IAT 814 Assignment 2: Research Topics in Visualization

A Critical Précis of
Perception of Elementary Graphical Elements in Tabletop and Multi-Surface Environments, CHI 2007

For Dr. Lyn Bartram

[redacted]

OVERVIEW

As tabletop displays become more common, they are increasingly being used in contexts where quick, accurate perception of displayed data is vital. The shift from the vertical orientations of projections and monitors to the horizontal orientation of tabletops alters the perspective of graphical elements used to present data (eg. diagrams, tables, and visualizations), which in turn affects the perception of these elements. To investigate these perceptual differences, Daniel Wigdor, Chia Shen, Clifton Forlines and Ravin Balakrishnan carried out research on the effects of display orientation on the perception of graphical elements. This research was presented at CHI 2007 [7].

The research presented consists of two experiments: the first experiment examines the effects of display orientation on graphical element perception on a single display, while the second examines the perception and comparison of graphical elements across two displays. Wigdor et al base their perception task design on previous work by Bertin [2], Cleveland and McGill [3,4]. This previous research defines a set of graphical elements in which data is encoded visually as “visual variables” (Bertin) or “elementary perceptual tasks” (Cleveland & McGill). Wigdor et al note that Cleveland and McGill’s work is most relevant to their research, as it recognizes more granular graphical element categories such as length and area, where Bertin uses broader categories like size. Research by Baird [1] has shown that these more granular categories are distinct in terms of human perceptual acuity, introducing ambiguity in research where broad uses that combine multiple elements into a greater element are examined.

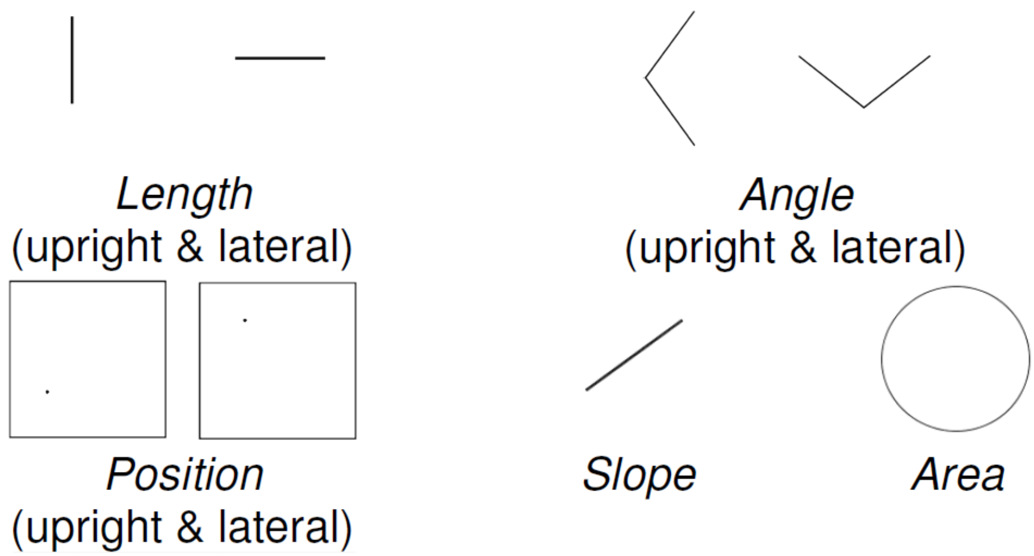

Figure 1 – Wigdor et al's Subset of Graphical Elements from Cleveland and McGill

A subset of Cleveland and McGill's graphical elements are used by Wigdor et al. This subset excludes elements such as volume and curvature that Cleveland and McGill removed from later research [4]. This set of graphical elements is best suited to evaluating human perceptual ability and accuracy "at a glance". As Tufte[6] has noted, there are multiple levels of engagement with visual data: A quick, initial, "first-look" engagement that allows a general, high-level understanding of the data, and a focused, scrutinized engagement that facilitates a more detailed, nuanced understanding. As a result, Wigdor et al do not include additional data decoding cues like numerical scales or grid lines in their testing in order to focus on "at a glance" perception.

At the task-level, Wigdor et al's research tests the accuracy of participant's perception of the relative magnitude of elements shown on display(s) in various orientations. This per-element, in-depth approach allows the authors to carefully examine which graphical elements are affected by perceptual distortion the most, and which are the most perceptually robust across different orientations.

