Towards Understanding the Impact of Guidance in Data Visualization Systems for Domain Experts

Sherry Qiu, Holly Rushmeier, Kim R.M. Blenman

Yale New Haven Health Smilow Cancer Hospital

Yale CANCER CENTER — A Comprehensive Cancer Center Designated by the National Cancer Institute

Abstract

Guided data visualization systems are highly useful for domain experts to highlight important trends in their large-scale and complex datasets. However, more work is needed to understand the impact of guidance on interpreting data visualizations as well as on the resulting use of visualizations when communicating insights. We conducted 2 studies with 11 proteomics domain experts consisting of end user biologists, computational biologists, and clinicians. We found that experts benefit from a guided coarse-to-fine structure when using data visualization systems, as this is the same structure in which they communicate findings.

Figure 1. visAPPprot interface for generating proteomics visualizations: Volcano Plot, Pathway Map, Overview Heatmap, and WGCNA clustering figures.

Study Details

Study 1: We divided participants into a guided and an unguided cohort. The guided cohort received instructions with labeled diagrams directing them to complete tasks and view the visualizations in a coarse-to-fine order: *1) Volcano Plot, 2) Pathway Map, 3) Overview Heatmap, 4) WGCNA clustering on all data, 5) WGCNA clustering on subset*. The unguided cohort received no visualization order.

