A Novel Generalization between Verbal Judgments and Perceptual Discrimination of 3D Space

Prachi Mahableshwarkar, John Philbeck, and Dwight Kravitz
Department of Psychological and Brain Sciences, The George Washington University

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC

Introduction

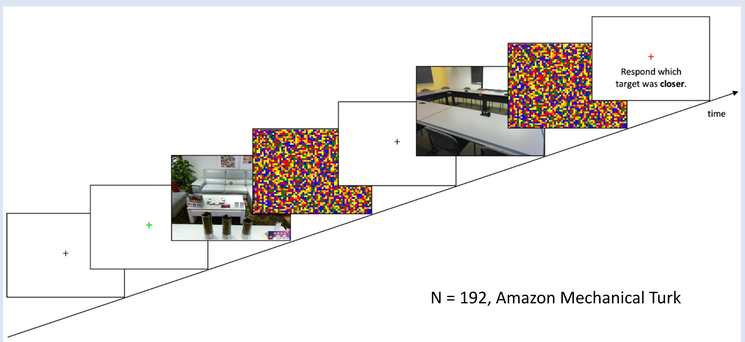
- Based on prior work in this field, we hypothesize that human visual distance perception in pictured scenes optimally integrates visual distance cues with temporal processing differences.
- Crowd-sourced data collection is leveraged to understand spatial perception within pictured scenes.
- We use two matched tasks focusing on depth perception (discrimination, verbal report) and quantify their shared variance across a large set of images

Stimuli

- All images in the dataset were taken with the Kinectv2 sensors, so the "ground truth" distance of each target was extracted from the depth array for each image in the SUN-RGBD dataset (Fig 1).
- All scenes are indoor and of diverse categories

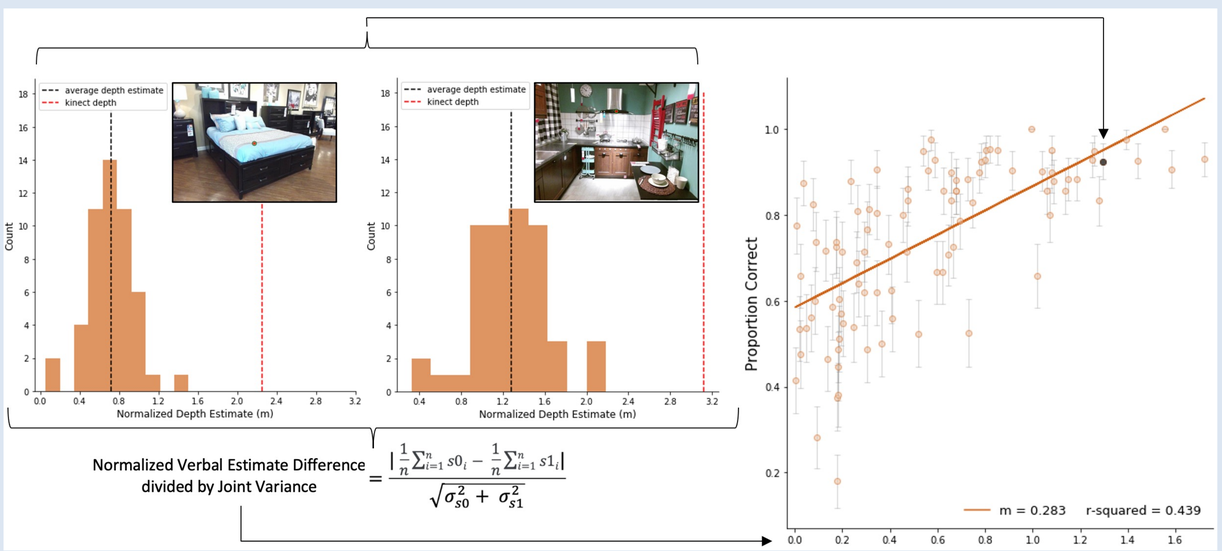
Experiment 2: Depth Discrimination

PARADIGM: Participants report which image's target was closer to them if they were to imagine themselves in the scene. There are 78 trials with images from the same stimulus set as Exp 1.

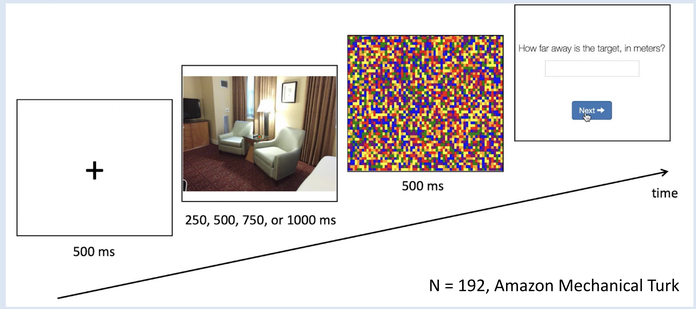


RESULTS

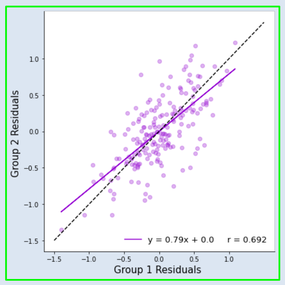
- Discrimination performance was strongly predicted by signal detection analyses applied to the verbal estimates from Experiment 1 for the stimuli in each discrimination (R2 = .57, p < 10^-14).