EXPERIMENTAL DESIGN AND OUTCOMES

Both experiments performed by Wigdor et al in the course of this research measure perceptual distortion by asking participants to complete magnitude estimations on pairs of various types of graphical elements while varying display orientation. Tested display orientations include the relatively novel tabletop or horizontal orientation, which the authors feel requires clarification of terminology in order to discuss the layout of graphical elements relative to the orientation of the display.

A display in *tabletop* orientation is at 0 degrees relative to floor, with the display surface facing up. A *vertical* display is at 90 degrees relative to the floor. *Left* and *right* along the display correspond to the x axis of the display's coordinate space, while *up* and *down* correspond to the y axis. Graphical elements are *upright* when aligned along y axis, and *lateral* when aligned along x axis.

The magnitude estimations, based on tasks from Cleveland and McGill's research, require the comparison of two graphical elements from the same set – a modulus elements and a stimulus element. Participants are shown the modulus element (at the position marked **M** in Figure 2) and are asked to evaluate the relative magnitude of the stimulus element (at one of the positions marked **S** in Figure 2) as a percentage of the modulus.

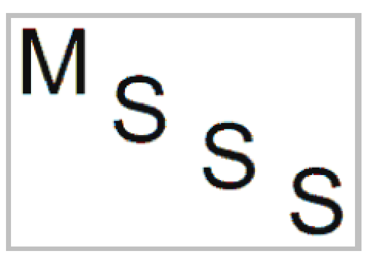

Figure 2 – Modulus and Stimulus positions from Cleveland and McGill

The participant's magnitude estimation is subtracted from the actual difference of the modulus and stimulus, giving an error value for that response. Thus, accuracy is defined as the difference between perceived magnitude and actual magnitude of the stimuli compared to the modulus, independent of task time, individual preference, or other testable parameters.

Wigdor et al note that this task design results in *relative* magnitude reports from participants. There is some debate as to whether this is an appropriate measure of perceived magnitude, as the accuracy of magnitude estimation reports usually depends on participants correctly mentally converting their perception into a numerical quantity at some point during the task. Uncertainty regarding participant's quantification abilities could confound studies that are concerned with actual vs. reported perceived magnitude. However, Wigdor et al are interested in the *change* in perceived magnitude as display orientation changes; they require only that a perceived change be reported, not that reported results precisely match a participant's actual perception.

Wigdor et al make two major changes worth noting to the Cleveland and McGill's tasks. First, a new modulus is shown along with the stimulus each time the task is repeated. Cleveland and McGills' original task compared a single modulus to multiple subsequent stimuli, which more recent research suggests may lead to an increase in increase perceptual error over time [5]. Second, they extended the original Cleveland & McGill element layout pattern by tripling the possible modulus and stimulus positions vertically, to allow for greater up/down visual distortion. The authors predict that the up/down position of graphical elements will strongly affect perceptual error in tabletop orientations due to the increased perspective distortion introduced by tabletop displays (compared to a vertical display as perceived by a standing or sitting viewer).

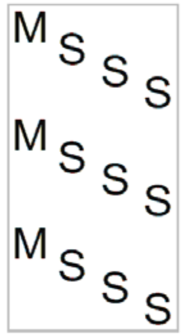

Figure 3 - Modified modulus/stimulus layout used by Wigdor et al

EXPERIMENT 1

The first experiment examines the perception of graphical elements on a single display, which is set at various angles. 12 participants (7 men, 5 women) were introduced to the graphical elements and run through a practice session with the facilitator to ensure a clear understanding of the task. Participants were then shown modulus/stimulus pairs and asked to enter the magnitude of the stimulus as a percentage of the modulus. This task was repeated throughout 2 hour long experimental sessions, with each participant evaluating 4 of the 8 possible graphical elements in matched lateral and upright pairs. During their session, participants viewed the elements in each of the 3 modulus positions and each of the 9 stimulus positions, repeated three times with the display in one of four positions: vertical, 60 degrees, 30 degrees, and tabletop. This resulted in a total of 1,296 magnitude estimations per participant, for a total of 15,552 total comparisons for Experiment 1.

A statistical analysis of error data from these comparisons yields the following outcomes:

- Magnitude estimation of the graphical elements is less accurate on tabletop orientations when elements are not presented at the same up/down distance from the user.
- The greater the up/down separation between elements on a tabletop-oriented display, the less accurate comparisons between those elements will be.

- Larger left/right distances between compared elements do not yield an increase in magnitude estimation error, with the exception of the *slope* graphical element.
- Of the graphical elements which are more robust for tabletop display, two (*position*, *angle*) require that they be presented laterally relative to the viewer in order to maintain their robustness.
- If elements must be perceived both upright and laterally, *position* is more accurate than *angle*.

