

# **Interpretive Discourse of Data Visualizations: An Exploratory Case Study from a Middle School English Language Arts Classroom**

Victor R. Lee, Nichole Nomura  
vrlee@stanford.edu, nnomura@stanford.edu  
Stanford University

**Abstract:** We report findings from an exploratory case study in a middle-school classroom that illustrates the use of a data visualization to support English Language Arts (ELA) learning outcomes. Students used visualizations of sentence length before and after reading a short story. Students not only initially noticed features of the data visualization, but they also immediately proposed explanations for those features without being able to situate those features in the story, advancing possible causal explanations before aligning the story to the visualization. This case illustrates an additional pathway for interpretive discourse that integrates story and data visualization and suggests a larger opportunity space for interpretative data discussion research in the future.

## **Introduction**

Text-based data and visualizations built from them, especially with data of high-interest origin (popular music, novels, or social media) are appealing tools for supporting student learning about data and data science. Responding to calls to expand and integrate work with data into other subject areas (Jiang et al., 2022), we consider the inverse: how data visualizations can be used in service of learning about other topics and subject areas. We explore intersections of data visualizations of text and their potential uses in English Language Arts (ELA), which includes instruction about literature, language, and composition, and present a case of a specific sequence for how data visualizations can be used in service of ELA instruction. In this sequence, an alternative pathway for interpretive discourse illustrates how visualizations of text can mediate learning interactions in ELA, suggesting further work in the integration of data visualizations into different subject areas.

## **Analytical approach: Understanding interpretive talk around data visualizations**

Learning sciences is producing new literature related to how we teach and learn data and data science by focusing interactionally on ways that people enact conversations and discussions, including the work of Roberts & Lyon (2020), Nemirovsky et al. (1998), Axelrod & Kahn (2023), Lee et al. (2021), and Philip et al. (2016), although the pedagogy in ELA around such visualizations has been relatively understudied.

In ELA contexts, discussion and conversation is one of the primary pedagogical moves, and we therefore turn to Axelrod and Kahn's CSAT model (2023) for our analysis of both data science and ELA learning. Three types of discursive actions constitute CSAT – Co-construct, Situate, and Advance a Theory. CSAT was derived from microanalyses of conversations whereby youth and families produced geobiographies – stories of their families' migration histories (Kahn, 2020). In CSAT, co-construction and situating can happen iteratively, and this can ultimately lead to the advancing of a theory – by which Axelrod and Kahn appear to refer to an explanatory structure linking data to story. In this paper, we draw from the components of CSAT to examine interactions in a different story situation and describe how it differs.

## **Methods**

We use an exploratory case study approach (Yin, 2008) so that salient phenomena and candidate hypotheses can be generated and later tested, describing the case through the lens of the CSAT model. We use video footage from two different 7<sup>th</sup>-grade class periods, taught by the same teacher on the same day, and deployed video data collection and review techniques like those described in Derry et al., (2010). Given the relative brevity of this paper format, we provide illustrative excerpts of broader utterance types that we emphasized in the course of video and transcript review. These utterances made reference to specific points, shapes, trends, colors, or regions in a data visualization like those named in Börner et al (2019).

## **Classroom activity context**

This seventh-grade teacher was emphasizing story arcs that consist of rising tension and techniques by which authors create suspense, using was the short story, "The Lottery" by Shirley Jackson (1948). In a long-building story arc, it is ultimately revealed that the "winner" of the lottery is stoned to death by the rest of the town.

## Results

### Before reading the short story

Before students read the story, their interpretative work involved noting features of the data visualization and then offering rationales in terms of how stories *in general* are structured and how authors might use different length sentences for different purposes. Using a “notice and wonder” routine, students noticed salient features of the visualizations and drew on prior knowledge about literature to wonder. Student A first talked about segments, suggesting they referring to a data point as a “segment” or “section”. Students B and C were discussing tendencies or patterns. For C, it was the “up and then down” and then “it” starts to even out.

When sharing what they noticed with the whole class, students began to offer conjectures about the structure of the text. For example, during the discussion, Student D offered that the reason there were more words in sentences was during the exposition portion of the story arc, and it was how the author was introducing characters and adding details:

D: I think the reason why there's more words in each sentence during the exposition, I think that is trying to introduce the characters and that's why they want to add so much [sic] details. Making the number of words kind of longer.

### Following completion of the short story

After reading, we saw students situate their previous data feature observations in the story. Student H situated the events of Tessy walking and the townspeople talking as dialogue near the end of the story, which were the reasons for shorter sentences and lower values in the data visualization. While situating some features from the visualization, others remained as questions and points for further discussion: H remained “confused why there was a sudden jump” in reference to an increased value in sentence length near the middle of the story.

H: Well, I guess, when Tessy was talking towards the end, like her, it's a dialogue, was much shorter. But also, yeah, I also, I was very confused why there was a sudden jump.

