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

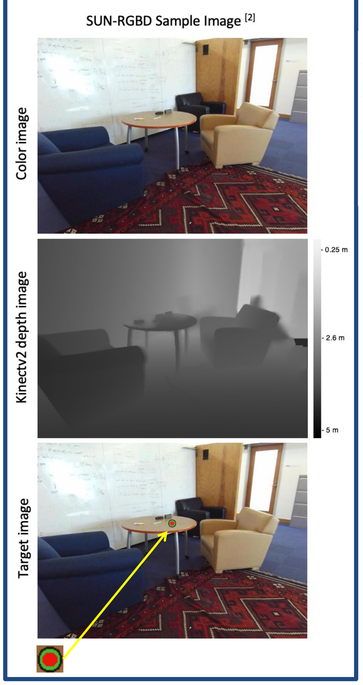
Prachi Mahableshwarkar, Lindsay Houck, John W. Philbeck, and Dwight J. Kravitz
Department of Psychological and Brain Sciences, The George Washington University

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC

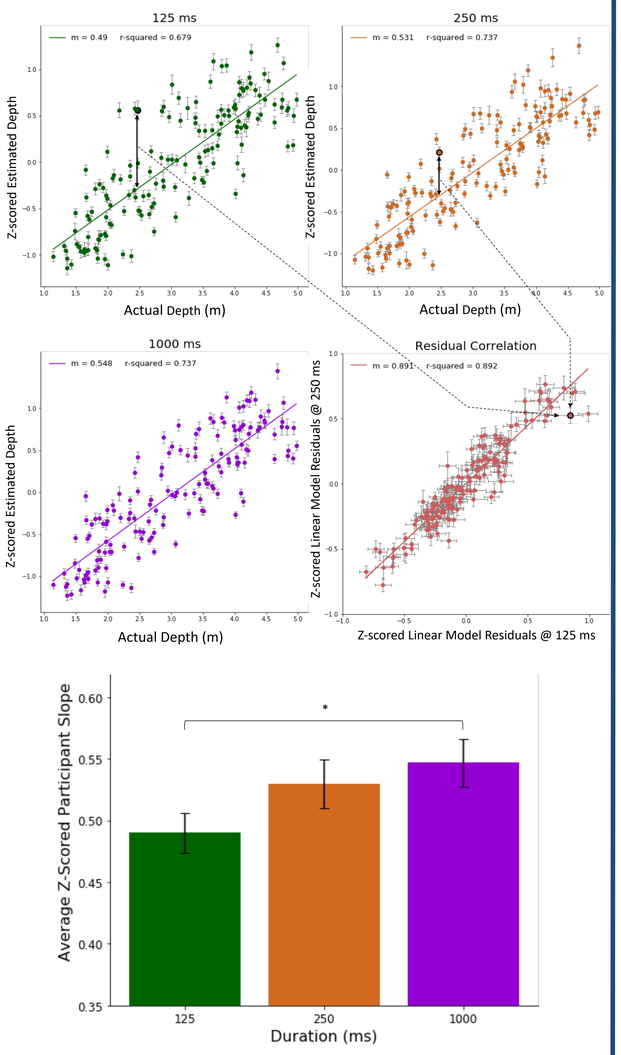
Introduction

Despite the lack of oculomotor and binocular disparity cues, humans are capable of visual distance perception in pictured scenes. Here we leverage crowd-sourced data collection to model the influence of spatial cues with temporal processing differences.

