A Qualitative Evaluation and Taxonomy of Student Annotations on Bar Charts

Md Dilshadur Rahman* — University of Utah

Ghulam Jilani Quadri† — University of North Carolina at Chapel Hill

Paul Rosen‡ — University of Utah

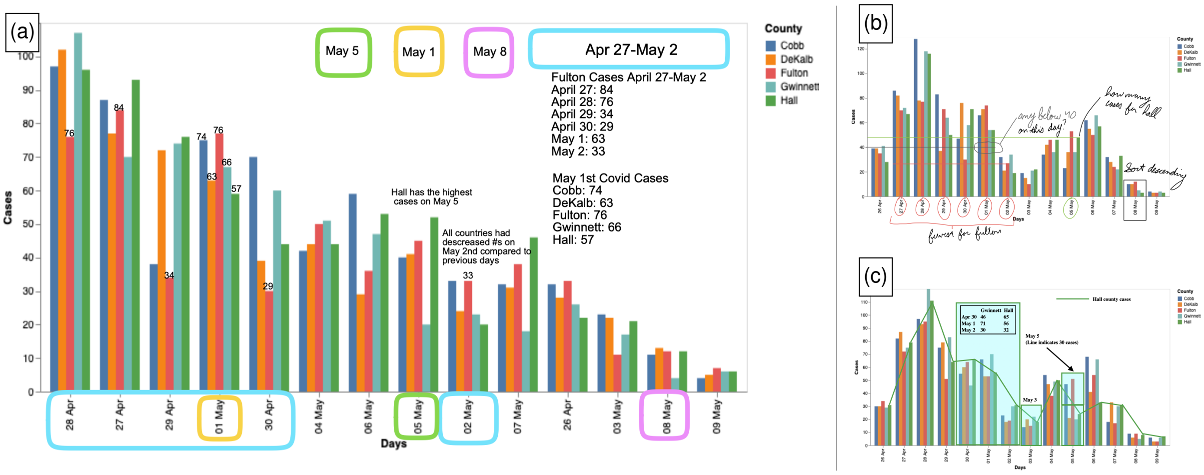

Figure 1: Three examples of annotated bar chart submissions. (a) One student used rectangles for filtering, text on bars for retrieving values and pointing out the extrema, and a legend for filtering important dates. (b) A second student annotated the chart with ellipses, rectangles, and lines for filtering, and text for filtering and as an identifier. (c) A final student used rectangular shapes and highlights for filtering, texts as an identifier, a trend line for finding extrema, and a legend for filtering.

Abstract

When sharing visualizations, annotations provide valuable insights into the data by focusing attention on important visual elements. As a result, annotations have become an essential part of visualizations, primarily when externalizing data or engaging in collaborative analysis. Therefore, it is crucial to understand how people annotate visualizations. This two-phase study used individual and group settings to investigate how visualization students annotate bar charts when asked to answer high-level questions about the data in the charts. The resulting annotations were coded and summarized into a taxonomy with several interesting findings. For example, we notice that several annotation types were broadly applied, while others were just used in special cases. In addition, ensembles of annotations were required to sufficiently annotate specific tasks. Our findings provide an early framework for contextualizing the usage of annotations, further providing guidance for best practice uses.

Index Terms: Human-centered computing—Visualization—Visualization design and evaluation methods

1 Introduction

Annotations are essential for better interpretation and understanding of data in visualizations. Studies demonstrated that visualizations that contain annotations help viewers' mental organization [3]. Past studies have also identified annotation as one of the "critical tasks that enable iterative visual analysis" [10, 22]. However, a natural question arose to us, which is how might someone annotate the graphic to make it clearer, and what are the limits to the effectiveness of those annotations?

As an example, in 2020, a Georgia Department of Public Health (GDPH) visualization became a viral sensation for its misleading nature (see Fig. 2) [7]. While technically accurate and well-labeled, the grouped bar chart was seen as deceptive for its unorthodox ordering of bars, ordered highest-to-lowest instead of chronologically. One principled approach to resolving such ambiguity is clear labeling of the graphic [21], which leads to the question of what constitutes clear labeling of the graphic and can annotations overcoming bad design?

