98$). Participants were more correct at answering questions with **CHARAGRAPH** ($M=98.61$, $SD=11.7$) than **BASELINE** ($M=92.59$, $SD=26.19$). In fact, aside from three participants who correctly answered all questions regardless of the condition, the other nine participants had consistently higher accuracy with **CHARAGRAPH** than **BASELINE**.

8.6.2 Time. The t-test did not find a significant effect of the condition on time ($p=.15$, $t=1.46$). Participants spent a similar amount of time to answer questions with **CHARAGRAPH** ($M=1\text{min } 38\text{s}$, $SD=55\text{s}$) than **BASELINE** ($M=1\text{min } 28\text{s}$, $SD=1\text{min } 13\text{s}$).

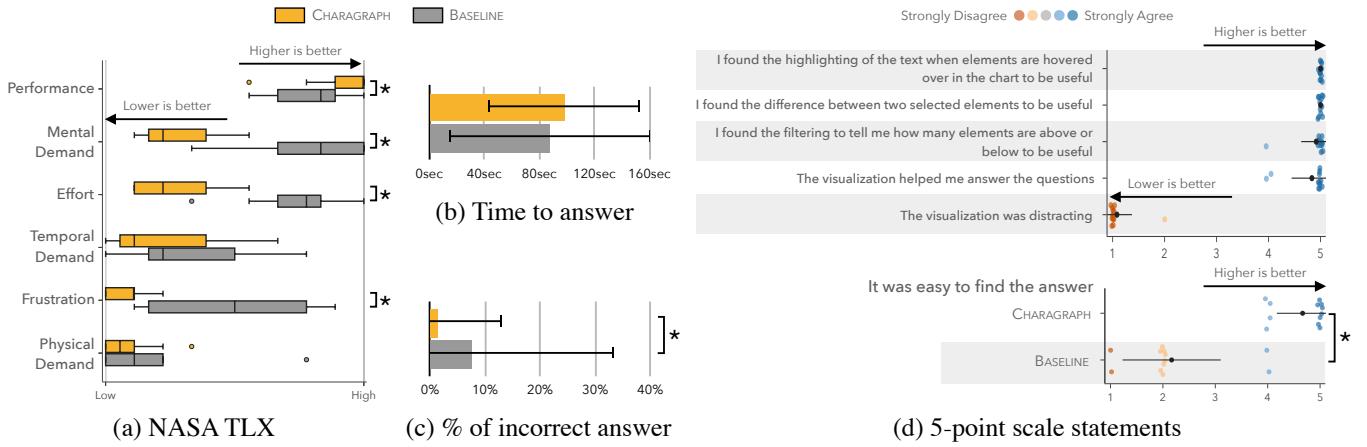


Figure 10: Results to (a) the raw NASA-TLX; (b) the average time to answer; (c) the average percent of incorrect answers; and (d) the 5-point scale statements. Error bars represented standard deviation. Statistical significance represented by a “*”

8.6.3 Workload (NASA-TLX). Wilcoxon Signed-Rank found a significant effect of condition on mental demand ($p<.003$, $z=-2.95$), frustration ($p=.007$, $z=-2.67$), effort ($p=.002$, $z=-3.09$) and performance ($p=.03$, $z=-2.14$). No significant effect for temporal ($p=.17$, $z=-1.38$) and physical demand ($p=.07$, $z=-1.84$). Compared to BASELINE, CHARAPGRAPH had lower mental demand ($M=7$ vs. $M=16.2$), frustration ($M=3.7$ vs. $M=10.8$), effort ($M=6.7$ vs. $M=15.5$), and higher performance ($M=18.5$ vs. $M=16.5$).

8.6.4 Questionnaire. All 12 participants preferred CHARAPGRAPH over BASELINE. They found the answers easier to find with CHARAPGRAPH ($Mdn=5$, $SD=0.5$, █) compared to BASELINE ($Mdn=2$, $SD=0.9$, █) and the difference was significant ($p=.003$, $z=-2.97$). Importantly, participants did not find Charagraphs to be distracting ($Mdn=1$, $SD=0.3$, █) and instead all agreed that it helped them answer the questions ($Mdn=5$, $SD=0.4$, █).

8.6.5 Interactions Used. On average, participants hovered over the chart while answering 83% of the questions, and over the text in 44% of the questions. Counting the number of interactions in each context, clicking a value in the chart was the top interaction when answering EXTREMUM (78%), SIMILARITY (59%), VARIABILITY (65%), and TREND (75%) questions. In contrast, clicking values in-text was preferred for DERIVE questions (64%). The filter interaction was the top interaction only for FILTER questions (100%). RETRIEVE questions were most often answered without interacting with the text (3%) nor chart (11%).

9 DISCUSSION

We discuss the results by first answering our two research questions and then open the discussion to other themes that stem from observations and participants' comments.

People can use our system to create Charagraphs. All participants were able to create Charagraphs and use their main features to answer questions. Participants even customized the Charagraphs to their preferences: some participants kept visualizations separate

because they disliked error bars, changed their style, and renamed ticks and legends.

People benefit from having Charagraphs compared to only text. Participants answered the questions more correctly with Charagraphs (98.61%) compared to text only (92.59%) and used and found useful all the main features of Charagraphs. All of them preferred Charagraphs and would like to use the system again. This was also reflected in the questionnaires. Participants had lower mental demand, effort, and frustration with Charagraph than when using text alone. Although harder to measure, participants' reactions suggest that Charagraphs benefit other areas. For example, despite having already read the text, P6 commented “Wow! It's them that produce the most [films]? It's mind-blowing” right after creating a Charagraph. This suggests the potential of Charagraphs to discover facts that would have been missed if only the text had been read.

9.1 Observations and Participants' Comments

Charagraphs might change reading behaviours. Two participants mentioned that knowing that they could create Charagraphs changed how they read the text: P9 “When I know I can create a visualization [...] it's kind of like I already have an imagined figure in my mind when I read a paragraph” and P4 “It was me thinking about oh, this information would be useful for my graph [...] so I did this kind of filtering while I was reading, so I was thinking a lot more about what I was reading”. Researchers have suggested that the benefit of active reading might not be in the marks left on paper but on the process of creating those marks [2, 68, 125, 143, 157]. These comments suggest a similar effect with Charagraphs.

Labels are not necessary but would have made answering questions easier and faster. When designing the utility part of the study, we decided against having participants add labels to Charagraphs. The idea was to simulate a perhaps more realistic scenario in which a user would not have to bother adding labels because they are not absolutely necessary. We also thought that the alternative (with labels) was too powerful considering that the questions become answerable just from the Charagraph, without looking at the text.

Participants commented that not having labels was not an issue because of the interactive highlight linking text and chart. P3 “*The [missing] labels... I don't think it bothered me. Because the highlighting will just link the data with the text. Because the data is very close to the text*”. However, some participants also commented that they would have preferred having labels: P6 “*If I had labelled the legends and ticks, I'd have been even faster*”; P8 “[labels] would have been easier than referring to the text”; P9 “*if I can directly find the answer in the visualization that would be better*”.