Stimuli

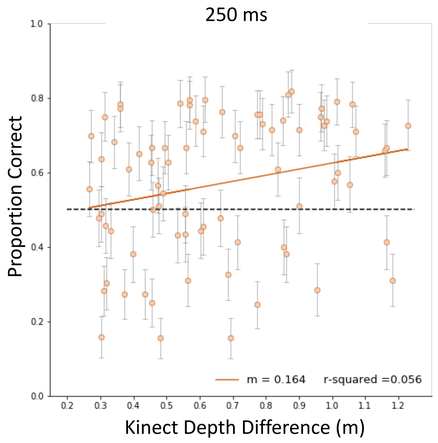


Verbal Estimate

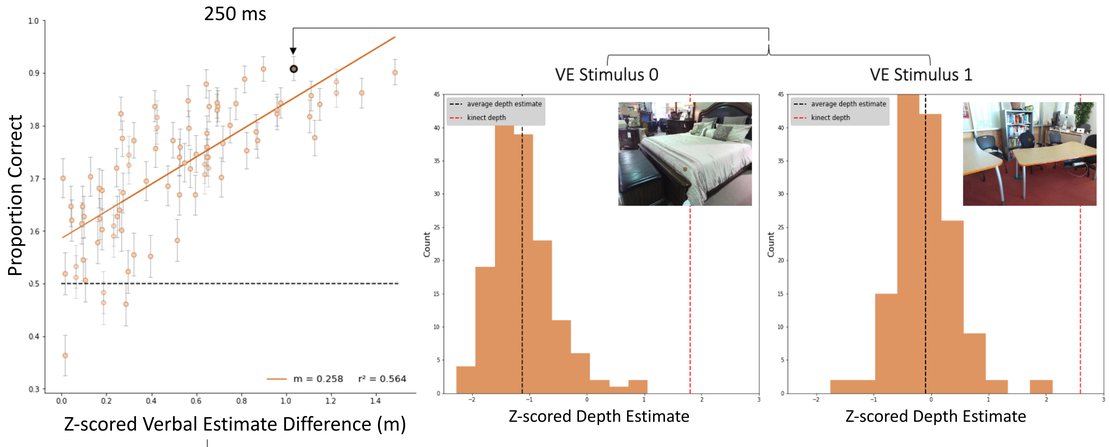


Discrimination

How predictable is discrimination performance from Kinect tagged depths?

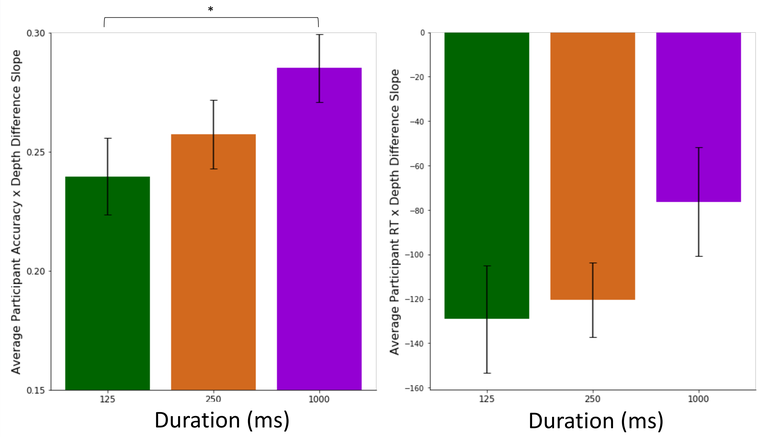


How predictable is discrimination performance from independent verbal judgments (Exp 1)?



Z-scored Verbal Estimate Difference = $|\frac{1}{n}\sum_{i=1}^{n} s0_i - \frac{1}{n}\sum_{i=1}^{n} s1_i|$

Discrimination Summary Plots

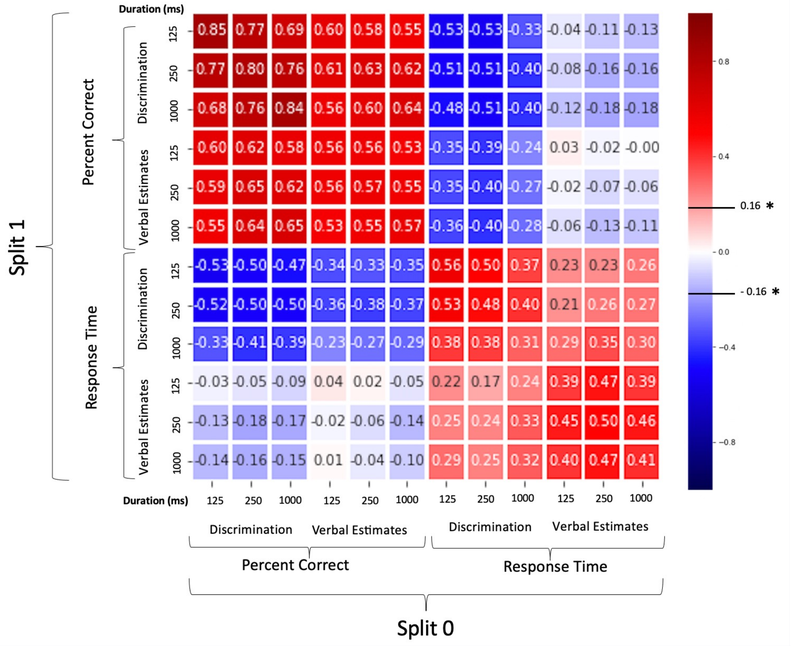


Stimulus Driven Task Generalization

We found high reliability in the relationships between tasks, participants, and viewing duration in the complete dataset and after repeatedly splitting the data into independent subsets and evaluating the average correlation matrix. Groups of subjects were selected repeatedly at random within each combination of task and duration. Independent correlations across all combinations of task (VE, Discrimination), outcome (percent correct, RT), and duration (125 ms, 250 ms, 1000 ms) were calculated and averaged across all random splits. We find a strong degree of generalization between task, duration, and outcome as a product of the stimulus-driven signal.