To better understand the process and options available for annotating a graphic, we set about to study the forms of annotations used by undergraduate and graduate visualization students when prompted with high-level questions about the data. In our two-phase study, we provided students with 3 grouped bar chart visualizations, each with 4 high-level questions. In the first phase, we asked students to individually enumerate which of 5 low-level tasks (retrieve value, filter, compute derived value, find extremum, and sort) were required to answer those questions and annotate the bar charts to make the questions as easy as possible to answer. The goal of this phase was ideation, i.e., we wanted students to creatively experiment with possible annotation types. The second phase repeated the assignment, except students were put into groups to discuss and annotate the visualizations together. The goal of this second phase was to have students discuss and decide which of the annotations they believe were most effective.

The contribution of this paper is that we coded and summarized the resulting annotated visualizations to build a design space, which showed that 5 primary annotation categories were used: enclosure, connectors, text, marks, and color, on the 5 low-level tasks. Within

*e-mail: dilshadur@sci.utah.edu
†e-mail: jiquad@cs.unc.edu
‡e-mail: prosen@sci.utah.edu

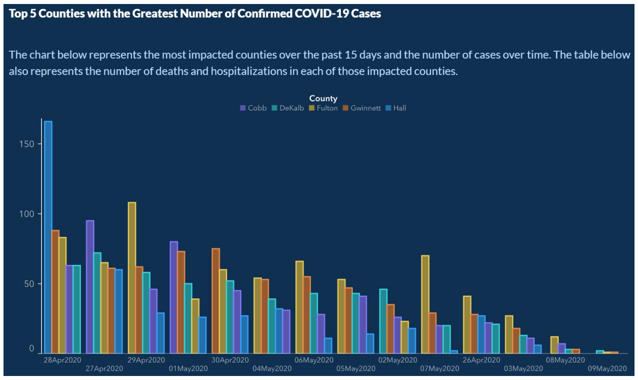

Figure 2: The viral visualization created by the Georgia Department of Public Health (GDPH)

this context, we found that most annotation types could be used for several tasks, specifically, retrieve value and filter. For the other low-level tasks, specifically, compute derived value, find extremum, and sort, we observed more targeted use of a subset of annotation types. Additionally, while the use of annotations did not differ significantly between the two phases, we found more targeted use of certain annotation types were used in the group phase. Further investigation also revealed that ensembles of annotations were sometimes used for a given task, e.g., enclosure, connector, and text were used for annotating computing derived value tasks. These results provide an early framework for contextualizing annotations and low-level tasks they support, which provide guidance for best practice usage.

2 Prior Work

Several roles for annotations have been previously investigated.

Externalization and Exploration Creating annotations is one of the most critical steps that analysts perform in exploratory data analysis. In addition, annotations serve as a vital part of externalizing an analyst's thought process, which allows an analyst to highlight important facts about the data in a visualization [12]. For example, a prior study that visualized user-authored annotation graphs concluded that annotations play a pivotal role in making sense of data and help with meta-analysis and data externalization [22]. Another study used annotations (e.g., text, brackets, arrows, etc.) in the tool VisInReport, a visual analysis tool for insight report generation from discourse transcripts, to clarify different events in different views [18]. They found that annotation played a crucial role in the reports—participants mentioned that it was impossible to include all aspects of the data without the annotations.

