



Selected Papers of #AoIR2025:
The 26th Annual Conference of the
Association of Internet Researchers
Niterói, Brazil / 15 – 18 Oct 2025

DATA, SENSE, AND SENSIBILITY: HOW DATA JOURNALISM STYLE SHAPES INTERACTIVITY

Avner Kantor
University of Haifa

Sheizaf Rafaeli
Shenkar College and Samuel Neaman Institute

Data journalism (DJ) aims to enhance audience comprehension and engagement by integrating statistical information, information sources, data visualizations, and journalistic style (Weber et al., 2018). To make data accessible to a broad audience, DJ may balance precision and engagement, aiming not only for comprehension but also to encourage exploration and interaction (Hao et al., 2024). This raises the question of whether data stories should prioritize analytical clarity or emotional resonance. To address this question, the study examines the role of journalistic style in audience interactivity.

Studies on the *Guardian* and the *New York Times* (NYT) indicate that DJ stories generate higher audience engagement than traditional news formats, with statistical information, information sources and data visualization playing a key role in this effect (Kantor & Rafaeli, 2021, 2025). However, these studies also suggest that while DJ attracts a subset of highly engaged users, its appeal to the broader public remains limited.

Journalistic style plays a crucial role in shaping audience engagement, influencing how information is perceived and interacted with (Weber et al., 2018). While an analytic style, emphasizing precision and logical reasoning, is central to professional journalism, a more affective style, which appeals to emotions, is often associated with tabloid journalism. Both styles shape audience interactivity, but their specific impact in the context of DJ has yet to be fully explored.

Using computational text analysis and mediation modeling, this study investigates three key questions. First, what journalistic styles do data journalists adopt? Do they employ an analytic approach that emphasizes precision and comprehension, or do they connect data to everyday experiences to foster engagement? Second, how do these stylistic choices influence audience interactivity? Third, to what extent is their role compatible

Suggested Citation (APA): Kantor, Avner and Rafaeli, Sheizaf. (2025, October). *Data, sense and sensibility: How data journalism shapes interactivity*. Paper presented at AoIR2025: The 26th Annual Conference of the Association of Internet Researchers. Niterói, Brazil: AoIR. Retrieved from <http://spir.aoir.org>.

with the analytic and affective effects of data visualization? These questions are examined using a model that allows for assessing their role and contribution.

Methodology

This study analyzes data collected from the NYT, a primary source of public information with a dedicated section for DJ. Stories and their associated comments were gathered using the NYT Community API and a web scraper. The final dataset includes 6,400 news stories and over 1.6 million comments, published between 2014 and 2022 across various topics, with DJ stories comprising 48.2% of the sample. A total of 137 stories with interactive visualizations were excluded because their comment section buttons were located in a different area of the webpage, which could introduce bias into the analysis.

The variables were computed using various methods, with their definitions omitted as they fall outside the scope of this paper. Table 1 presents the descriptive statistics. DJ stories were identified using the Upshot label provided by the NYT, which distinguishes DJ from traditional reporting. Text analysis was performed using the Linguistic Inquiry and Word Count tool and its built-in dictionaries. Analytic style was measured using the *analytic* dictionary, which captures logical and formal thinking (Pennebaker et al., 2022). Affective style was assessed using the *affect* dictionary, which identifies both positive and negative emotions. Static visualizations, identified through HTML feature crawling and quantified via web analysis, were found in 1,443 stories comprising 22.5% of the dataset.

Table 1. Descriptive statistics

	Min	Max	<i>M</i>	<i>SD</i>
Analytic style	12.9	99.0	91.9	6.4
Affective style	0.0	11.3	4.1	1.3
Static visualizations	0.0	14.0	0.4	0.9
Comment number (log)	2.1	8.1	4.4	1.1
Conversation	0.0	3.5	0.8	0.4

Interactivity was measured in two ways. Comment number represents the level of user interaction with the story content and was measured by the number of top comments per story (Ksiazek et al., 2016). Conversation represents user-to-user interactivity and was calculated as the ratio of responses to comments for each story. Control variables included story topic, year of publication, and story length in words.

The unit of analysis is the story, with all dependent variables aggregated at the story level. Since comment volume exhibited a skewed distribution, it was log-transformed for normalization. Each story topic was dummy-coded as a separate independent variable.

Statistical analysis was conducted using Hayes' PROCESS macro Model 80 (Hayes, 2022). The model included three parallel mediation paths connecting DJ to conversation, each with two serial mediators. The analysis was applied to 5,000 bias-corrected bootstrap samples with 95 percent confidence intervals. A significant effect (p

< .05) was determined when both the lower and upper confidence interval bounds did not include zero.

Results

Figure 1 shows that DJ stories tend to use a less analytic style and rely more on affective language and static visualizations. Analytic style and static visualizations are both positively associated with conversation, whereas affective style is negatively associated. DJ stories are associated with fewer comments overall, but when comments are present, they are more likely to develop into conversations, as comment number is positively related to conversation. Additionally, analytic style is associated with fewer comments, while affective style corresponds with a higher number of comments. Static visualizations, however, do not significantly affect comment number.