Key Findings:

1. Strong correlations along the main diagonal
2. Significant negative correlations between accuracy and RT in discrimination and VE
3. Percent correct from VE is significantly correlated with percent correct from discrimination at each duration
4. At each duration, response time from discrimination can significantly predict accuracy in VE



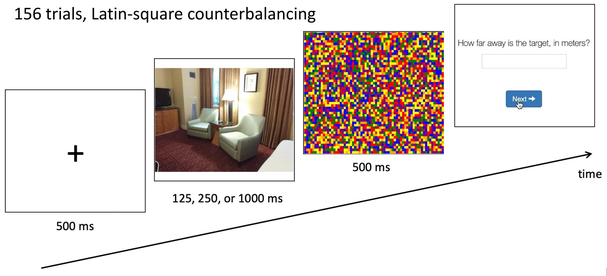
Experimental Paradigms

OSF PREREGISTRATION
https://osf.io/28vjd

amazon mechanical turk beta

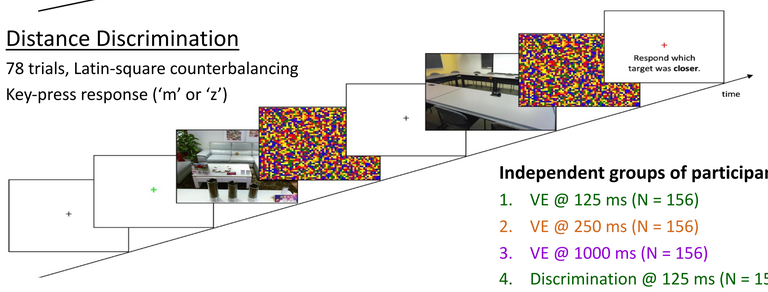
Verbal Estimate (VE) of Distance

156 trials, Latin-square counterbalancing



Distance Discrimination

78 trials, Latin-square counterbalancing
Key-press response ('m' or 'z')



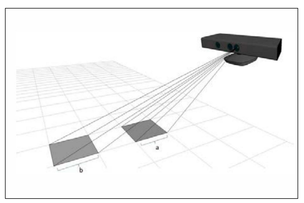
Independent groups of participants

1. VE @ 125 ms (N = 156)
2. VE @ 250 ms (N = 156)
3. VE @ 1000 ms (N = 156)
4. Discrimination @ 125 ms (N = 156)
5. Discrimination @ 250 ms (N = 156)
6. Discrimination @ 1000 ms (N = 156)

Conclusion

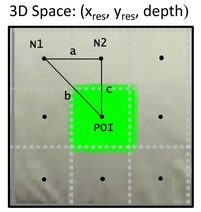
The generalization across tasks and participants suggests a stable and strong relationship between verbal report and discrimination, allowing for the creation of an explicit model that relates stimulus-based visual cues and 3D picture perception. Future work will quantify and model the influence of various visual cues, such as familiar size, on behavioral performance.

Familiar Size Algorithm



Calculating real world object dimensions from Kinect RGB-D image using dynamic resolution. [1]

3D Space: (x_{res}, y_{res}, depth)



$s = (a + b + c)/2$

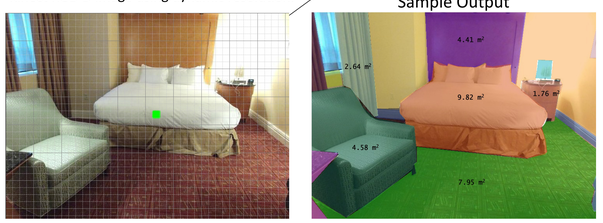
Pixel Surface Area

$SA = \sqrt{s(s-a)(s-b)(s-c)}$

Object Surface Area

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

Sample Output



References

[1] Anwer, A., Baig, A., & Nawaz, R. (2015). Calculating Real World Object Dimensions from Kinect RGB-D image using dynamic resolution. 2015 12th International Bhurban Conference on Applied Sciences and Technology (IBCAST).

[2] Song, S., Lichtenberg, S. P., & Xiao, J. (2015). Sun RGB-D: A RGB-D scene understanding benchmark suite. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Contact: pmahable@gwmail.gwu.edu