Robustness is defined by lower error scores and consistent variability in perceptual error. Based on the data from Experiment 1, the order of graphical elements from most to least robust is: *length* (lateral), *length* (upright), *position* (lateral), *angle* (lateral), *area*, *position* (upright), *angle* (upright), *slope*.

EXPERIMENT 2

The second experiment examines the perception of graphical elements across two displays, with both displays set at a vertical position, or one display vertical and the other in tabletop orientation.

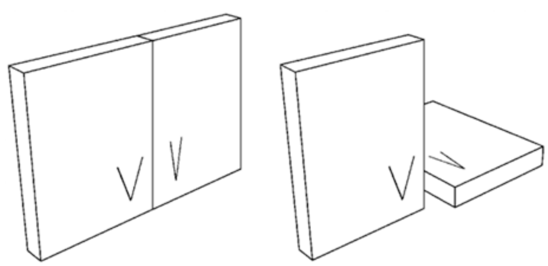

Figure 4 – Display orientations used for Experiment 2: Vertical/Vertical or Vertical/Horizontal

For Experiment 2, 8 participants (all male) were introduced to the graphical elements and run through a practice session with the facilitator. Modulus/stimulus pairs shown to the participants were distributed across both displays, with the modulus on one display and the stimulus on the other. As in Experiment 1, participants were asked to enter the magnitude of the stimulus as a percentage of the modulus. Participants evaluated all 8 of the possible graphical elements, and completed 31 magnitude estimates in both display conditions, resulting in a total of 496 magnitude estimations per participant, for a total of 3,968 total comparisons for Experiment 2.

A statistical analysis of error data from these comparisons, using the same error calculation as used in Experiment 1, yields the following outcomes:

- If possible, information visualizations should not be compared across display orientations. Error in magnitude perception across different display orientations was consistently higher than across the vertical/vertical orientation (by as much 4 times for certain graphical elements).
- If comparisons across display orientations are unavoidable, the first three in this are far better than the rest (the list is ordered by accuracy): *length (lateral)*, *length (upright)*, *position (lateral)*, *angle (lateral)*, *area*, *angle (upright)*, *position (upright)*, *slope*.
- If a graphical element is required to be perceived both upright and laterally, *angle* is more accurate than *position* across differently-oriented displays, unlike on a tabletop.

DESIGN RECOMMENDATIONS

Wigdor et al conclude with some high-level design recommendations drawn from both experiments. They recommend that when using tabletop displays, graphical elements being compared should be positioned at the same up/down position from a viewer – this minimizes distortion differences caused by perspective on tabletop displays. As length and angle graphical elements emerged as the most perceptually robust elements from the experiments, they are recommended for use in data visualizations over other elements. Widgor et al provide examples of less-robust elements that can be translated into these more robust forms (Figure 5, 6).

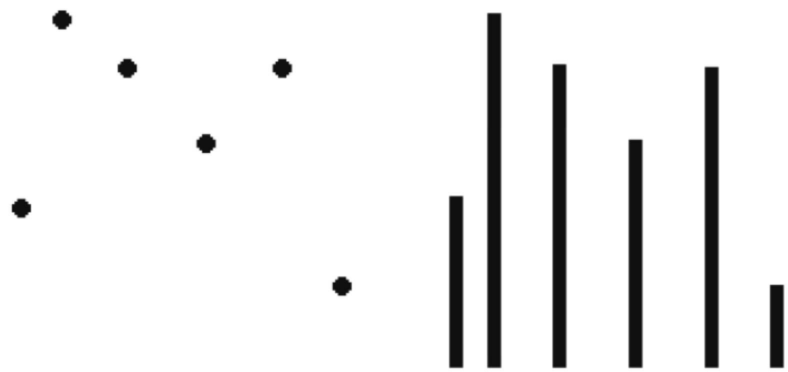

Figure 5 – Data encoded as position elements (left) shown as perceptually robust length elements (right)