Collaborative Analysis Annotation can also help in collaborative analysis by externalizing the thoughts of the analyst and sharing them with collaborators [16]. Chen et al. discussed the importance of facilitating annotations in Asynchronous Collaborative Visual Analysis to establish common ground so that the insights from the individual work can be used to develop a shared understanding [4]. They used semi-automatic annotation combined with a semi-automatic common ground construction approach using the tool Click2Annotate [5]. It demonstrated that insight annotations could capture low-level analytics task results (e.g., clusters and outliers) with higher semantic richness over actions and events. Furthermore, an analytical processes study of practices in Intelligence Analysis found that annotations were frequently used in different forms (i.e., pens, highlighters, etc.) to understand the facts or events in the dataset, which suggests that sketching tools that support annotations may be helpful in this kind of data analysis and that annotation tools should facilitate collaborative data analysis [6]. Breslav et al. talked about a tool named UberTagger that provided a comment panel where users could use comments as annotations or tags to highlight any event or issue and communicate them to the other members of the team [2, 19].

Annotation Techniques A variety of annotation types appear in prior work. Kong et al. introduced graphical overlays, visual elements that are layered onto charts, such as reference structures (e.g., gridlines), highlights (e.g., outlines around important marks), redundant encodings (e.g., numerical data labels), summary statistics (e.g., min or max), and annotations (e.g., descriptive text), that help the reader with several tasks [14]. Additionally, they proposed an automated method for applying these graphic overlays to existing charts. Free-form ink annotations (e.g., circles, lines, arrows, etc.) and text were used in line charts to highlight the visually salient perceptual features (e.g., peaks, valleys, rising slopes, declining slopes, etc.) during asynchronous collaborations [13]. These annotations identified and highlighted salient aspects of the visualization and attracted attention. In another study, annotations on bar charts, specifically trends in bar heights and text values on bars, were used to optimize visualization layout by utilizing an area of interest measure [15].

Applications Utilizing Annotation The application of annotations was suggested in the visual analytics system Jigsaw [20], a comprehensive tool that helps with sense-making during investigative analysis. The evaluation of the system demonstrated how it facilitated annotations [11]. Another application of annotations was demonstrated using a tool that reveals deception (e.g., truncated y-axis, inverted y-axis, aspect ratio distortion, etc.) in line charts [9]. The browser extension read line chart images and generated text and visual annotations to help readers assess the distortions in line charts and come to an accurate understanding of the data. Finally, a comparison of simple visualizations in physical and virtual reality using annotations was conducted to see whether virtual annotation and filtering effect viewers' ability to solve fundamental data analysis tasks [8]. The study investigated whether participants preferred the physical or VR visualizations and how the VR experience might be shaped by using virtual annotations.

Apart from all of these different investigations of annotations, our work introduces a new design space of annotations, specifically focused on bar charts. Such a design space forms an important foundation towards codifying the best practice for annotating graphics to improve their readability.

3 Methodology & Study

The goal of our study was to understand how people annotate bar charts when prompted with specific questions about the data. To do this, we utilized and analyzed a course assignment in a mixed undergraduate and graduate data visualization course. In particular, we conducted a two-phase study where we evaluated annotation patterns on grouped bar charts based on the GDPH visualization (see Fig. 2).

3.1 Study Procedure

Individual Phase In the first phase, the students completed the assignment individually. We began with a brief in-class discussion (10-15 minutes) about the GDPH visualization. We then verbally reviewed the written assignment instructions and answered any questions. Students had 7 days to complete the assignment, which was done out of class. We did not require them to use any specific tool to complete the work.

Group Phase In the second phase, students came together in groups to share ideas and submit a collaborative work. To do this, we asked them to share and discuss each individual's annotations with the group and then use that discussion to determine a better set of annotations for the charts. Students had 7 days to complete the assignment. Some class time was allocated ($\sim$15 minutes) to start the assignment, but it was mostly completed out of class.