Charagraphs might help even when data is already visualized. Readers may have preferences that might not be met by canonical visualizations [82, 101]. Even in our controlled scenarios during the usability task, participants expressed different preferences: 2 participants preferred having two separate visualizations for means and standard deviation instead of a combined representation with error bars. P9 explained “*for me the bar chart is easier to read than error bars [because error bars] do not start at 0*”. If not preferences, a participant also mentioned that existing visualizations might be overblown, P12 “*Even if they [documents] do have graphs, they don't have the specific information that you are looking for... so if you want, like, focus on a specific thing like the mean for example*”. Similarly, an existing static visualization might be deceptive because it has truncated axes, or because the aspect ratio exaggerates a trend [97, 109]. Or the visualization has accessibility issues because of its colour scheme and style [69, 156]. In these cases, as long as the data is presented in text, a Charagraph can recreate the static visualization, but with the possibility of tweaking its scale and style. The resulting “reader-created” visualizations can be used to compare or replace the static canonical visualizations already present in the document.

9.2 Limitations

Some in-text data might be difficult or impossible to extract. The suggestion algorithm forms groups based on the text before and after numbers. This requires numbers to share prefixes (e.g., M=, SD=) or suffixes (e.g., %, ms, cm) or sentences to follow similar structures with identical prepositions before numbers. When this assumption fails, users have to resort to adding missing values manually. Additionally, data other than numbers and certain sentence constructions are currently unsupported. For example, a sentence might refer to a quantity mentioned previously, such as “half of the participants”. Lastly, combining two Charagraphs or adding error bars relies on the order of the data points. As such, this requires the data to be presented in text in the same order or to manually re-organize the data in the Charagraphs so that they match.

The instructions in the utility study might have made participants less engaged with the text. In the utility part of the study, we gave participants instructions to follow to create a Charagraph. This was done to remove the usability component (which had already been tested in the previous part). Several participants mentioned that this made them less engaged with the text: P4 “*I definitely noticed, with the instruction I was not reading the text [...] I think I actually preferred the sort of exploratory mode [usability part] where I could just make the graphs I wanted without the instructions*” and P5 “*I didn't like having the instructions [...] I think, in the last condition [baseline with only text] and in the condition without instructions [usability part] I read the text more*”.

Participants were well-versed in data literacy. All the participants in our study were graduate students and researchers trained in reading scientific documents and interpreting numeric values. Thus, participants could understand the content with text only. Given findings that charts help casual readers understand numeric data [31, 114], it is plausible that people who are less versed in reading scientific documents might benefit from Charagraphs as long as they are familiar with data visualizations.

9.3 Future Work

Exploring the design space of Charagraphs. In our characterization (Section 4), we mentioned two other level of integration that are currently not supported by our system: chart in text using word-scale visualizations [47, 149] and text in chart using annotated charts [122]. Our study focused on separate representations that let readers choose the one that is most appropriate. It would be interesting to observe the impact of the two other modes on reader's behaviours. Similarly, we believe there is an opportunity to extract the information contained within the text using NLP to guide the generation of Charagraphs and give better defaults. From our characterization (Section 4), statistical-terms and units could be extracted to design embellished charts that are more memorable [9, 14]; labels could be recovered to annotate charts and add names to legends, ticks and axes; and the semantic could decide the visualization.

Support for advanced Natural Language Processing (NLP). Adding NLP has the potential of making the creation of Charagraphs easier and quicker. First, the suggestion algorithm could rely on named entity recognition (NER) to filter numbers that are not data points such as references to sections, dates and figures. Semantically related groups of numbers could be formed by analyzing the sentence structure and content which might result in better suggestions. Second, the extraction of data points could be made more robust by understanding common natural language structures. For example, to understand “half of the participants”, a system would need to keep track of the mentioned values (here, the number of participants), understand the described arithmetic (half of), and calculate the final number. Third, NLP could extract meta-information about the data, such as p-values to automatically decorate the visualization with the paired significance level. Last, NLP could automatize the extraction of labels by parsing the relation between the number and the label. Such a system could draw from approaches to parse textual analogies [36, 77]. However, these remain challenging even with modern techniques (see the difficulties faced by approaches to automatically connect text and chart [76, 113]).

Support for other visualizations and tasks. In addition to the common bar, line, and pie charts already supported, the system could be extended to include scatter plots, choropleth maps, box plots, and radar charts, which are relatively common [10]. The challenge is whether required data can be mined or calculated from the text. For example, a scatter plot is designed to show raw data points, but text typically reports on aggregated values representing raw data (e.g., means and standard deviations). A scatter plot also requires data in two-dimensions (e.g. x and y), which may not be reported in-text. Regarding other tasks, a Charagraph often represents a subset of data, or a combination of some data of interest. This suggests other tasks for a reader, such as to characterize the distribution of the

selected data, find correlations, clusters, or anomalies. These tasks could be made easier by automatically identifying data facts and remarkable features [146] of the generated Charagraphs, and offer suggestions or annotations [76, 140].

Evaluating the impact of Charagraphs on information retention and engagement. Other benefits of graphical representations over text might transfer to Charagraphs. For example, previous work has shown better recall of information [15, 58] and greater reader engagement [96] when using data visualization compared to text alone. A user study that specifically evaluates these aspects could further inform us on the benefits of Charagraphs.

Charagraphs with printed text and physical documents. The system currently works with digital documents but could be extended to the physical world. Using a smartphone, readers could take a photo of a document containing number in the text. The system would process this photo using Optical Character Recognition (OCR), offer similar selection features to extract the values, and generate a Charagraph on the smartphone. Future work could investigate the design and implementation of such a system or a more advanced version that would overlay Charagraphs on the physical document, similar to DuallyNoted that uses augmented reality on a smartphone to annotate paper documents [118].

10 CONCLUSION

Documents seldom include visualizations despite the advantages of graphics compared to textual representations. To help readers when visualizations are missing or poorly designed, we introduced Charagraphs, dynamically generated interactive charts and annotations for in-situ visualization, comparison, and manipulation of data included within text. We also presented a document viewer to create Charagraphs in existing PDF documents. In a user study, participants could quickly create their own imagined visualizations from data in text and were more correct when answering questions using Charagraphs compared to only reading text. Charagraphs provide an immediate solution for readers wanting to visualize and manipulate numeric data in existing documents, and explore the concept of reader-generated visualizations.

ACKNOWLEDGMENTS

This research received ethics clearance from the Office of Research Ethics, University of Waterloo. This work was made possible by NSERC Discovery Grant 2018-05187 and the LAI Réapp.