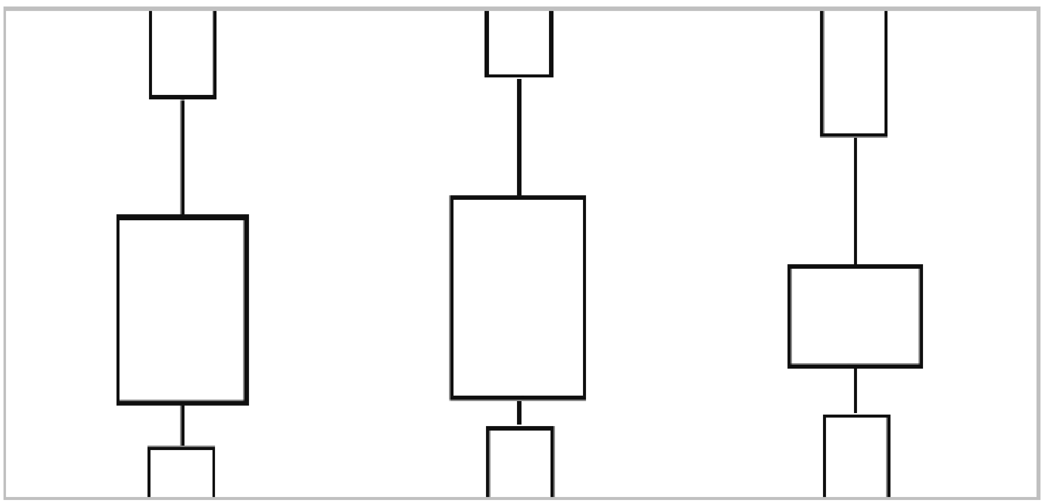

Figure 6 – A modified box plot which uses length elements instead of position for inter-quartile bounds

Finally, data encoded as area and slope elements are not recommended for use in tabletop oriented displays, as they are prone to large and highly variable errors in magnitude perception.

CRITIQUE

Overall, this research appears to be well thought-out, well grounded, and thorough. Grounding the task design in Cleveland and McGill's research strengthens the construct validity and internal validity of the author's perception measurements. As this area of perception has been subject to further study by psychologists and psychophysicists, the research can be linked to current findings and refined further as needed. The authors also seem to consider ecological validity by including a brief discussion of the unique challenges multi-user perception around tabletops present. (Their conclusion – use small multiples oriented to each user if at all possible). Altering the task to account for greater up/down distortion that tabletop displays may introduce also strengthens the ecological validity of their findings. Lastly, their design recommendations are simple and clear, facilitating easy comprehension of the implications of their results.

In their discussion, Wigdor et al discuss some shortcomings with this research and outline some areas for further study. The authors state early in the paper that although there is a significant body of psychophysical research examining the underlying mental processes used in perceiving and understand graphical elements like those used here, they have not attempted to explain their results in the context of this research, or even attempt to match their results to a known perceptual model. They suggest that this is an important next step, as it may allow researchers to create a predictive model that could be used to evaluate more complex data displays on tabletops. The authors also note that their emphasis on "at-a-glance" perception ignores the role that other graphical or perceptual cues could play in magnitude estimation or data comparison. Examining the effects of gridlines, scale-bars, or other data visualization labelling and support approaches on the perception of tabletop data displays would allow this research to be even more ecologically valid. Lastly, the authors briefly discuss the omission of time measurement in their task design – it is possible that constraining the amount of time allowed for each magnitude comparison or recording time taken for each comparison would yield noteworthy results.

The only obvious shortcoming that authors fail to address is the small participant group size used for both studies. Although participant groups of 12 – 15 people are common in HCI research, repeated trials within these small groups do not provide the statistical power needed to make generalizable claims about human perception. Although repeating a task numerous times provides a large data set for statistical analysis, that analysis is still limited by individual participant difference within the data set – in effect, conclusions drawn from small participant groups may only apply to that specific group.

Bibliography

1. Baird, J.C. (1970), *Psychophysical Analysis of Visual Space*, New York: Pergamon Press.

2. Bertin, J. (1977). *La Graphique et le traitement graphique de l'information.* Flammarion, Paris, France.

3. Cleveland, W.S., McGill, R. (1984). Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. *Journal of the American Statistical Association*, 79(387). p. 531-553.

4. Cleveland, W.S., McGill, R. (1985). Graphical perception and graphical methods for analyzing and presenting scientific data. *Science,* 229(4716). p. 828-833.

5. Gescheider, G.A. (1988). Psychophysical Scaling. *Annual Review of Psychology*, 39. p. 159-200.

6. Tufte, E. (1983). *Visual Display of Quantitative Information.* Graphics Press, Cheshire, Connecticut.

7. Wigdor, D., Shen, C., Forlines, C., and Balakrishnan, R. (2007). Perception of elementary graphical elements in tabletop and multi-surface environments. In *Proc* CHI '07. ACM, New York, NY, 473-482.