Task	Enclosure				Connector		Text			Mark	Color				Other
	Ellipse	Bracket	Half Box	Rectangle	Arrow	Line	Descprition	Values	Legend	Identifier	Highlights	Category	Questions	Values	
Retrieve value	15	4	0	14	18	52	19	11	4	1	0	33	14	2	0
Filter	20	18	10	31	34	47	35	4	11	27	7	30	55	2	0
Compute Derived Value	1	6	0	8	0	9	11	8	4	1	0	15	3	1	5
Find Extremum	11	0	0	5	11	15	21	7	3	3	0	14	9	0	0
Sort	0	0	0	1	6	9	14	1	3	1	0	6	3	0	0
Total	47	28	10	59	69	132	100	31	25	33	7	98	84	5	5

(a) Summary of Individual Phase Annotations

Enclosure				Connector		Text			Mark	Color				Other
Ellipse	Bracket	Half Box	Rectangle	Arrow	Line	Description	Values	Legend	Identifier	Highlights	Category	Questions	Values	
3	2	2	2	5	34	6	2	0	1	0	12	9	1	0
5	5	5	22	11	21	16	0	11	7	0	25	45	0	0
4	0	0	10	2	1	17	1	0	0	0	5	3	1	0
2	0	0	3	0	8	9	1	1	0	0	3	4	1	0
2	0	0	0	1	5	4	2	0	0	0	0	0	0	2
16	7	7	37	19	69	52	6	12	8	0	45	61	3	2

(b) Summary of Group Phase Annotations

Figure 3: Summary of the two-level taxonomy of annotations with the frequency of different tasks for the (a) individual and (b) group phases. Box colors indicate frequency: 1-9 10-19 20+ 1-5 6-14 15+.

3.2 Participants

The participants are the students from the Interactive Data Visualization course that one of the authors taught at the University of South Florida in Spring 2022. The course was a cross-listed elective for senior-level undergraduate and Master's- and Ph.D.-level graduate students with a total of 39 students (21 undergraduate and 18 graduate). The assignment was given approximately halfway through the semester. By that point, course material had covered most visualization foundations, but there had been no specific lecture focusing on bar charts. For the group phase, the students were divided into 10 groups of 3-5 students they had collaborated with throughout the semester. Students self-selected groups, and while they could mix undergraduate and graduate, generally, they matched with students of the same level.

3.3 Assignment

Each assignment had 3 bar charts, each with 4 questions/high-level tasks (12 questions in total). For each question in the assignment, subjects were asked to: 1) identify low-level tasks and 2) annotate the visualization to make answering that question easier. A sample assignment is available in the supplemental materials.

3.3.1 Visualizations

Datasets We wanted each student to have different data but similar trends within the data. Therefore, the datasets we used were generated using a random number for each county from the GDPH visualization from April 26 to May 9. We ensured that each random number fell within a specific range, e.g., for Fulton on May 4, the range was set to a minimum of 33 and a maximum of 56.

Visualization We used Vega-Lite [17] to create the grouped bar charts from the generated data. As with the GDPH visualization, we had five bars, one per county, of different colors for each day from April 26 to May 9. Like the GDPH visualization, we had 2 variants of the x-axis, one non-chronological (like the viral GDPH visualization) and one chronological (like the corrected GDPH visualization). In total, we generated 150 bar charts, half chronological and half non-chronological. For the individual phase, we built 40 unique assignments of 3 bar charts each. Half of them had 2 bar charts with chronological dates, 1 with non-chronological dates, whereas the other half had 2 bar charts with non-chronological dates and 1 with chronological dates. For the group phase, we generated 10 unique assignments in a similar manner to the individual phase. Each had 3 visualizations—half had 2 bar charts with chronological dates and 1 with non-chronological dates and vice versa.

3.3.2 High-Level Tasks

Students were asked to annotate the visualization based on questions about the data in the charts. There were 4 types of questions:

- finding a specific value (e.g., how many COVID cases are there in Hall County on May 5?);
- filtering some values from others (e.g., which counties have fewer than 40 cases of COVID on May 1?);
- aggregating (e.g., how many total COVID cases does Dekalb County have from May 1 to May 4?); and
- sorting (e.g., sort the counties in descending order based on the number of COVID cases on May 8).