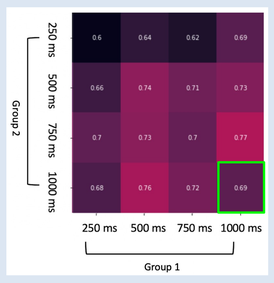


Experiment 1: Verbal Judgements of Distance



- When the effect of the actual distance was regressed out, significant and consistent residuals for each scene replicated across durations (R2 ~ .56, p ~ 10^-35), despite no participant overlap (Fig 5A, B).

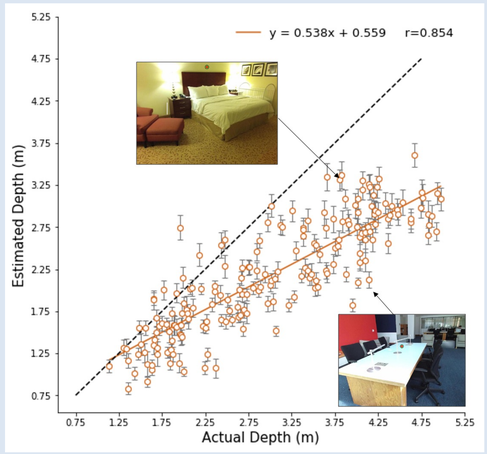


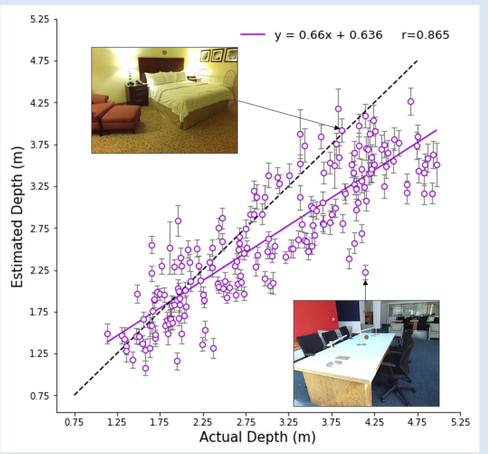


PARADIGM: Participants viewed 192 images with targets at variable distances (1m – 5m) and made egocentric distance estimations (Fig 3).

RESULTS

- Distance estimates to individual scenes showed increasing sensitivity to distance with longer viewing durations (F = 15.32, p < 10^-9) (Fig 4A, B).
- Participants underestimate distance, but there is strong systematicity in responses





TITLE?

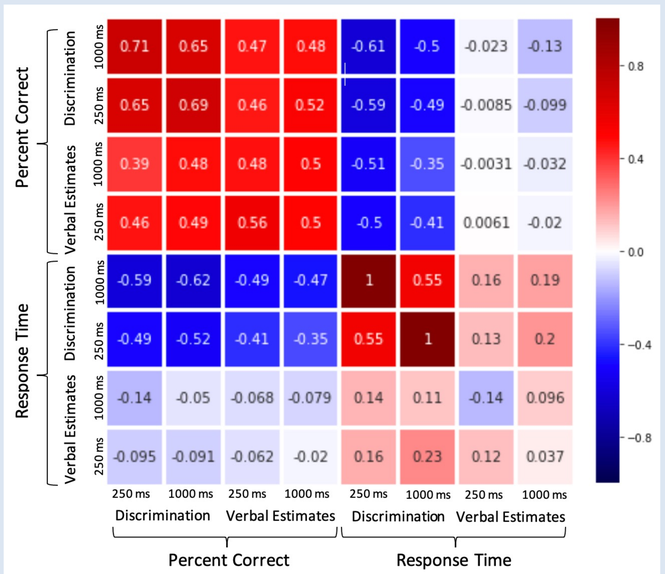
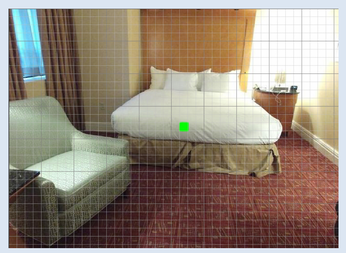


Fig X.

What spatial cues facilitate depth perception in pictured scenes (e.g., ground plane, familiar size, clutter)?



$$Object\ Surface\ Area = \sum_{i}^{n} SA_i = SA_1 + \ldots + SA_n$$

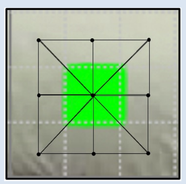


Fig X. Algorithmic quantification of size to determine the degree to which familiar size drives task-general spatial perception and the amount of it's influence over time.

Conclusion

The generalization across tasks and participants suggests a stable and strong relationship between verbal report and discrimination, allowing for the creation of an explicit model that relates stimulus-based visual cues and 3D picture perception.

REFERENCES