REFERENCES

- [1] Acland Acland. 2012. From International Blockbusters to National Hits: Analysis of the 2010 UIS Survey on Feature Film Statistics - UNESCO Digital Library. <https://unesdoc.unesco.org/ark:/48223/pf0000217103>.
- [2] Mortimer Jerome Adler and Charles Van Doren. 2014. *How to Read a Book* (touchstone hardcover edition ed.). Simon & Schuster, New York.
- [3] D. G. Altman and J. M. Bland. 1996. Presentation of Numerical Data. *BMJ (Clinical research ed.)* 312, 7030 (March 1996), 572. <https://doi.org/10.1136/bmj.312.7030.572>
- [4] R. Amar, J. Eagan, and J. Stasko. 2005. Low-Level Components of Analytic Activity in Information Visualization. In *IEEE Symposium on Information Visualization, 2005. INFOVIS 2005*. IEEE, Minneapolis, MN, USA, 111–117. <https://doi.org/10.1109/INFVIS.2005.1532136>
- [5] American Psychologic Association. 2020. *Publication Manual of the American Psychological Association* (6th edition ed.). American Psychological Association, Washington, DC.
- [6] Paul Ayres and Gabriele Cierniak. 2012. Split-Attention Effect. In *Encyclopedia of the Sciences of Learning*, Norbert M. Seel (Ed.). Springer US, Boston, MA, 3172–3175. https://doi.org/10.1007/978-1-4419-1428-6_19
- [7] Sriram Karthik Badam, Zhicheng Liu, and Niklas Elmqvist. 2019. Elastic Documents: Coupling Text and Tables through Contextual Visualizations for Enhanced Document Reading. *IEEE Transactions on Visualization and Computer Graphics* 25, 1 (Jan. 2019), 661–671. <https://doi.org/10.1109/TVCG.2018.2865119>
- [8] Aaron Bangor, Philip T. Kortum, and James T. Miller. 2008. An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction* 24, 6 (July 2008), 574–594. <https://doi.org/10.1080/10447310802205776>
- [9] Scott Bateman, Regan L. Mandryk, Carl Gutwin, Aaron Genest, David McDine, and Christopher Brooks. 2010. Useful Junk?: The Effects of Visual Embellishment on Comprehension and Memorability of Charts. In *Proceedings of the 28th International Conference on Human Factors in Computing Systems - CHI '10*. ACM Press, Atlanta, Georgia, USA, 2573. <https://doi.org/10.1145/1753326.1753716>
- [10] Leilani Battle, Peitong Duan, Zachery Miranda, Dana Mukusheva, Remco Chang, and Michael Stonebraker. 2018. Beagle: Automated Extraction and Interpretation of Visualizations from the Web. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–8. <https://doi.org/10.1145/3173574.3174168>
- [11] Aaron Bauer and Kenneth R. Koedinger. 2008. Note-Taking, Selecting, and Choice: Designing Interfaces That Encourage Smaller Selections. In *Proceedings of the 8th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL '08)*. Association for Computing Machinery, New York, NY, USA, 397–406. <https://doi.org/10.1145/1378889.1378961>
- [12] Fabian Beck and Daniel Weiskopf. 2017. Word-Sized Graphics for Scientific Texts. *IEEE Transactions on Visualization and Computer Graphics* 23, 6 (June 2017), 1576–1587. <https://doi.org/10.1109/TVCG.2017.2674958>
- [13] Jacques Bertin. 2013. *Sémiologie graphique: Diagrammes, les réseaux, les cartes*. EHESS, Paris.
- [14] Michelle A. Borkin, Zoya Bylinskii, Nam Wook Kim, Constance May Bainbridge, Chelsea S. Yeh, Daniel Borkin, Hanspeter Pfister, and Aude Oliva. 2016. Beyond Memorability: Visualization Recognition and Recall. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (Jan. 2016), 519–528. <https://doi.org/10.1109/TVCG.2015.2467732>
- [15] Michelle A. Borkin, Azalea A. Vo, Zoya Bylinskii, Phillip Isola, Shashank Sunkavalli, Aude Oliva, and Hanspeter Pfister. 2013. What Makes a Visualization Memorable? *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (Dec. 2013), 2306–2315. <https://doi.org/10.1109/TVCG.2013.234>
- [16] Matthew Brehmer and Tamara Munzner. 2013. A Multi-Level Typology of Abstract Visualization Tasks. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (Dec. 2013), 2376–2385. <https://doi.org/10.1109/TVCG.2013.124>
- [17] Victor Bret. 2011. Explorable Explanations. <http://worrydream.com/ExplorableExplanations/>.
- [18] John Brooke. 1995. SUS: A Quick and Dirty Usability Scale. *Usability Eval. Ind.* 189 (Nov. 1995), 7.
- [19] L. F. Carter. 1947. An Experiment on the Design of Tables and Graphs Used for Presenting Numerical Data. *The Journal of Applied Psychology* 31, 6 (Dec. 1947), 640–650. <https://doi.org/10.1037/h0054246>
- [20] Olivier Chapuis and Nicolas Roussel. 2007. Copy-and-Paste between Overlapping Windows. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 201–210. <https://doi.org/10.1145/1240624.1240657>
- [21] Zhihui Chen and Hajun Xia. 2022. CrossData: Leveraging Text-Data Connections for Authoring Data Documents. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–15. <https://doi.org/10.1145/3491102.3517485>
- [22] Ed H. Chi, Lichan Hong, Michelle Gumbrecht, and Stuart K. Card. 2005. SceneHighlights: Highlighting Conceptually-Related Sentences during Reading. In *Proceedings of the 10th International Conference on Intelligent User Interfaces (IUI '05)*. Association for Computing Machinery, New York, NY, USA, 272–274. <https://doi.org/10.1145/1040830.1040895>
- [23] William S. Cleveland. 2004. *The Elements of Graphing Data* (2 edition ed.). Hobart Press, Murray Hill, NJ.
- [24] Matthew Conlen and Jeffrey Heer. 2018. Idyll: A Markup Language for Authoring and Publishing Interactive Articles on the Web. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*. ACM, Berlin Germany, 977–989. <https://doi.org/10.1145/3242587.3242600>
- [25] Matthew Conlen, Alex Kale, and Jeffrey Heer. 2019. Capture & Analysis of Active Reading Behaviors for Interactive Articles on the Web. *Computer Graphics Forum* 38, 3 (2019), 687–698. <https://doi.org/10.1111/cgf.13720>
- [26] Patricia Costigan-Eaves. 1986. Edward R. Tufte *The Visual Display of Quantitative Information*. *Information Design Journal* 4, 3 (Jan. 1986), 235–236. <https://doi.org/10.1075/ijd.4.3.12cos>
- [27] Weiwei Cui, Xiaoyu Zhang, Yun Wang, He Huang, Bei Chen, Lei Fang, Haidong Zhang, Jian-Guan Lou, and Dongmei Zhang. 2020. Text-to-Viz: Automatic Generation of Infographics from Proportion-Related Natural Language Statements. *IEEE Transactions on Visualization and Computer Graphics* 26, 1 (Jan.