3.3.3 Subject Tasks

Low-Level Task Identification We developed the high-level tasks so that the students could answer them by performing one or more low-level analysis tasks. We used Amar et al.'s low-level task taxonomy, which enumerated 10 tasks people frequently use to understand data in visualizations [1]. From that set, we selected 5 that fit into our study, including: *retrieve value (RV)*, *filter*, *compute derived value (CDV)*, *find extremum (FE)*, and *sort*. Students were asked to enumerate which of these low-level tasks were used to answer each of the high-level questions. There are 12 total questions for each assignment. Among these, RV and filter appear most commonly in our questions. RV was associated with every question, and at least half had a filter task. For other tasks, there were at least 3 CDV tasks, 2 FE tasks, and 2 sort tasks.

Annotating the Visualization We instructed the students to annotate the charts while looking for answers to the questions associated with the charts. Their goal was to annotate the charts in a way that made the questions as easy as possible to determine without necessarily writing the answer on the visualization. We did *not* ask the participants to directly relate their annotations to any low-level task.

4 Evaluation

4.1 Data

A total of 38 students completed the study's individual phase, and all 10 groups completed the second phase. Once we received all the worksheets from both phases, we anonymized them, assigned each a random number, and performed quality checks. In total, our evaluation includes 20 individual and 7 group assignments. We excluded 18 individual submissions from the analysis (13 had no annotations, and 5 had minimal annotations). One student annotated only 2 of the 3 charts, but their work was included in the evaluation. Similarly, 3 group submissions had no annotations and were not evaluated. All anonymized submissions are included in our supplemental materials.

4.2 Taxonomy

To evaluate and summarize the submitted assignments, 2 co-authors went through all the annotated charts and hand-coded them in several iterations. The coding involved identifying specific annotation marks and associating them with low-level tasks. In the first iteration, they separately identified a set of annotation types which they were able to link with different tasks performed by the participants by carefully examining the submissions. Next, all authors discussed their findings and agreed on an initial taxonomy of annotation types. In the second and subsequent iterations, the 2 coding co-authors independently revisited each submitted chart, re-categorizing the annotations, and the group revised the taxonomy. This process continued for several iterations until the group reached a consensus on a common taxonomy and agreed to all codings. Ultimately, we identified 14 annotation types which were grouped into 5 top-level annotation types: *Enclosure*, *Connector*, *Text*, *Mark*, and *Color*. The resulting two-level taxonomy is shown in Fig. 3.

4.3 Annotations

4.3.1 Enclosure

Enclosure is an annotation that uses a partially or fully closed boundary, including the *ellipse*, *bracket*, *half-box*, and *rectangle* categories. The enclosure annotations were used in a variety of situations. For example, in Fig. 1a, the rectangle is used for filtering the number of cases in Fulton county on May 1. Similarly, in Fig. 1b, ellipses are used for the filtering, half-boxes are used to support the filtering, and a rectangle is used for a sorting task. Brackets are used to mark the range indicating the bar/axes in Fig. 1b from April 27 to May 2. In the individual phase, enclosure was generally used for RV and filter tasks, and CDV and FE tasks to a lesser extent. Ellipse and rectangle annotations were used most frequently. In contrast, during the group phase, enclosure was primarily used for filter and CDV tasks, with rectangles being used most frequently.

4.3.2 Connector

A connector is an annotation that uses a line (e.g., solid, dotted, or directional line) to answer questions about data in the bar charts, which includes *arrow* (directional) and *line* (undirected) categories. Line annotations were used, e.g., to mark the height of a bar relative to the axis (see Fig. 4a) or to represent the trend in the height of bars (see Fig. 1c). Similarly, arrow annotations were used for pointing text to a particular bar or group of bars (see Fig. 1c and Fig. 4a) or to point the enclosure annotation to a bar, axis-value, or legend, or vice-versa. In both the individual and group phases of the assignment, students used connector annotations for RV, filter, and, to a lesser extent, FE tasks. In the individual phase, connectors were used for some sort tasks, but that pattern diminished in the group phase. Notably, most of the time, connectors did not appear alone but were combined with other annotations for a given task.