Study 2: Participants were asked to order the visualizations to tell a story about their findings and list insights they gained into the dataset for each visualization.
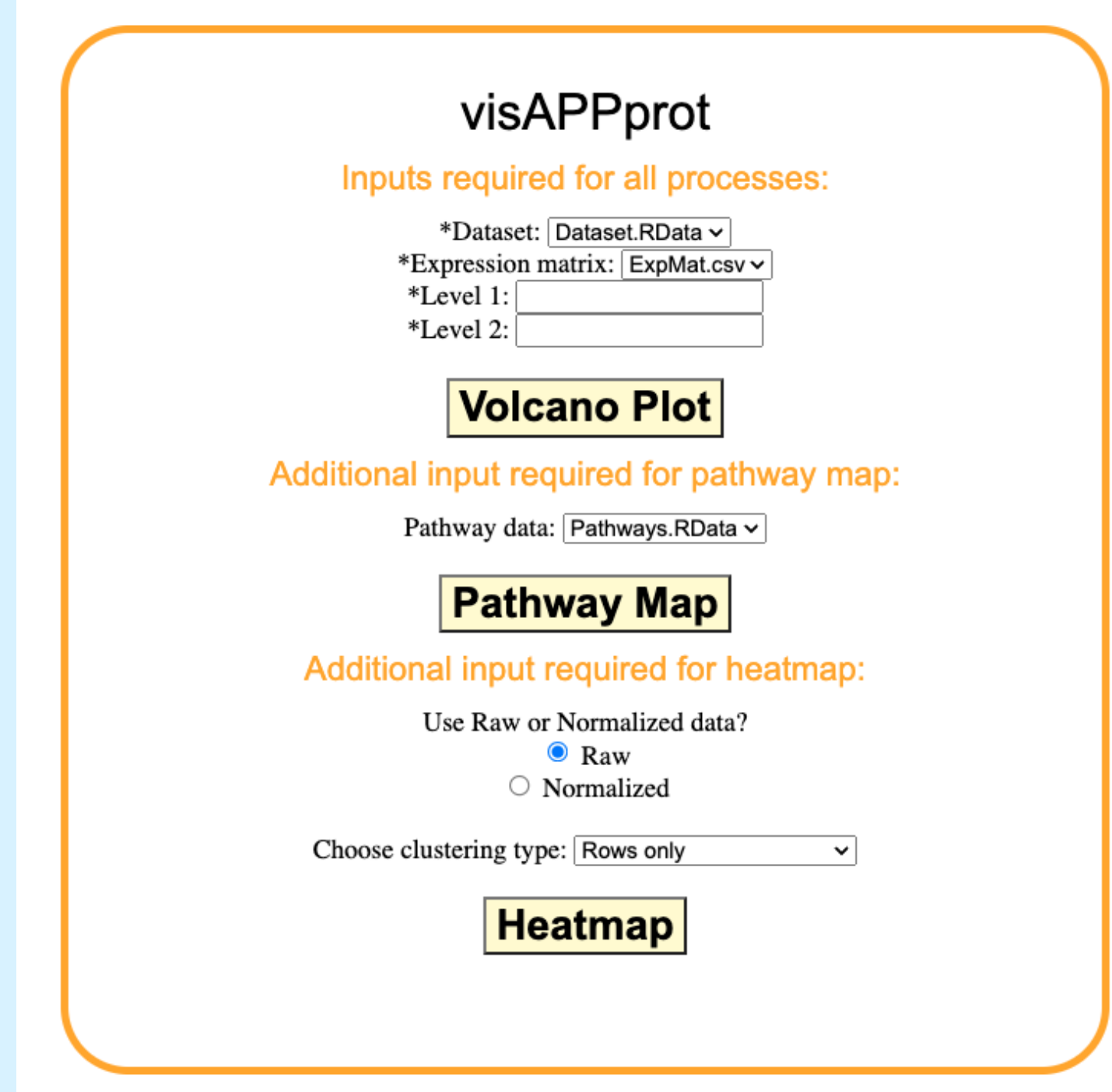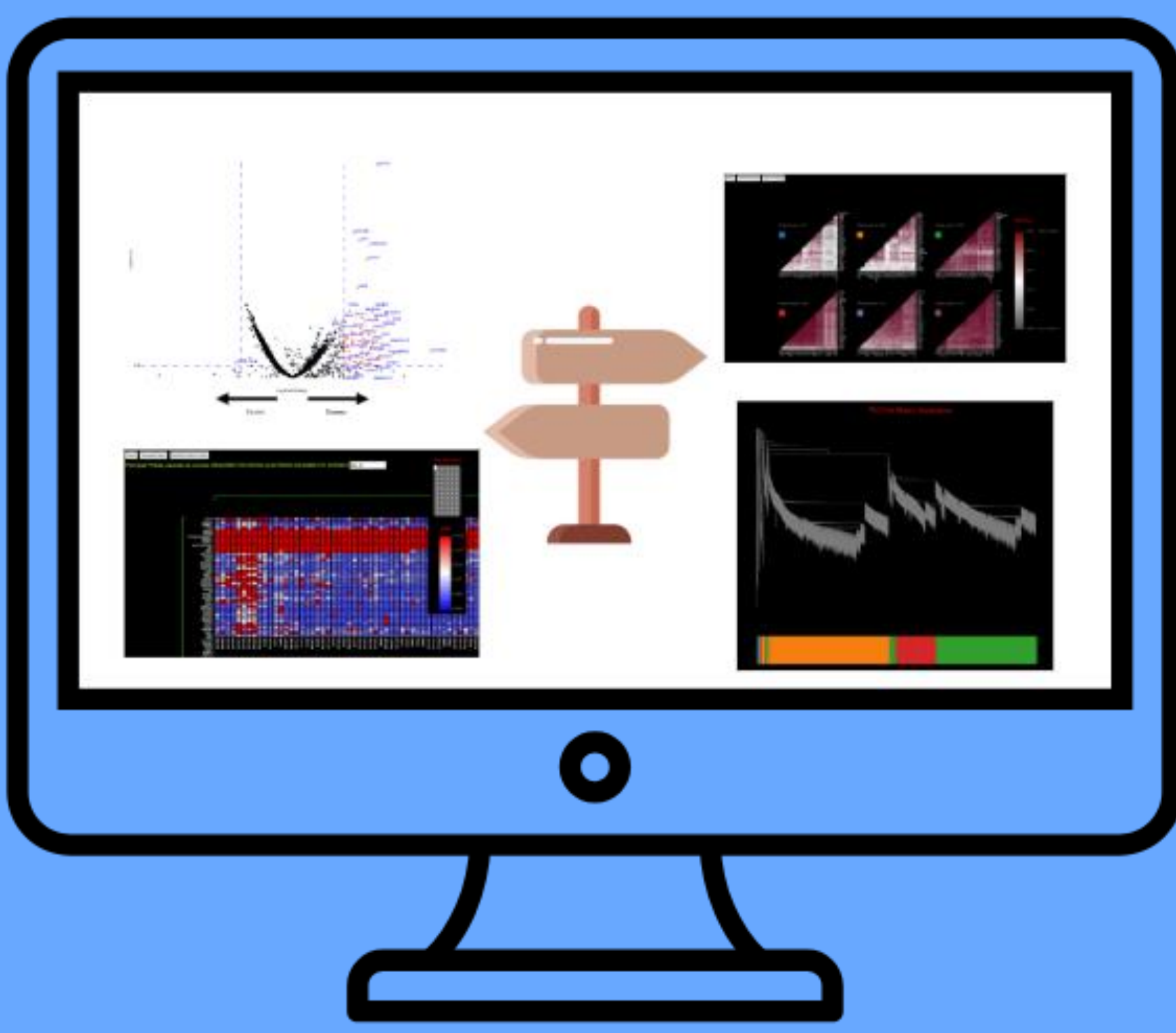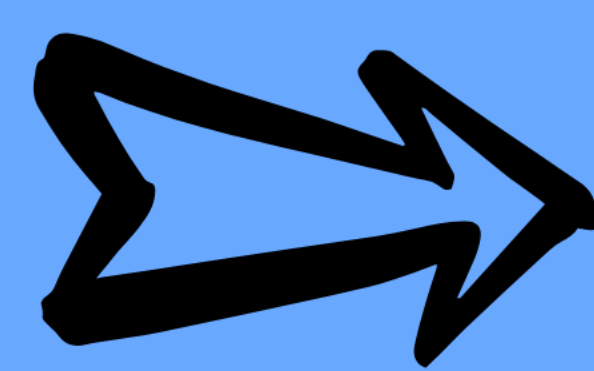
Guided Analysis → Coarse-to-fine Storytelling