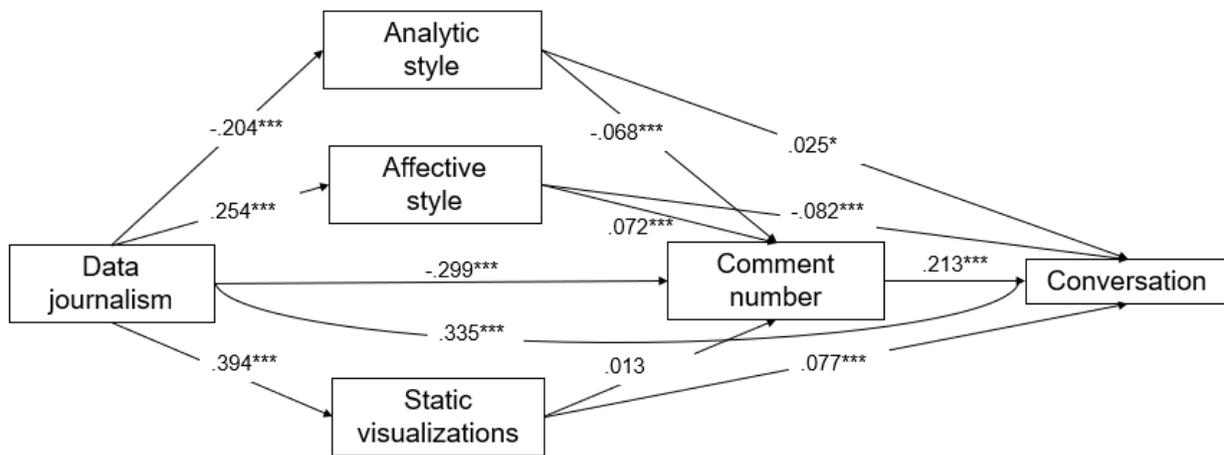


Figure 1. Results of multiple serial mediation analysis
 $*p < .05$; $**p < .01$; $***p < .001$; all effects are standardized

Table 2 shows that DJ has a direct positive effect on conversation. The mediation effects differ by style: both analytic and affective styles negatively mediate conversation, whereas static visualizations positively mediate it. Comment number negatively mediates conversation, meaning that higher comment number does not necessarily lead to richer discussions. Analytic and affective styles both exhibit serial mediation effects, though their pathways differ. The effect of static visualizations on conversation is not significant in this context.

Table 2. Total, direct and indirect model effects

Model pathways	Effect	SE	LLCI	ULCI
Total effect	.118	.012	.094	.142
Direct effect	.139	.012	.115	.163
Partially standardized indirect effects				
Total	-.052	.010	-.072	-.032
DJ - Analytic style - Conversation	-.005	.003	-.011	-.000
DJ - Affective style - Conversation	-.021	.004	-.029	-.014
DJ - Static visualizations - Conversation	.030	.006	.020	.042
DJ - Comment number - Conversation	-.064	.007	-.078	-.051
DJ - Analytic style - Comment number - Conversation	.003	.001	.002	.005
DJ - Affective style - Comment number - Conversation	.004	.001	.003	.006
DJ - Static vis. - Comment number - Conversation	.001	.001	-.001	.003

Discussion

The results indicate that NYT DJ employs a less analytic and more affective style compared to traditional journalism. This pattern aligns with the increase in comment number, suggesting that data journalists may adopt their style to reach a broader audience.

The relationship between journalistic style and conversation shows that analytic style and static visualizations are positively associated with conversation, while affective style is negatively associated. The mediation results reveal a complex interplay between analytic processing and emotional engagement in shaping conversation. While both analytic and affective styles negatively mediate conversation, they contribute positively through serial mediation via comment number. This suggests that, despite shifts in analytic and emotional processing, participation is sustained. However, while static visualizations directly contribute to conversation, they do not initially drive audience engagement. This insight is relevant to data journalists and other fields that convey complex content, such as education.

These findings also indicate that the subset of highly engaged users identified in previous studies (Kantor & Rafaeli, 2021, 2025) emerges due to their distinct conversational preferences compared to the broader public, suggesting that analytic story content encourages more in-depth conversation.

While this study represents a single case, further research should examine similar patterns across different newspapers, platforms, and additional factors not explored here, such as the role of story topic. Nonetheless, these findings present a challenge for news organizations: fostering engagement through emotionally resonant storytelling may come at the cost of reducing in-depth conversation, whereas an emphasis on analytic style and static visualizations may appeal to a smaller but more engaged audience. This highlights the ongoing need for data journalists to refine their storytelling style with sense and sensibility, continually reflecting on and exploring new approaches to balance engagement and depth (Bhargava, 2023).

Data availability

The data are available at the Open Science Framework: <https://osf.io/m6p9q/>.

References

- Bhargava, R. (2023). Data representation as epistemological resistance. *AoIR Selected Papers of Internet Research*.
- Hao, S., Wang, Z., Bach, B., & Pschetz, L. (2024). Design patterns for data-driven news articles. *Conference on Human Factors in Computing Systems - Proceedings, March*.
- Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (3rd ed.). Guilford publications.
- Kantor, A., & Rafaeli, S. (2021). Independence through data journalism. *AoIR Selected Papers of Internet Research*.
- Kantor, A., & Rafaeli, S. (2025). Conversations with data: How data journalism affects online comments in the New York Times. *Proceedings of the 58th Hawaii International Conference on System Sciences*, 8, 2386–2395.
- Ksiazek, T. B., Peer, L., & Lessard, K. (2016). User engagement with online news: Conceptualizing interactivity and exploring the relationship between online news videos and user comments. *New Media and Society*, 18(3), 502–520.
- Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., & Booth, R. J. (2022). The development and psychometric properties of LIWC-22. *LIWC-22 Development Manual*, 1(2), 1–22.
- Weber, W., Engebretsen, M., & Kennedy, H. (2018). Data stories. Rethinking journalistic storytelling in the context of data journalism. *Studies in Communication Sciences*, 18(1), 191–206.