- 2020), 906–916. <https://doi.org/10.1109/TVCG.2019.2934785>
- [28] Zhe Cui, Sriram Karthik Badam, M Adil Yalçın, and Niklas Elmqvist. 2019. DataSite: Proactive Visual Data Exploration with Computation of Insight-Based Recommendations. *Information Visualization* 18, 2 (April 2019), 251–267. <https://doi.org/10.1177/1473871618806555>
- [29] Hugh M. Culbertson and Richard D. Powers. 1959. A Study of Graph Comprehension Difficulties. *Audio Visual Communication Review* 7, 2 (1959), 97–110.
- [30] Pierre Dragicevic. 2016. Fair Statistical Communication in HCI. In *Modern Statistical Methods for HCI*, Judy Robertson and Maurits Kaptein (Eds.). Springer International Publishing, Cham, 291–330. https://doi.org/10.1007/978-3-319-26633-6_13
- [31] Pierre Dragicevic and Yvonne Jansen. 2018. Blinded with Science or Informed by Charts? A Replication Study. *IEEE Transactions on Visualization and Computer Graphics* 24, 1 (Jan. 2018), 781–790. <https://doi.org/10.1109/TVCG.2017.2744298>
- [32] Pierre Dragicevic, Yvonne Jansen, Abhraneel Sarma, Matthew Kay, and Fanny Chevalier. 2019. Increasing the Transparency of Research Papers with Explorable Multiverse Analyses. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–15. <https://doi.org/10.1145/3290605.3300295>
- [33] Nell K. Duke and P. David Pearson. 2009. Effective Practices for Developing Reading Comprehension. *Journal of Education* 189, 1–2 (Jan. 2009), 107–122. <https://doi.org/10.1177/0022057409189001-208>
- [34] Karolina Duklan and Michael A Martin. 2002. Communicating Effectively with Words, Numbers, and Pictures: Drawing on Experience. *Journal of Actuarial Practice* 10 (2002), 59.
- [35] Apache 2015. *Apache ECharts*. Apache. <https://echarts.apache.org/en/index.html>
- [36] Lei Fang and Jian-Guang Lou. 2021. Part & Whole Extraction: Towards A Deep Understanding of Quantitative Facts for Percentages in Text. arXiv:2110.13505 [cs]
- [37] Gloria D. Feliciano, Richard D. Powers, and Bryant E. Kearn. 1963. The Presentation of Statistical Information. *Audio Visual Communication Review* 11, 3 (1963), 32–39.
- [38] Fiona Fidler, Geoff Cumming, Neil Thomason, Dominique Pannuzzo, Julian Smith, Penny Fyffe, Holly Edmonds, Claire Harrington, and Rachel Schmitt. 2005. Toward Improved Statistical Reporting in the Journal of Consulting and Clinical Psychology. *Journal of Consulting and Clinical Psychology* 73, 1 (Feb. 2005), 136–143. <https://doi.org/10.1037/0022-006X.73.1.136>
- [39] Fiona Fidler and Geoffrey R. Loftus. 2009. Why Figures with Error Bars Should Replace p Values: Some Conceptual Arguments and Empirical Demonstrations. *Zeitschrift für Psychologie/Journal of Psychology* 217, 1 (2009), 27–37. <https://doi.org/10.1027/0044-3409.217.1.27>
- [40] Raymond Fok, Andrew Head, Jonathan Bragg, Kyle Lo, Marti A. Hearst, and Daniel S. Weld. 2022. Scim: Intelligent Faceted Highlights for Interactive, Multi-Pass Skimming of Scientific Papers. arXiv:2205.04561 [cs]
- [41] Tong Gao, Mira Dontcheva, Eytan Adar, Zhicheng Liu, and Karric G. Karahalios. 2015. DataTone: Managing Ambiguity in Natural Language Interfaces for Data Visualization. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. ACM, Charlotte NC USA, 489–500. <https://doi.org/10.1145/2807442.2807478>
- [42] Tong Gao, Jessica R. Hullman, Eytan Adar, Brent Hecht, and Nicholas Diakopoulos. 2014. NewsViews: An Automated Pipeline for Creating Custom Geovisualizations for News. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 3005–3014. <https://doi.org/10.1145/2556288.2557228>
- [43] Andrew Gelman, Cristian Pasarica, and Rahul Dodhia. 2002. Let's Practice What We Preach. *The American Statistician* 56, 2 (May 2002), 121–130. <https://doi.org/10.1198/000313002317572790>
- [44] Douglas J. Gillan, Christopher D. Wickens, J. G. Hollands, and C. Melody Carswell. 1998. Guidelines for Presenting Quantitative Data in HFES Publications. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 40, 1 (March 1998), 28–41. <https://doi.org/10.1518/001872098779480640>
- [45] Arthur M Glenberg and William E Langston. 1992. Comprehension of Illustrated Text: Pictures Help to Build Mental Models. *Journal of Memory and Language* 31, 2 (April 1992), 129–151. [https://doi.org/10.1016/0749-596X\(92\)90008-L](https://doi.org/10.1016/0749-596X(92)90008-L)
- [46] Pascal Goffin, Tanja Blascheck, Petra Isenberg, and Wesley Willett. 2020. Interaction Techniques for Visual Exploration Using Embedded Word-Scale Visualizations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–13. <https://doi.org/10.1145/3313831.3376842>
- [47] Pascal Goffin, Jeremy Boy, Wesley Willett, and Petra Isenberg. 2017. An Exploratory Study of Word-Scale Graphics in Data-Rich Text Documents. *IEEE Transactions on Visualization and Computer Graphics* 23, 10 (Oct. 2017), 2275–2287. <https://doi.org/10.1109/TVCG.2016.2618797>
- [48] Pascal Goffin, Wesley Willett, Anastasia Bezerianos, and Petra Isenberg. 2015. Exploring the Effect of Word-Scale Visualizations on Reading Behavior. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15)*. Association for Computing Machinery, New York, NY, USA, 1827–1832. <https://doi.org/10.1145/2702613.2732778>
- [49] 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 (Dec. 2014), 2291–2300. <https://doi.org/10.1109/TVCG.2014.2346435>
- [50] Pascal Goffin, Wesley Willett, Jean-Daniel Fekete, and Petra Isenberg. 2015. Design Considerations for Enhancing Word-Scale Visualizations with Interaction. In *Posters of the Conference on Information Visualization (InfoVis)*. IEEE, Chicago, Illinois, USA, 2.
- [51] Pascal Goffin, Wesley Willett, and Petra Isenberg. 2015. Sharing Information from Personal Digital Notes Using Word-Scale Visualizations. In *Proceedings of the IEEE VIS Workshop on Personal Visualization: Exploring Data in Everyday Life*. IEEE, Chicago Illinois USA, 5.
- [52] John Gramlich. 2020. What the 2020 Electorate Looks like by Party, Race and Ethnicity, Age, Education and Religion.
- [53] Transparent Statistics in Human-Computer Interaction Working Group. 2019. Transparent Statistics Guidelines. <https://doi.org/10.5281/zendeno.1186169>
- [54] Andrew Head, Kyle Lo, Dongyeop Kang, Raymond Fok, Sam Skjonsberg, Daniel S. Weld, and Marti A. Hearst. 2021. Augmenting Scientific Papers with Just-in-Time, Position-Sensitive Definitions of Terms and Symbols. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–18. <https://doi.org/10.1145/3411764.3445648>
- [55] Jeffrey Heer and George Robertson. 2007. Animated Transitions in Statistical Data Graphics. *IEEE Transactions on Visualization and Computer Graphics* 13, 6 (Nov. 2007), 1240–1247. <https://doi.org/10.1109/TVCG.2007.70539>
- [56] Jouni Helske, Satu Helske, Matthew Cooper, Anders Ynnerman, and Lonnii Besançon. 2021. Can Visualization Alleviate Dichotomous Thinking? Effects of Visual Representations on the Cliff Effect. *IEEE Transactions on Visualization and Computer Graphics* 27, 8 (Aug. 2021), 3397–3409. <https://doi.org/10.1109/TVCG.2021.3073466>
- [57] Ken Hinckley, Shengdong Zhao, Raman Sarin, Patrick Baudisch, Edward Cutrell, Michael Shilman, and Desney Tan. 2007. InkSine: In Situ Search for Active Note Taking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. Association for Computing Machinery, New York, NY, USA, 251–260. <https://doi.org/10.1145/1240624.1240666>
- [58] William E. Hockley. 2008. The Picture Superiority Effect in Associative Recognition. *Memory & Cognition* 36, 7 (Oct. 2008), 1351–1359. <https://doi.org/10.3758/MC.36.7.1351>
- [59] Fred Hohman, Matthew Conlen, Jeffrey Heer, and Duen Horng (Polo) Chau. 2020. Communicating with Interactive Articles. *Distill* 5, 9 (Sept. 2020), e28. <https://doi.org/10.23915/distill.00028>
- [60] Kasper Hornbaek and Erik Frokjær. 2003. Reading Patterns and Usability in Visualizations of Electronic Documents. *ACM Transactions on Computer-Human Interaction* 10, 2 (June 2003), 119–149. <https://doi.org/10.1145/772047.772050>
- [61] Kevin Hu, Michiel A. Bakker, Stephen Li, Tim Kraska, and César Hidalgo. 2019. VizML: A Machine Learning Approach to Visualization Recommendation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland UK, 1–12. <https://doi.org/10.1145/3290605.3300358>
- [62] Jessica Hullman, Nicholas Diakopoulos, and Eytan Adar. 2013. Contextifier: Automatic Generation of Annotated Stock Visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. Association for Computing Machinery, New York, NY, USA, 2707–2716. <https://doi.org/10.1145/2470654.2481374>
- [63] Jessica Hullman, Yea-Seul Kim, Francis Nguyen, Lauren Speers, and Maneesh Agrawala. 2018. Improving Comprehension of Measurements Using Concrete Re-expression Strategies. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–12. <https://doi.org/10.1145/3173574.3173608>
- [64] Junyong In and Sangseok Lee. 2017. Statistical Data Presentation. *Korean Journal of Anesthesiology* 70, 3 (2017), 267. <https://doi.org/10.4097/kjae.2017.70.3.267>
- [65] René Jorna. 1988. A Comparison of Presentation and Representation: Linguistic and Pictorial. In *Human-Computer Interaction*, Gerrit C. van der Veer and Gijsbertus Mulder (Eds.). Springer, Berlin, Heidelberg, 172–185. https://doi.org/10.1007/978-3-642-73402-1_11
- [66] Jonathan P. Kastellec and Eduardo L. Leon. 2007. Using Graphs Instead of Tables in Political Science. *Perspectives on Politics* 5, 4 (Dec. 2007), 755–771. <https://doi.org/10.1017/S1537592707072209>
- [67] Dave Kelly, Jaap Jasperse, and Ian Westbrooke. 2005. Designing Science Graphs for Data Analysis and Presentation. *Science & Technical Publishing, Department of Conservation, Wellington, New Zealand* 32 (2005), 68.
- [68] Alison Kidd. 1994. The Marks Are on the Knowledge Worker. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '94)*. Association for Computing Machinery, New York, NY, USA, 186–191. <https://doi.org/10.1145/191666.191740>
- [69] N. W. Kim, S. C. Joyner, A. Riegelhuth, and Y. Kim. 2021. Accessible Visualization: Design Space, Opportunities, and Challenges. *Computer Graphics Forum* 40, 3 (2021), 173–188. <https://doi.org/10.1111/cgf.14298>