4.3.3 Text

ABC

Text is an annotation that uses words or sentences to answer questions about the data, with second-level categories of *descriptions*, *values*, and *legends*. A *description* annotation is defined in our taxonomy as text that describes a process, information, computation, or derivation that supports a particular task annotation, e.g., in Fig. 1a where the text "Hall has the highest cases on May 5." *Value* is another annotation that is a specific text-based annotation used to highlight exact data of the bar, e.g., in Fig. 1b and Fig. 1c where a number is used to represent the bar length. *Legend* is the final text-based annotation, which was used for any labels on legends, e.g., the text enclosed in the colored boxes on Fig. 1a or annotating the trend in a number of cases in the Hall County in Fig. 1c. From Fig. 3, description-based text annotations were broadly applied for all 5 tasks. Text enumerating values was used frequently for RV tasks in the individual but not group phase. Finally, legend text was frequently used for filter tasks. Overall, text annotations seemed to be used for elaborating on what other annotations were highlighting and as a last resort when no other annotations were suitable.

(a)

(b)

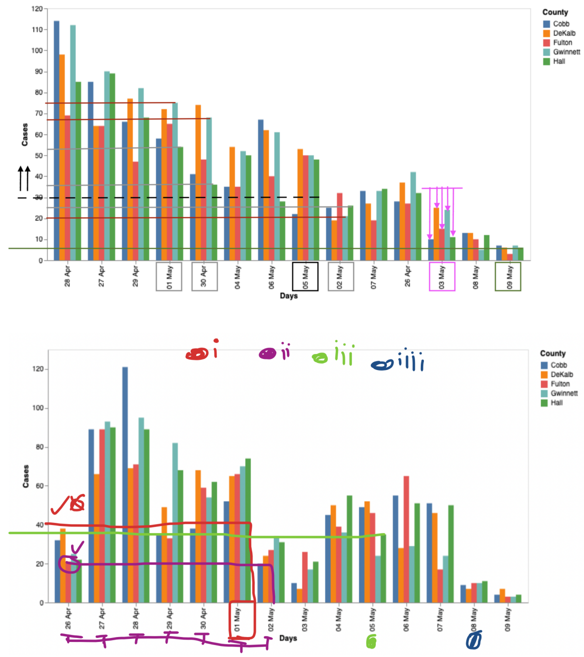

Figure 4: (a) An example of student annotations of the chart, including rectangles, lines, and arrows used for filtering and arrows used for sorting, and (b) another example including lines, circles, T-shaped, and circular marks.

4.3.4 Mark

Marks are annotations that use symbols or shapes to answer questions about data, which perform the function of *identifying* a particular object or category (e.g., a county or date). For example, in Fig. 4b, different markers (i.e., i, ii, iii, etc.) were used to denote the question numbers. In the same submission, circular and T-shaped marks are used to denote different dates. Further, marks were always accompanied by some sort of color for differentiating the markers, which is captured by the color category. From Fig. 3, mark-based annotations were primarily used for filter tasks in both individual and group phases.

4.3.5 Color

Color annotations are any property of color (generally, but not exclusively, hue) to answer questions about data, which is divided into second-level hierarchy categories, including *highlights*, *category*, *questions*, and *values*. The *highlight*-based annotations use different colors to highlight an area, bars, or group of bars (see Fig. 1c) that helps annotate a given task. Color for *values* annotations are an annotation type that aligns with text for value annotations (similar to a color map). The *category*-based color annotations highlight different categories (e.g., counties, dates/months, etc.) and are typically used in collaboration with mark- and enclosure-based annotations. For example, in Fig. 1a, four different colors are used for the date enclosures. The *question*-based color annotates the different questions (e.g., Q1, Q2, etc.) used in coordination with mark-based annotations. For example, roman numerals used in Fig. 4b are colored. From Fig. 3, color-based annotations were used heavily for all tasks, except sort, in the individual phase, while highlights were only used for the filter task. In contrast, color was used mainly for RV and filter tasks for the group phase.

