

Against tasks.

Matthew Kay*

Northwestern University

A prodigious volume of research is dedicated to determining how “well” people perform visualization “tasks”, such as “retrieving values”, “identifying clusters”, or “identifying outliers”. Typical task taxonomies paradoxically make tasks specific enough to measure but not specific enough for those measurements to generalize. “Retrieve value”, purportedly a low-level task, sounds specific enough. Yet it is usually measured in data units we have no reason to believe should generalize to other datasets, viewing conditions, or chart types (“retrieve value” performance on a scatterplot tells me nothing about that same “task” on a bar chart or pie chart). Ignoring cross-chart generalization, we should at least expect tasks to carry a meaningful theoretical structure that yields generalization within-chart. Why can’t I use knowledge of how good “retrieving values” is on a scatterplot to predict how good “find extremum” is? Or “identify clusters”? Presumably those tasks involve also retrieving values. Yet even a “low-level” task like retrieve value is underspecified: it forms no part of a coherent structure determining how tasks relate to each other or how they might compose.

Tasks are the structureless relationshipless objects of a naive empiricism: visualization researchers, terminally task-brained, act as primitive taxonomists, collecting and classifying “task” specimens without building deeper knowledge or theory. In VIS reviewing season we show off our specimens in overstuffed Victorian menageries called “papers”. Observe my 6 hypotheses based on flimsy verbal theory; behold my wondrous full-page 30-row coefficient plot; gaze upon the Cartesian product of my 7 tasks and the 4 visualizations I pulled out of my ass 30 minutes before submitting my preregistration – *I fucking dare you* to form a single succinct, coherent conclusion from the Baroque mess of my results section.¹

Maybe these task menageries would be more helpful if isolated tasks told us something about what makes a visualization “good”. But we don’t use visualizations to retrieve a single value, or look for a single cluster, or whatever. A visualization is not a slow, serial sequence of task-facts. A visualization is a magical parallel pipe of information and structure directly into the front of your brain. As teachers we know this even if as researchers we don’t: Vis classes commonly start with Anscombe’s quartet, a visual example the point of which is that visualizations are *more useful to people* than single statistics or data tables (from which one could more accurately “retrieve values”). We use visualizations to understand the *structure* of data, to understand complex data relationships *together* in ways we otherwise could not.

A good visualization not only helps us understand data but also helps us understand *the right way* to understand data: it uses its visual form to push us towards visual operations that correspond to valid data operations and away from visual operations that do not. An adherent of algebraic visualization design [7] might call good visualizations structure-preserving maps from data (and data operations) to visual representations (and perceptual operators [8]). A more poetically-inclined designer might say a visualization carries

its own visual *language*: semantics and metaphor that tell us what the data means and what we can or cannot infer from it. A good visualization, in other words, is not just a structure-*preserving* map, it is a structure-*communicating* map.

Studying tasks in isolation tells us little about how well a chart communicates structure. Consider a Treemap [5], which retrieve-value-ists might insist is a bad visualization: it maps all values onto *area*, a perceptually horrible encoding! Meanwhile, the *raison d’être* of a Treemap is that it enacts a structure-preserving map: the data is a tree where the value of sibling nodes *sum* to the value of their parent; the visualization depicts areas for each node that *visually sum* to the area of their parent. What’s interesting about a Treemap lies not in asking “can people retrieve values from it?” but “do they understand the structure and relationships it depicts?”

On the other hand, consider the bar chart, which retrieve-value-ists might insist is a good default visualization since it uses a positional encoding: one could Cleveland-and-McGill [1] untold numbers [3] of new participants [4] and keep rediscovering [2] that they “retrieve” bar chart values better than anything. Meanwhile if you show people a bar chart of a mean and ask them to draw the underlying data, a bunch of them will draw all the data points within the bar [6], a mathematical impossibility! Why? A bar’s visual form suggests an accumulation of values up to its tip; it is the wrong *visual metaphor* for a mean. Task-brained visualization researchers, forgetting to ask people what they see, spent decades asking people to “retrieve values” from bars; noting that values were biased downward, they posited the existence of a “within-the-bar bias”. This early “bias” presaged countless papers on “visualization biases”, giving names to countless additional underexplained “biases” we purportedly possess. Yet no matter how many biases we discover, each new chart permutation yields another new bias; like tasks, we collect them but do not understand them. Our brains do not contain a bajillion bias knobs, one for bias when bars are upright, upside down, sideways, grouped or ungrouped, close or distant. There must be a simpler, unifying explanation; if we could just *stop collecting specimens* long enough to look for it.

As someone who is (I think) a decent visualization designer, and who has also collected many a task in his papers, my great disappointment is that my research informs so little of my design. Visualization design is so combinatorial and visualization research has so few mechanisms of generalization that if a minor design variation does not have a paper about *exactly that thing*, you’re shit out of luck. Even if it does, the effects of different encodings on whatever task can easily be dwarfed by people just failing to understand the *meaning* of a visualization; i.e., the stuff we don’t have good theory about. What theories we have either make precise predictions about overly-specific combinations of task $\times$ chart $\times$ context or are too high-level to make precise predictions about anything at all.

We need precise, compositional theories that can predict measurements of human understanding on never-before-seen visualizations. We need theories grounded in generalizable visual operations people can do, what visual forms and metaphors suggest what operations, and how those operations correspond to data-space operations – theories that use those operations to predict the direction, magnitude, and variance of people’s responses. We need experiments that adjudicate between those theories, not between chart type performance on some task or another. It will be hard work, but the first step is clear: we need to give up tasks.

*e-mail: matthew.kay@u.northwestern.edu

¹The next time I am asked to review a paper with a Hypothesis 5(c) I will burn PCS to the goddamn ground.

REFERENCES

- [1] W. S. Cleveland and R. McGill. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American statistical association*, 79(387):531–554, 1984. [1](#)
- [2] R. Davis, X. Pu, Y. Ding, B. D. Hall, K. Bonilla, M. Feng, M. Kay, and L. Harrison. The risks of ranking: Revisiting graphical perception to model individual differences in visualization performance. *IEEE Transactions on Visualization and Computer Graphics*, 30(3):1756–1771, 2022. [1](#)
- [3] J. Heer and M. Bostock. Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 203–212, 2010. [1](#)
- [4] J. Hullman, E. Adar, and P. Shah. The impact of social information on visual judgments. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 1461–1470, 2011. [1](#)
- [5] B. JOHNSON. Tree-maps: a space-filling approach to the visualization of hierarchical information structures. In *VIS’91: Proceedings of the 2nd conference on Visualization’91*, pp. 284–291. IEEE Computer Society Press, 1991. [1](#)
- [6] S. H. Kerns and J. B. Wilmer. Two graphs walk into a bar: Readout-based measurement reveals the bar-tip limit error, a common, categorical misinterpretation of mean bar graphs. *Journal of vision*, 21(12):17–17, 2021. [1](#)
- [7] G. Kindlmann and C. Scheidegger. An algebraic process for visualization design. *IEEE transactions on visualization and computer graphics*, 20(12):2181–2190, 2014. [1](#)
- [8] S. Long, R. Chang, E. Wu, A. Kale, and M. Kay. Visual decoding operators: Towards a compositional theory of visualization perception. *arXiv preprint arXiv:2604.02220*, 2026. [1](#)