## Discussion

In their presentation of CSAT, Axelrod and Kahn found data interpretation involves processes of co-construction and situating of a story to ultimately generate an explanation/theory of what was happening that was reflected in the data. Their case involved geospatial data and family geobiographies, and they use “situating” to specifically mean the finding of people they are related to in the dataset. In a different domain (ELA) with different affordances, we observed a teacher facilitating a different sequence of events in the examination of a data visualization before and after reading the story the visualization was based on. At the start, the students not only initially noticed features of the data visualization, but they also immediately proposed explanations for those features without being able to situate those features in the story, instead advancing tentative theories with possible causal explanations based on generalized knowledge about stories. Later, they were able to make connections to the story and situate the data visualization features that they had noted into the events and composition of the story, while also situating elements of the story in the data. In doing so, they were able to maintain a dual focus on the comprehension of the story and the authorial craftwork in the sentence structure and story arc.

This observation of a different sequence of interpretation from data visualization in CSAT's terms expands Axelrod & Kahn's model of data visualization discourse—expanding of their model of interpretive discourse to a different situation with different stories and different degrees of familiarity with the story. In their work, they started with students working from prior knowledge of the family migration stories. In the present work, the students were working without prior story knowledge and thus had graphical features and general knowledge or ideas about stories as they were learning in class to rely upon. This was a product of how the sequence of events were structured in the classroom, as noticing and wondering were encouraged before the story was read. It suggests that there may be multiple pathways that could place different initial emphases for students' interpretative work as revealed in their utterances.

Ultimately, as we explore data visualization integration and interpretation in multiple settings, we can appreciate that the pedagogical opportunities and sense-making processes may vary from one another in observable ways, and future work can explore more about the nature, occurrence and affordances of these differences and how sequencing, story type, and learning goals interplay to structure those variations.

## References

- Axelrod, D., & Kahn, J. (2023). Families' discursive practices in data discussions about migration histories. *Discourse Processes*, 60(6), 457-477. doi:10.1080/0163853X.2023.2225981
- Börner, K., Bueckle, A., & Ginda, M. (2019). Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. *Proceedings of the National Academy of Sciences*, 116(6), 1857-1864. doi:10.1073/pnas.1807180116
- Chao, J., Ellis, R., Jiang, S., Rosé, C., Finzer, W., Tatar, C., . . . Wiedemann, K. (2024). Exploring Artificial Intelligence in English Language Arts with StoryQ. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(13), 15999-16003. doi:10.1609/aaai.v37i13.26899
- Curcio, F. R. (1987). Comprehension of Mathematical Relationships Expressed in Graphs. *Journal for Research in Mathematics Education JRME*, 18(5), 382-393. doi:10.5951/jresmetheduc.18.5.0382
- D'Ignazio, C., & Bhargava, R. (2016). DataBasic: Design principles, tools and activities for data literacy learners. *The Journal of Community Informatics*, 12, 83-107.
- Jiang, S., Lee, V. R., & Rosenberg, J. M. (2022). Data science education across the disciplines: Underexamined opportunities for K-12 innovation. *British Journal of Educational Technology*, 53(2), 1073-1079. doi:10.1111/bjet.13258
- Kahn, J. (2020). Learning at the Intersection of Self and Society: The Family Geobiography as a Context for Data Science Education. *Journal of the learning sciences*, 20(1), 57-80. doi:10.1080/10508406.2019.1693377
- Lee, V. R., Drake, J., Cain, R., & Thayne, J. (2021). Remembering what produced the data: Reflective reconstruction in the context of a 'quantified self' elementary data and statistics unit. *Cognition & Instruction*, 39(4), 367-408. doi:10.1080/07370008.2021.1936529
- Nemirovsky, R., Tierney, C., & Wright, T. (1998). Body motion and graphing. *Cognition and Instruction*, 16(2), 119-172.
- Philip, T. M., Olivares-Pasillas, M. C., & Rocha, J. (2016). Becoming Racially Literate About Data and Data-Literate About Race: Data Visualizations in the Classroom as a Site of Racial-Ideological Micro-Contestations. *Cognition and Instruction*, 34(4), 361-388. doi:10.1080/07370008.2016.1210418
- Roberts, J., & Lyons, L. (2020). Examining Spontaneous Perspective Taking and Fluid Self-to-Data Relationships in Informal Open-Ended Data Exploration. *Journal of the learning sciences*, 20(1), 32-56. doi:10.1080/10508406.2019.1651317
- Sinclair, S., Ruecker, S., & Radzikowska, M. (2013). Information Visualization for Humanities Scholars. In K. M. Price & R. Siemens (Eds.), *Literary Studies in the Digital Age*. Modern Language Association of America.
- Yin, R. K. (2008). *Case Study Research: Design and Methods*. Thousand Oaks, CA: Sage Publications.

## Acknowledgments

This work was supported in part by funding from the National Science Foundation under Grant No. 2241483. The opinions expressed herein are those of the authors and do not necessarily reflect those of the National Science Foundation. The authors thank Sarah Levine, Christine Bywater, Deepak Varuvel Dennison for their assistance.