4.3.6 Other

The other category comprises some special symbols used in the annotation. In the individual study, we found plus (+) used for adding up different values for the CDV task, delta (Δ) for differentiating groups of bars, and pound (#) used with an enclosure bracket for the CDV task. In the group phase, *overview+detail* was used to show a zoomed portion of the bar chart.

4.4 Discussion

Annotation Ensembles Several annotations were used as an ensemble, e.g., mark-color, enclosure-connector-text, connector-text. These combinations occur because individual annotations cannot stand on their own, e.g., the text might need a connector for spacing reasons. Our taxonomy does not directly capture these, but classifying them remains an interesting direction for future work.

The Sorting Challenge The sorting task appeared to be challenging for students to annotate. Some students used the arrow connector for the sorting task (see Fig. 4a), but most often, students relied upon enclosure to show the part of the charts that needed to be sorted. We also observed the use of lines and text to annotate sorting tasks, but that can be challenging when data are dense, or the number of data to sort is large.

Individual vs. Group Designs While we wanted to understand how students would use annotations individually versus group, we did not expect a large difference in the annotations used. We were a little surprised that most of the groups appeared to focus on a similar limited set of annotations when they worked together. For example, with enclosure, much more variety was used in the individual phase in terms of encodings and tasks, while rectangles with filter and CDV tasks were the majority in the group phase. This pattern can be observed with other encodings and tasks too. One possible reason behind this may be that the students used their individual annotated charts as the basis to make their group annotations more coherent. However, further study is required to better understand the cause of this behavior.

Annotation Effectiveness One important aspect of understanding the utility of annotations is to understand the efficacy of the different options available. While our study has provided a comprehensive taxonomy of different annotation types used on bar charts, it remains an important piece of future work to evaluate which are best and in what situations.

5 Conclusion

Annotations on visualizations support the visual exploration of data and provide a salient visual aid to users. They can be used to help generate hypotheses, communicate information, to assist in sense-making, and to collaboratively analyze charts. Understanding the different annotations used on visualizations (bar charts in our case) could help in developing a comprehensive list of annotation encodings, which could, in turn, provide design guidelines for collaborative user-analytic tasks and for knowledge transfer. To address these needs, we developed a two-level taxonomy of annotations used on bar charts. The taxonomy grouped annotations into 5 top-level categories, enclosure, connectors, text, marks, and color, and related them to 5 low-level tasks. The taxonomy highlights the diversity in available annotations, but it also points to a more limited subset being seen as popular or more effective in the eyes of the student groups. Ultimately, the use of annotation is heavily dependent on the nature of the tasks being performed.

Acknowledgments

This work was partially supported by the National Science Foundation (IIS-1845204 and DUE-2216227).

References

[1] R. Amar, J. Eagan, and J. Stasko. Low-level components of analytic activity in information visualization. In *IEEE Symposium on Information Visualization (InfoVis)*, pp. 111–117, 2005. doi: 10.1109/INFVIS.2005.1532136

[2] S. Breslav, R. Goldstein, A. Khan, and K. Hornbæk. Exploratory sequential data analysis for multi-agent occupancy simulation results. In *Symposium on Simulation for Architecture & Urban Design*, pp. 59–66, 2015.

[3] A. Cedilnik and P. Rheingans. Procedural annotation of uncertain information. In *IEEE Visualization*, pp. 77–84, 2000. doi: 10.1109/VISUAL.2000.885679

[4] Y. Chen, J. Alsakran, S. Barlowe, J. Yang, and Y. Zhao. Supporting effective common ground construction in asynchronous collaborative visual analytics. In *IEEE Symposium on Visual Analytics Science and Technology (VAST)*, pp. 101–110, 2011. doi: 10.1109/VAST.2011.6102447