- [70] Sunjung Kim and Linda J. Lombardino. 2015. Comparing Graphs and Text: Effects of Complexity and Task. *Journal of Eye Movement Research* 8, 3 (Oct. 2015), 17. <https://doi.org/10.16910/jemr.8.3.2>
- [71] Yea-Seul Kim, Jessica Hullman, and Maneesh Agrawala. 2016. Generating Personalized Spatial Analogies for Distances and Areas. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 38–48. <https://doi.org/10.1145/2858036.2858440>
- [72] Torsten L Klein. 2014. Communicating Quantitative Information: Tables vs Graphs. *MPRA* 60514 (2014), 20.
- [73] Nicholas Kong and Maneesh Agrawala. 2012. Graphical Overlays: Using Layered Elements to Aid Chart Reading. *IEEE Transactions on Visualization and Computer Graphics* 18, 12 (Dec. 2012), 2631–2638. <https://doi.org/10.1109/TVCG.2012.229>
- [74] Markus Konkol, Christian Kray, and Jan Suleiman. 2019. Creating Interactive Scientific Publications Using Bindings. *Proceedings of the ACM on Human-Computer Interaction* 3, EICS (June 2019), 1–18. <https://doi.org/10.1145/3331158>
- [75] Bum Chul Kwon, Florian Stoffel, and Dominik Jäckle. 2014. VisJockey: Enriching Data Stories through Orchestrated Interactive Visualization. In *Computation + Journalism Symposium*. New York, NY, USA, 5.
- [76] Chufai Lai, Zhiqian Lin, Ruike Jiang, Yun Han, Can Liu, and Xiaoru Yuan. 2020. Automatic Annotation Synchronizing with Textual Description for Visualization. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13.
- [77] Matthew Lamm, Arun Tejasvi Chaganty, Christopher D. Manning, Dan Jurafsky, and Percy Liang. 2018. Textual Analogy Parsing: What's Shared and What's Compared among Analogous Facts. *arXiv:1809.02700* [cs]
- [78] Jill H. Larkin and Herbert A. Simon. 1987. Why a Diagram Is (Sometimes) Worth Ten Thousand Words. *Cognitive Science* 11, 1 (1987), 65–100. <https://doi.org/10.1111/j.1551-6708.1987.tb00863.x>
- [79] Shahid Latif, Diao Liu, and Fabian Beck. 2018. Exploring Interactive Linking Between Text and Visualization. In *EuroVis 2018 - Short Papers*. The Eurographics Association, Brno Czech Republic, 4 pages. <https://doi.org/10.2312/EUROVISSHORT.20181084>
- [80] Shahid Latif, Kaidie Su, and Fabian Beck. 2019. Authoring Combined Textual and Visual Descriptions of Graph Data. In *EuroVis*. IEEE, Porto, Portugal, 5.
- [81] Shahid Latif, Zheng Zhou, Yoon Kim, Fabian Beck, and Nam Wook Kim. 2022. Kori: Interactive Synthesis of Text and Charts in Data Documents. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (Jan. 2022), 184–194. <https://doi.org/10.1109/TVCG.2021.3114802>
- [82] Elsie Lee-Robbins, Shiqing He, and Eytan Adar. 2022. Learning Objectives, Insights, and Assessments: How Specification Formats Impact Design. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (Jan. 2022), 676–685. <https://doi.org/10.1109/TVCG.2021.3114811>
- [83] Yuxuan Lei, Qi Lu, and Yingqing Xu. 2022. O&O: A DIY Toolkit for Designing and Rapid Prototyping Olfactory Interfaces. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–21. <https://doi.org/10.1145/3491102.3502033>
- [84] Deqing Li, Honghui Mei, Yi Shen, Shuang Su, Wenli Zhang, Junting Wang, Ming Zu, and Wei Chen. 2018. ECharts: A Declarative Framework for Rapid Construction of Web-Based Visualization. *Visual Informatics* 2, 2 (June 2018), 136–146. <https://doi.org/10.1016/j.visinf.2018.04.011>
- [85] Veronica Liesaputra and Ian H. Witten. 2012. Realistic Electronic Books. *International Journal of Human-Computer Studies* 70, 9 (Sept. 2012), 588–610. <https://doi.org/10.1016/j.ijhcs.2012.02.003>
- [86] Ziming Liu. 2005. Reading Behavior in the Digital Environment: Changes in Reading Behavior over the Past Ten Years. *Journal of Documentation* 61, 6 (Jan. 2005), 700–712. <https://doi.org/10.1108/00220410510632040>
- [87] Geoffrey R. Loftus. 1993. A Picture Is Worth a Thousand Values: On the Irrelevance of Hypothesis Testing in the Microcomputer Age. *Behavior Research Methods, Instruments, & Computers* 25, 2 (June 1993), 250–256. <https://doi.org/10.3758/BF03204506>
- [88] Min Lu, Jie Liang, Yu Zhang, Guozheng Li, Siming Chen, Zongru Li, and Xiaoru Yuan. 2017. Interaction+: Interaction Enhancement for Web-Based Visualizations. In *2017 IEEE Pacific Visualization Symposium (PacificVis)*. IEEE, Seoul Republic of Korea, 61–70. <https://doi.org/10.1109/PACIFICVIS.2017.8031580>
- [89] Yuyu Luo, Xuedi Qin, Nan Tang, and Guoliang Li. 2018. DeepEye: Towards Automatic Data Visualization. In *2018 IEEE 34th International Conference on Data Engineering (ICDE)*. IEEE, Paris, 101–112. <https://doi.org/10.1109/ICDE.2018.00019>
- [90] Yuyu Luo, Nan Tang, Guoliang Li, Jiawei Tang, Chengliang Chai, and Xuedi Qin. 2022. Natural Language to Visualization by Neural Machine Translation. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (Jan. 2022), 217–226. <https://doi.org/10.1109/TVCG.2021.3114848>
- [91] Shuai Ma, Taichang Zhou, Fei Nie, and Xiaojuan Ma. 2022. Glancee: An Adaptable System for Instructors to Grasp Student Learning Status in Synchronous Online Classes. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–25. <https://doi.org/10.1145/3491102.3517482>