Key Findings

Guidance is needed to orient users in detecting high level trends.
Over half of the unguided participants tried to view the entire heatmap quickly through attempting bidirectional navigation of the marker in the navigation map to scroll the heatmap, while none of the guided participants exhibited this behavior.

Guidance is needed to orient users in exploring detailed information.
The unguided cohort took longer to view the subset figures and lacked direction in their exploration compared to the guided cohort.

Visualization dynamicism and interactivity are more effective for displaying large datasets compared to the static domain standard visualizations in current practice.
The Module Heatmap displayed the same data as the Overview Heatmap but was organized by color-coded modules, making the Module Heatmap more intuitive than the Overview Heatmap.

Participants communicated their findings in a coarse-to-fine structure.
Participants wanted a "big picture" and "broad view of all data" for the first visualization. For latter visualizations, participants wanted to show "relationships between proteins" and "groups of proteins".

Coarse-to-fine guidance is effective for interpreting visualizations and subsequent storytelling of insights. We reveal a need for a continuous dialogue between domain experts and developers to create intuitive guided visualization systems.
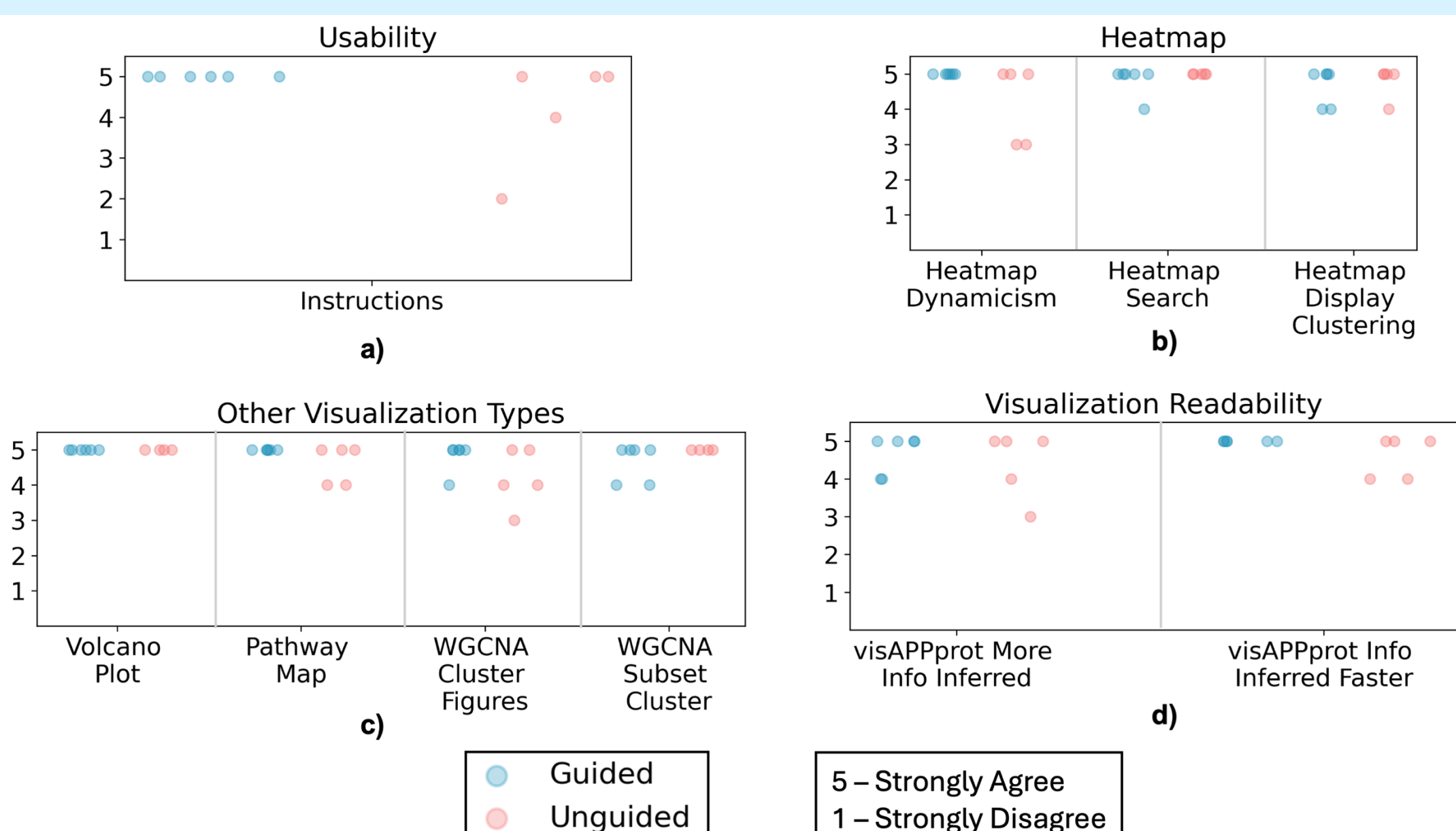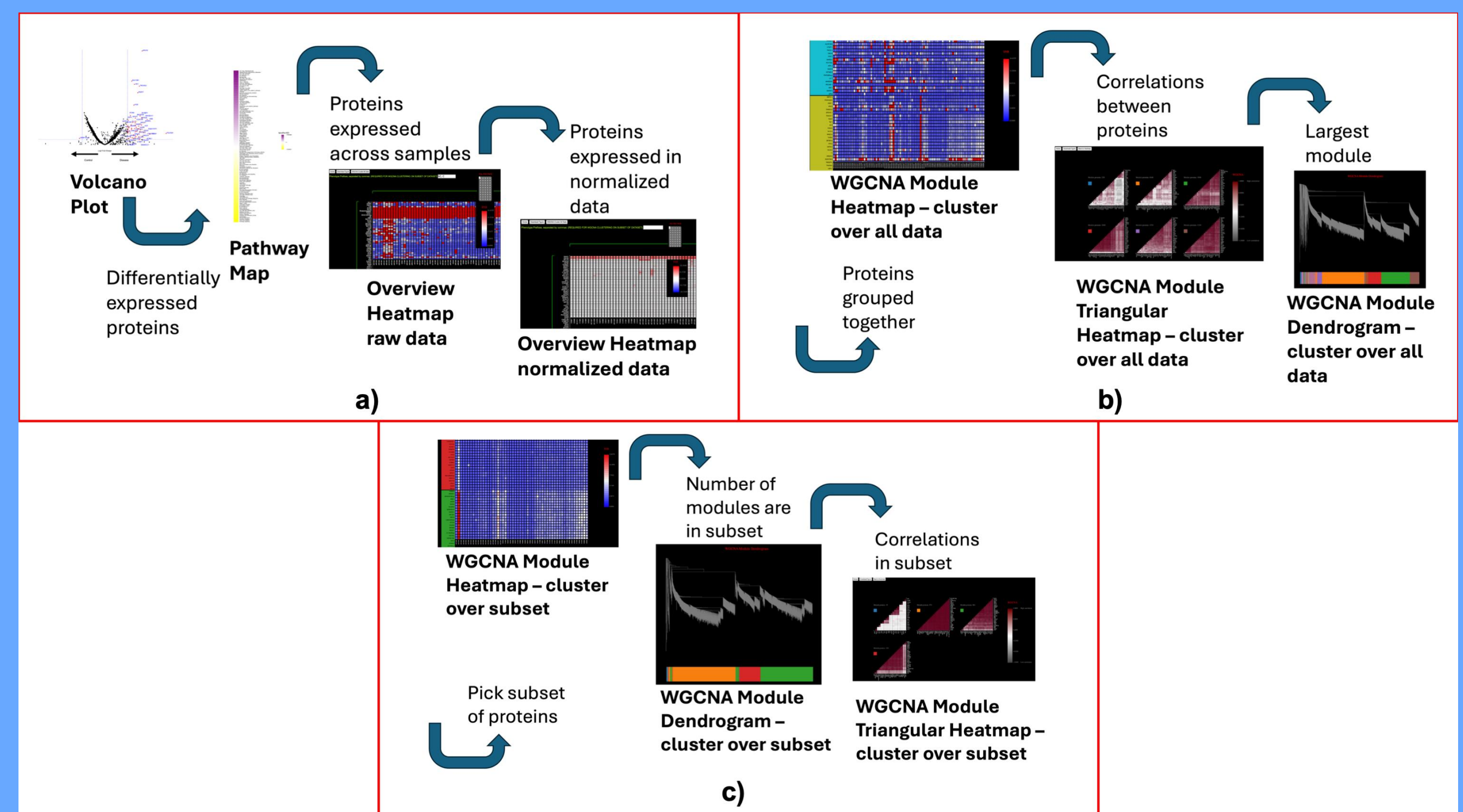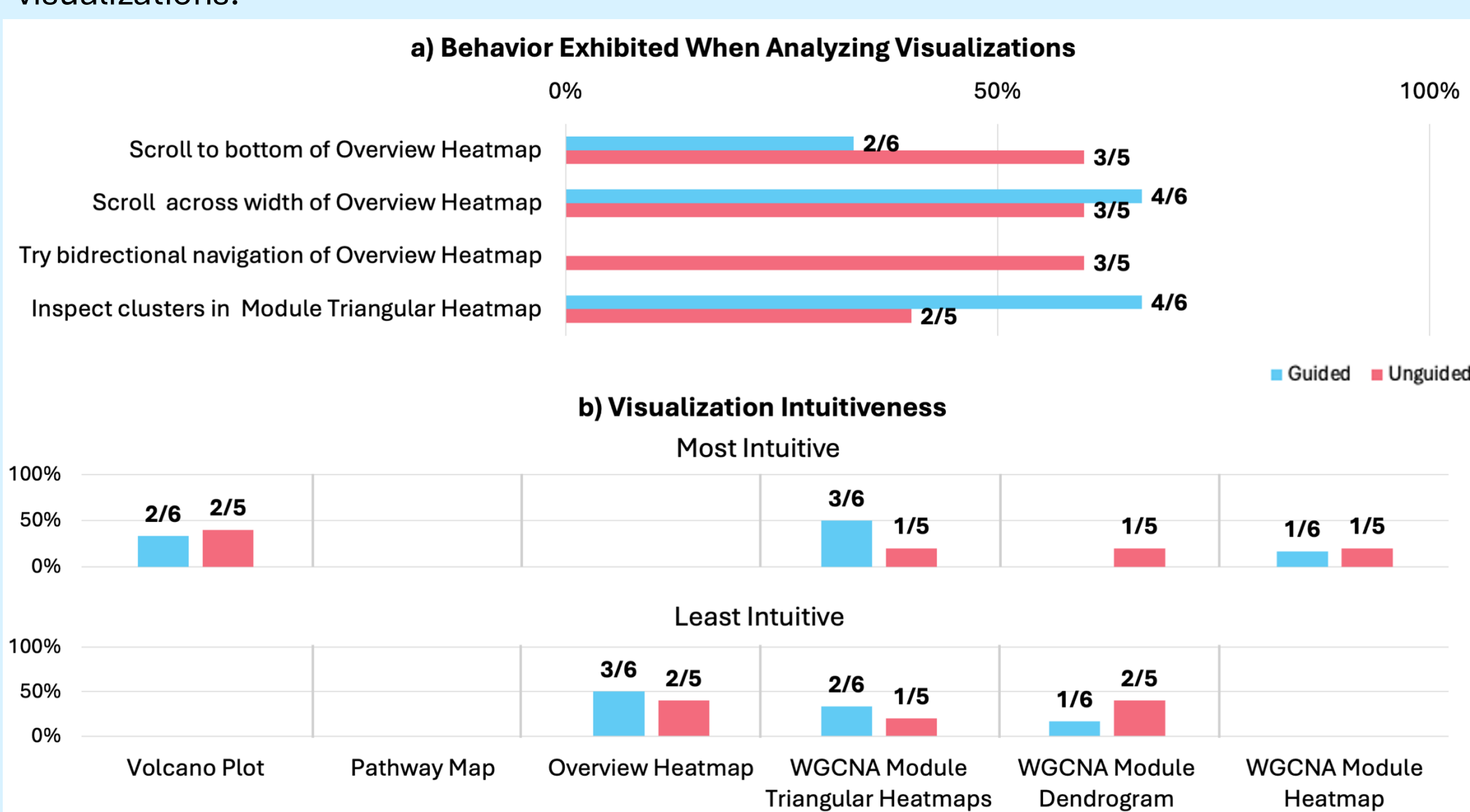