[5] Y. Chen, S. Barlowe, and J. Yang. Click2annotate: Automated insight externalization with rich semantics. In *IEEE Symposium on Visual Analytics Science and Technology (VAST)*, pp. 155–162, 2010. doi: 10.1109/VAST.2010.5652885

[6] G. Chin Jr, O. A. Kuchar, and K. E. Wolf. Exploring the analytical processes of intelligence analysts. In *ACM SIGCHI Conference on Human Factors in Computing Systems*, pp. 11–20, 2009. doi: 10.1145/1518701.1518704

[7] S. Collins. Georgia's covid-19 cases aren't declining as quickly as initial data suggested they were. Vox, 2020.

[8] K. Danyluk, T. T. Ulusoy, W. Wei, and W. Willett. Touch and beyond: Comparing physical and virtual reality visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 2020. doi: 10.1109/TVCG.2020.3023336

[9] A. Fan, Y. Ma, M. Mancenido, and R. Maciejewski. Annotating line charts for addressing deception. In *ACM SIGCHI Conference on Human Factors in Computing Systems*, 2022.

[10] J. Heer and B. Shneiderman. Interactive dynamics for visual analysis. *Communications of the ACM*, 55(4):45–54, 2012. doi: 10.1145/2133806.2133821

[11] Y.-a. Kang, C. Gorg, and J. Stasko. Evaluating visual analytics systems for investigative analysis: Deriving design principles from a case study. In *IEEE Symposium on Visual Analytics Science and Technology (VAST)*, pp. 139–146, 2009. doi: 10.1109/VAST.2009.5333878

[12] Y.-a. Kang and J. Stasko. Characterizing the intelligence analysis process through a longitudinal field study: Implications for visual analytics. *Information Visualization*, 13(2):134–158, 2014. doi: 10.1177/1473871612468877

[13] N. Kong and M. Agrawala. Perceptual interpretation of ink annotations on line charts. In *ACM Symposium on User Interface Software and Technology (UIST)*, pp. 233–236, 2009. doi: 10.1145/1622176.1622219

[14] N. Kong and M. Agrawala. Graphical overlays: Using layered elements to aid chart reading. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2631–2638, 2012. doi: 10.1109/TVCG.2012.229

[15] M. Raschke, S. Strohmaier, T. Blascheck, and T. Ertl. Annotation of graphical elements in visualizations for an efficient analysis of visual tasks. In *ACM SIGCHI Conference on Human Factors in Computing Systems Extended Abstracts*, pp. 1387–1398, 2014. doi: 10.1145/2559206.2581305

[16] P. M. Sanderson and C. Fisher. Exploratory sequential data analysis: Foundations. *Human–Computer Interaction*, 9(3-4):251–317, 1994.

[17] A. Satyanarayan, D. Moritz, K. Wongsuphasawat, and J. Heer. Vega-lite: A grammar of interactive graphics. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):341–350, 2016. doi: 10.1109/TVCG.2016.2599030

[18] R. Sevastjanova, M. El-Assady, A. J. Bradley, C. Collins, M. Butt, and D. Keim. Visinreport: Complementing visual discourse analytics through personalized insight reports. *IEEE Transactions on Visualization and Computer Graphics*, 2021. doi: 10.1109/TVCG.2021.3104026

[19] Simon-Gd. Simon-gd/ubertagger: Exploratory analysis and annotation

of sequntial data.

[20] J. Stasko, C. Görg, and Z. Liu. Jigsaw: supporting investigative analysis through interactive visualization. *Information visualization*, 7(2):118–132, 2008.

[21] E. R. Tufte. The visual display of quantitative information. *The Journal for Healthcare Quality (JHQ)*, 7(3):15, 1985. doi: 10.1097/01445442-198507000-00012

[22] J. Zhao, M. Glueck, S. Breslav, F. Chevalier, and A. Khan. Annotation graphs: A graph-based visualization for meta-analysis of data based on user-authored annotations. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):261–270, 2016. doi: 10.1109/TVCG.2016.2598543