The reliability of human clustering

Vijay Marupudi, V. N. Vimal Rao, Jimin Park, Rina Harsch, Jeffrey K. Bye, Sashank Varma
University of Minnesota

BACKGROUND

Traveling Salesman Problem

- Involves creating a tour which connects a set of cities while visiting each city exactly once.
- Computers find this problem hard.
- Humans are nearly optimal for low numbers of cities.
- Humans are pretty efficient, i.e., their solution time taken grows linearly with number of cities.
- It has been proposed that humans are guided by the external boundary of the cities (convex hull), but this strategy does not have a linear time complexity.

Hypothesis

- People might be ***clustering*** the problems into mini-TSPs, solving them, and then connecting the clusters together.
- This strategy could enable near-optimal, linear-time performance.

Research Questions

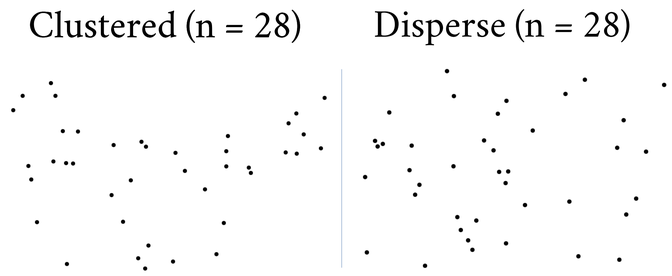
- **Is human clustering reliable to support this strategy?**
- Is human TSP performance reliable?
- Do people's clusters predict their TSP performance / path?

METHOD

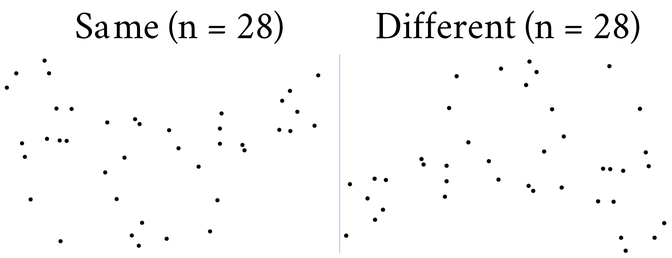
Participants: N= 13 (ongoing till 40), undergraduate students.
Procedure: Participants clustered 112 dot clouds at a computer using a mouse.
Measures: Number of clusters, cluster point membership, **calculated Fowlkes-Mallows index**, time to complete clustering.

MATERIALS

Dot clouds (n = 112)



The stimuli were shown again, with either the same or flipped orientation.



Fowlkes-Mallows Index
Measure of reliability

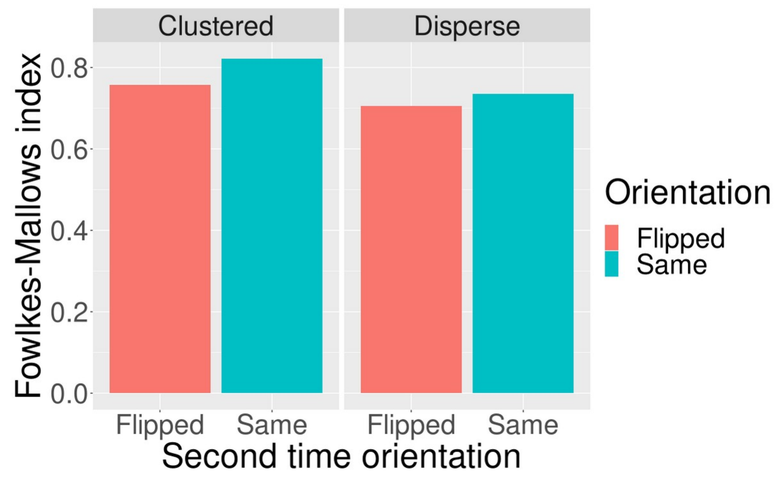
$$FM = \sqrt{\frac{TP}{TP+FP} \cdot \frac{TP}{TP+FN}}$$

Analysis: Linear Mixed Effects Model
Fixed Effects:

- Structure (Clustered/Disperse)
- Orientation (Same/Flipped)
- Number of points (10 – 40)
- Structure x Orientation

Random Effects: Participant, Stimulus

RESULTS



- High clustering reliability (M = 0.76)
- Main effect of orientation (p = 0.046), with higher reliability for same (M = 0.79) vs. different orientation (M = 0.72) on second viewing.
- No effect of structure (p = 0.21), with comparable reliabilities for clustered vs. dispersed stimuli.
- No effect of structure x orientation interaction (p = 0.36)

DISCUSSION

- High reliability suggests that clustering is a stable ability, and thus potentially a solid foundation for human TSP strategies.
- No duration differences for number of points suggests that clustering strategies might enable linear time complexity.