Figure 2. Visualization interpretation by guided and unguided cohort, separated by category and plotted by individual participant results.

Figure 3. Top: Behaviors exhibited by participants. Bottom: Most and least intuitive visualizations.

Contact

Sherry Qiu
Yale University
Email: sherry.qiu@yale.edu

Kim R.M. Blenman
Yale University
Email: kim.blenman@yale.edu

References

[1] D. Ceneda, T. Gschwandtner, T. May, S. Miksch, H.-J. Schulz, M. Streit, and C. Tominski. Characterizing guidance in visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):111–120, 2017. doi: 10.1109/TVCG.2016.2598464

[2] V. Filipov, D. Ceneda, D. Archambault, and A. Arleo. Timelighting: Guidance-enhanced exploration of 2d projections of temporal graphs. In *International Symposium on Graph Drawing and Network Visualization*, pp. 231–245. Springer, 2023

[3] P. Langfelder and S. Horvath. WGCNA: an r package for weighted correlation network analysis. *BMC bioinformatics*, 9:1–13, 2008.

[4] R.A. Leite, A. Arleo, J. Sorger, T. Gschwandtner, and S. Miksch. Hermes: Guidance-enriched visual analytics for economic network exploration. *Visual Informatics*, 4(4):11–22, 2020.

[5] J. Li, C. Lai, Y. Wang, A. Luo, and X. Yuan. Spectrumva: Visual analysis of astronomical spectra for facilitating classification inspection. *IEEE Transactions on Visualization and Computer Graphics*, 2023.

[6] M. Meuschke, U. Niemann, B. Behrendt, M. Gutberlet, B. Preim, and K. Lawonn. Gucci - guided cardiac cohort investigation of blood flow data. *IEEE Transactions on Visualization and Computer Graphics*, 29(03):1876–1892, mar 2023. doi: 10.1109/TVCG.2021.3134083

[7] B. Preim and K. Lawonn .A survey of visual analytics for public health. In *Computer Graphics Forum*, vol. 39, pp. 543–580. Wiley Online Library, 2020.

[8] B. Shneiderman. The eyes have it: a task by data type taxonomy for information visualizations. In *Proceedings 1996 IEEE Symposium on Visual Languages*, pp. 336–343, 1996. doi: 10.1109/VL.1996.545307

[9] C. Stoiber, D. Ceneda, M. Wagner, V. Schetinger, T. Gschwandtner, M. Streit, S. Miksch, and W. Aigner. Perspectives of visualization onboarding and guidance in va. *Visual Informatics*, 6(1):68–83, 2022. doi: 10.1016/j.visinf.2022.02.005 1