- [92] Michael MacDonald-Ross. 1977. How Numbers Are Shown: A Review of Research on the Presentation of Quantitative Data in Texts. *AV Communication Review* 25, 4 (1977), 359–409.
- [93] Heinz Mandl and Joel R. Levin. 1989. *Knowledge Acquisition from Text and Pictures*. North-Holland, Amsterdam ; New York : New York, N.Y., U.S.A.
- [94] Damien Masson, Sylvain Malacria, Edward Lank, and Géry Casiez. 2020. Chameleon: Bringing Interactivity to Static Digital Documents. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–13. <https://doi.org/10.1145/3313831.3376559>
- [95] Damien Masson, Sylvain Malacria, Daniel Vogel, Edward Lank, and Géry Casiez. 2023. ChartDetective: Easy and Accurate Interactive Data Extraction from Complex Vector Charts. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (CHI '23). Association for Computing Machinery, New York, NY, USA, 17 pages. <https://doi.org/10.1145/3544548.3581113>
- [96] Richard E. Mayer. 2002. Multimedia Learning. In *Psychology of Learning and Motivation*. Vol. 41. Academic Press, 85–139. [https://doi.org/10.1016/S0079-7421\(02\)80005-6](https://doi.org/10.1016/S0079-7421(02)80005-6)
- [97] Andrew McNutt, Gordon Kindlmann, and Michael Correll. 2020. Surfacing Visualization Mirages. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–16. <https://doi.org/10.1145/3313831.3376420>
- [98] Ronald Metoyer, Qiyu Zhi, Bart Janczuk, and Walter Scheirer. 2018. Coupling Story to Visualization: Using Textual Analysis as a Bridge Between Data and Interpretation. In *23rd International Conference on Intelligent User Interfaces*. ACM, Tokyo Japan, 503–507. <https://doi.org/10.1145/3172944.3173007>
- [99] Joachim Meyer, Marcia Kuskin Sham, and Daniel Gopher. 1999. Information Structure and the Relative Efficacy of Tables and Graphs. *Human Factors* 41, 4 (Dec. 1999), 570–587. <https://doi.org/10.1518/001872099779656707>
- [100] Pam A. Mueller and Daniel M. Oppenheimer. 2014. The Pen Is Mightier Than the Keyboard: Advantages of Longhand Over Laptop Note Taking. *Psychological Science* 25, 6 (June 2014), 1159–1168. <https://doi.org/10.1177/0956797614524581>
- [101] Tamara Munzner. 2014. *Visualization Analysis and Design* (1st edition ed.). A K Peters/CRC Press, Boca Raton.
- [102] Arpit Narrechania, Arjun Srinivasan, and John Stasko. 2021. NL4DV: A Toolkit for Generating Analytic Specifications for Data Visualization from Natural Language Queries. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (Feb. 2021), 369–379. <https://doi.org/10.1109/TVCG.2020.3030378>
- [103] C. North. 2006. Toward Measuring Visualization Insight. *IEEE Computer Graphics and Applications* 26, 3 (May 2006), 6–9. <https://doi.org/10.1109/MCG.2006.70>
- [104] Kenton O'Hara. 1996. Towards a Typology of Reading Goals.
- [105] Kenton O'Hara and Abigail Sellen. 1997. A Comparison of Reading Paper and On-Line Documents. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*. ACM, Atlanta Georgia USA, 335–342. <https://doi.org/10.1145/258549.258787>
- [106] World Health Organization. 2021. Weekly Epidemiological Update on COVID-19 - 22 June 2021. <https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19--22-june-2021>.
- [107] World Health Organization. 2021. Weekly Epidemiological Update on COVID-19 - 29 June 2021. <https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19--29-june-2021>.
- [108] Alvitte Ottley, Aleksandra Kaszowska, R. Jordan Crouser, and Evan M. Peck. 2019. The Curious Case of Combining Text and Visualization. In *EuroVis*. The Eurographics Association, Porto, Portugal, 5. <https://doi.org/10.2312/evs.20191181>
- [109] Anshul Vikram Pandey, Katharina Rall, Margaret L. Satterthwaite, Oded Nov, and Enrico Bertini. 2015. How Deceptive Are Deceptive Visualizations? An Empirical Analysis of Common Distortion Techniques. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. Association for Computing Machinery, New York, NY, USA, 1469–1478. <https://doi.org/10.1145/2702123.2702608>
- [110] Jonas Parnow and Marian Dörk. 2015. Micro Visualizations: Data-driven Typography and Graphical Text Enhancement. In *InfoVis*. IEEE, Chicago, Illinois USA, 2.
- [111] Mozilla 2011. *PDF.js*. Mozilla. <https://mozilla.github.io/pdf.js/>
- [112] Jennifer Pearson, George Buchanan, and Harold Thimbleby. 2014. *Designing for Digital Reading*. Springer International Publishing, Cham. <https://doi.org/10.1007/978-3-03-02330-9>
- [113] Joao Pinheiro and Jorge Poco. 2022. ChartText: Linking Text with Charts in Documents. *arXiv:2201.05043* [cs]
- [114] Zachary Poussman, John Stasko, and Michael Mateas. 2007. Casual Information Visualization: Depictions of Data in Everyday Life. *IEEE Transactions on Visualization and Computer Graphics* 13, 6 (Nov. 2007), 1145–1152. <https://doi.org/10.1109/TVCG.2007.70541>
- [115] Morgan N. Price, Bill N. Schilit, and Gene Golovchinsky. 1998. XLibris: The Active Reading Machine. In *CHI '98 Conference Summary on Human Factors in Computing Systems (CHI '98)*. Association for Computing Machinery, New York, NY, USA, 22–23. <https://doi.org/10.1145/286498.286510>
- [116] PrimeTek Informatics 2017. *PrimeReact / React UI Component Library*. PrimeTek Informatics. <https://www.primefaces.org/primereact>

- [117] Chunyao Qian, Shizhao Sun, Weiwei Cui, Jian-Guang Lou, Haidong Zhang, and Dongmei Zhang. 2021. Retrieve-Then-Adapt: Example-based Automatic Generation for Proportion-related Infographics. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (Feb. 2021), 443–452. <https://doi.org/10.1109/TVCG.2020.3030448>
- [118] Jing Qian, Qi Sun, Curtis Wigington, Han L. Han, Tong Sun, Jennifer Healey, James Tompkin, and Jeff Huang. 2022. Dually Noted: Layout-Aware Annotations with Smartphone Augmented Reality. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–15. <https://doi.org/10.1145/3491102.3502026>
- [119] Ghulam Jilani Abdul Rahim Quadri. 2021. *Constructing Frameworks for Task-Optimized Visualizations*. Ph. D. Dissertation. University of South Florida, United States – Florida.
- [120] Napol Rachatasumrit, Jonathan Bragg, Amy X. Zhang, and Daniel S Weld. 2022. CiteRead: Integrating Localized Citation Contexts into Scientific Paper Reading. In *27th International Conference on Intelligent User Interfaces (IUI '22)*. Association for Computing Machinery, New York, NY, USA, 707–719. <https://doi.org/10.1145/3490099.3511162>
- [121] Meta. 2013. *React – A JavaScript Library for Building User Interfaces*. Meta. <https://reactjs.org/>
- [122] Donghao Ren, Matthew Brehmer, Bongshin Lee, Tobias Höllerer, and Eun Kyung Choe. 2017. ChartAccent: Annotation for Data-Driven Storytelling. In *2017 IEEE Pacific Visualization Symposium (PacificVis)*. IEEE, Seoul Republic of Korea, 230–239. <https://doi.org/10.1109/PACIFICVIS.2017.8031599>
- [123] Nathalie Henry Riche, Christophe Hurter, Nicholas Diakopoulos, and Sheelagh Carpendale. 2018. *Data-Driven Storytelling*. CRC Press, Natick, Massachusetts, USA.
- [124] Yann Riche, Nathalie Henry Riche, Ken Hinckley, Sheri Panabaker, Sarah Fulling, and Sarah Williams. 2017. As We May Ink?: Learning from Everyday Analog Pen Use to Improve Digital Ink Experiences. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, Denver Colorado USA, 3241–3253. <https://doi.org/10.1145/3025453.3025716>
- [125] 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*. ACM, Glasgow Scotland UK, 1–13. <https://doi.org/10.1145/3290605.3300272>
- [126] Hugo Romat, Emmanuel Pietriga, Nathalie Henry-Riche, Ken Hinckley, and Caroline Appert. 2019. SpaceInk: Making Space for In-Context Annotations. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (UIST '19)*. Association for Computing Machinery, New York, NY, USA, 871–882. <https://doi.org/10.1145/3323165.3347934>
- [127] Bahador Saket, Alex Endert, and Cagatay Demiralp. 2018. Task-Based Effectiveness of Basic Visualizations. [arXiv:1709.08546 \[cs\]](https://arxiv.org/abs/1709.08546)
- [128] Manolis Savva, Nicholas Kong, Arti Chhajta, Li Fei-Fei, Maneesh Agrawala, and Jeffrey Heer. 2011. ReVision: Automated Classification, Analysis and Redesign of Chart Images. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology - UIST '11*. ACM Press, Santa Barbara, California, USA, 393. <https://doi.org/10.1145/2047196.2047247>
- [129] Mike Scaife and Yvonne Rogers. 1996. External Cognition: How Do Graphical Representations Work? *International Journal of Human-Computer Studies* 45, 2 (Aug. 1996), 185–213. <https://doi.org/10.1006/ijhc.1996.0048>
- [130] Bill N. Schilit, Gene Golovchinsky, and Morgan N. Price. 1998. Beyond Paper: Supporting Active Reading with Free Form Digital Ink Annotations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '98)*. ACM Press/Addison-Wesley Publishing Co., USA, 249–256. <https://doi.org/10.1145/274644.274680>
- [131] Wolfgang Schnottz and Maria Bannert. 2003. Construction and Interference in Learning from Multiple Representation. *Learning and Instruction* 13, 2 (April 2003), 141–156. [https://doi.org/10.1016/S0959-4752\(02\)00017-8](https://doi.org/10.1016/S0959-4752(02)00017-8)
- [132] Edward Segel and Jeffrey Heer. 2010. Narrative Visualization: Telling Stories with Data. *IEEE Transactions on Visualization and Computer Graphics* 16, 6 (Nov. 2010), 1139–1148. <https://doi.org/10.1109/TVCG.2010.179>
- [133] Vidya Setlur, Sarah E. Battersby, Melanie Tory, Rich Gossweiler, and Angel X. Chang. 2016. Eviza: A Natural Language Interface for Visual Analysis. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*. Association for Computing Machinery, New York, NY, USA, 365–377. <https://doi.org/10.1145/2984511.2984588>
- [134] Leixian Shen, Enya Shen, Yuyu Luo, Xiaocong Yang, Xuming Hu, Xiongshuai Zhang, Zhiwei Tai, and Jianmin Wang. 2022. Towards Natural Language Interfaces for Data Visualization: A Survey. *IEEE Transactions on Visualization and Computer Graphics* (2022), 1–1. <https://doi.org/10.1109/TVCG.2022.3148007> arXiv:2109.03506 [cs]
- [135] Danqing Shi, Xinyue Xu, Fuling Sun, Yang Shi, and Nan Cao. 2021. Calliope: Automatic Visual Data Story Generation from a Spreadsheet. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (Feb. 2021), 453–463. <https://doi.org/10.1109/TVCG.2020.3030403>
- [136] Herbert A Simon. 1978. On the Forms of Mental Representation. *University of Minnesota Press, Minneapolis* (1978), 16.
- [137] Lauren M. Singer and Patricia A. Alexander. 2017. Reading on Paper and Digitally: What the Past Decades of Empirical Research Reveal. *Review of Educational Research* 87, 6 (Dec. 2017), 1007–1041. <https://doi.org/10.3102/0034654317722961>
- [138] Ian Spence. 1990. Visual Psychophysics of Simple Graphical Elements. *Journal of Experimental Psychology: Human Perception and Performance* 16, 4 (1990), 683–692. <https://doi.org/10.1037/0096-1523.16.4.683>
- [139] Ian Spence and Stephan Lewandowsky. 1991. Displaying Proportions and Percentages. *Applied Cognitive Psychology* 5, 1 (1991), 61–77. <https://doi.org/10.1002/acp.2350050106>
- [140] Arjun Srinivasan, Steven M. Drucker, Alex Endert, and John Stasko. 2019. Augmenting Visualizations with Interactive Data Facts to Facilitate Interpretation and Communication. *IEEE Transactions on Visualization and Computer Graphics* 25, 1 (Jan. 2019), 672–681. <https://doi.org/10.1109/TVCG.2018.2865145>
- [141] Arjun Srinivasan, Nikhil Naipathy, Bongshin Lee, Steven M. Drucker, and John Stasko. 2021. Collecting and Characterizing Natural Language Utterances for Specifying Data Visualizations. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–10. <https://doi.org/10.1145/3411764.3445400>
- [142] Hendrik Strobelt, Daniela Oelke, Christian Rohrdantz, Andreas Stoffel, Daniel A. Keim, and Oliver Deussen. 2009. Document Cards: A Top Trumps Visualization for Documents. *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (Nov. 2009), 1145–1152. <https://doi.org/10.1109/TVCG.2009.139>
- [143] Hariharan Subramonyam, Colleen Seifert, Priti Shah, and Eytan Adar. 2020. texSketch: Active Diagramming through Pen-and-Ink Annotations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–13. <https://doi.org/10.1145/3313831.3376155>
- [144] Nicole Sultanum, Fanny Chevalier, Zoya Bylinskii, and Zhicheng Liu. 2021. Leveraging Text-Chart Links to Support Authoring of Data-Driven Articles with VizFlow. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–17. <https://doi.org/10.1145/3411764.3445354>
- [145] Kaiwen Sun, Carlo Sugatan, Tanisha Afnan, Hayley Simon, Susan A. Gelman, Jenny Radesky, and Florian Schaub. 2021. “They See You’re a Girl If You Pick a Pink Robot with a Skirt”: A Qualitative Study of How Children Conceptualize Data Processing and Digital Privacy Risks. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–34. <https://doi.org/10.1145/3411764.3445333>
- [146] Bo Tang, Shi Han, Man Lung Yiu, Rui Ding, and Dongmei Zhang. 2017. Extracting Top-K Insights from Multi-dimensional Data. In *Proceedings of the 2017 ACM International Conference on Management of Data*. ACM, Chicago Illinois USA, 1509–1524. <https://doi.org/10.1145/3035918.3035922>
- [147] Craig S. Tashman and W. Keith Edwards. 2011. Active Reading and Its Discontents: The Situations, Problems and Ideas of Readers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 2927–2936. <https://doi.org/10.1145/1978942.1979376>
- [148] Craig S. Tashman and W. Keith Edwards. 2011. LiquidText: A Flexible, Multi-touch Environment to Support Active Reading. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Vancouver BC Canada, 3285–3294. <https://doi.org/10.1145/1978942.1979430>
- [149] Edward Tufte. 2006. *Beautiful Evidence*. Graphics Press, Cheshire, Conn.
- [150] UNESCO Institute for Statistics. 2017. Literacy Rates Continue to Rise from One Generation to the next - UNESCO Digital Library. <https://unesdoc.unesco.org/ark:/48223/pf000258942?posInSet=1&queryId=826eb5c2-2600-4792-a3c8-83848b42df94>
- [151] U.S. Bureau of Labor Statistics. 2022. Employment Situation Summary - 2022 M08 Results. <https://www.bls.gov/news.release/empstat.nr.htm>.
- [152] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C. J. Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. 2020. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* 17 (2020), 261–272. <https://doi.org/10.1038/s41592-019-0686-2>
- [153] J. N. Washburne. 1927. An Experimental Study of Various Graphic, Tabular, and Textual Methods of Presenting Quantitative Material. *Journal of Educational Psychology* 18, 7 (1927), 465–476. <https://doi.org/10.1037/h0070054>
- [154] Alan J. Wecker, Joel Lanir, Osnat Mokry, Einat Minkov, and Tsvi Kuflik. 2014. Semantize: Visualizing the Sentiment of Individual Document. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces - AVI '14*. ACM Press, Como, Italy, 385–386. <https://doi.org/10.1145/2598153.2600056>

- [155] Kanit Wongsuphasawat, Dominik Moritz, Anushka Anand, Jock Mackinlay, Bill Howe, and Jeffrey Heer. 2016. Voyager: Exploratory Analysis via Faceted Browsing of Visualization Recommendations. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (Jan. 2016), 649–658. <https://doi.org/10.1109/TVCG.2015.2467191>
- [156] Keke Wu, Emma Petersen, Tahmina Ahmad, David Burlinson, Shea Tanis, and Danielle Albers Szafir. 2021. Understanding Data Accessibility for People with Intellectual and Developmental Disabilities. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–16. <https://doi.org/10.1145/3411764.3445743>
- [157] Sally P. W. Wu and Martina A. Rau. 2019. How Students Learn Content in Science, Technology, Engineering, and Mathematics (STEM) Through Drawing Activities. *Educational Psychology Review* 31, 1 (March 2019), 87–120. <https://doi.org/10.1007/s10648-019-09467-3>
- [158] Qian Yang, Gerard de Melo, Yong Cheng, and Sen Wang. 2017. HiText: Text Reading with Dynamic Salience Marking. In *Proceedings of the 26th International Conference on World Wide Web Companion (WWW '17 Companion)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 311–319. <https://doi.org/10.1145/3041021.3054168>
- [159] Ji Soo Yi, Youn ah Kang, John Stasko, and J.A. Jacko. 2007. Toward a Deeper Understanding of the Role of Interaction in Information Visualization. *IEEE Transactions on Visualization and Computer Graphics* 13, 6 (Nov. 2007), 1224–1231. <https://doi.org/10.1109/TVCG.2007.70515>
- [160] Qiyu Zhi, Alvitta Ottley, and Ronald Metoyer. 2019. Linking and Layout: Exploring the Integration of Text and Visualization in Storytelling. *Computer Graphics Forum* 38, 3 (2019), 675–685. <https://doi.org/10.1111/cgf.13719>