33.317 Human Object Detection in Natural Scenes: Evidence From a New Dot Probe Task

Colin S. Flowers & Mary A. Peterson

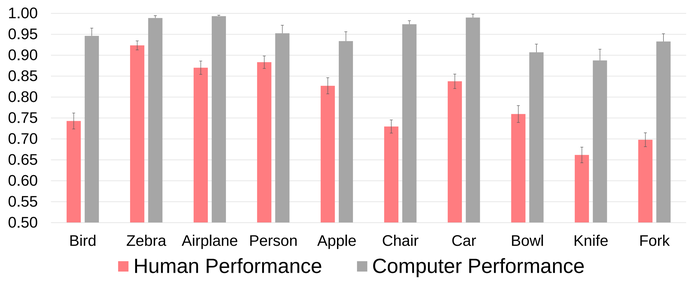
Contact: Colin S. Flowers
cflowers@email.arizona.edu

THE UNIVERSITY OF ARIZONA.

Background

Object detection entails “what” and “where” an object is

Human and Computer Accuracy on 2AFC[1]



2AFC method used in both human and computer vision
- Assesses “what” object was present
- Not “where” with respect to a scene border
- Allows context-, texture-, &/or feature-based guessing

Goal

Assess “where” an object is relative to a border in naturalistic scene photos using a dot probe on/off task

Methods

Brief masked exposures (741 photos from CoCo Set[2])

Flickering dot probes near object borders
- Half “on” objects; half “off” objects

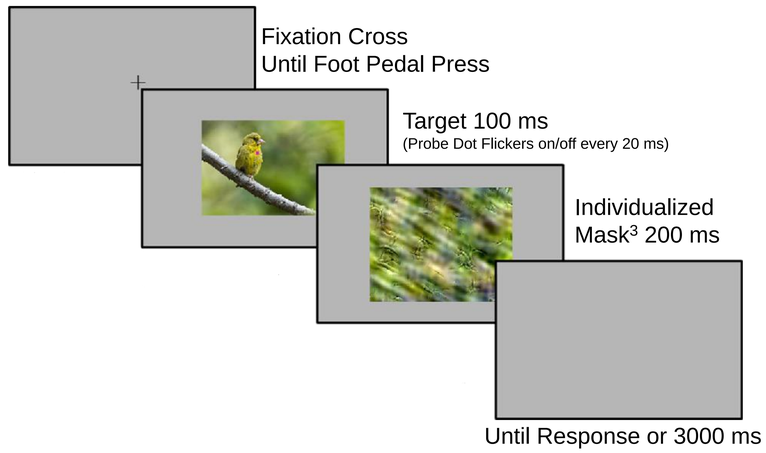
Task: Was dot probe “on” or “off” the object bounded by the nearest border
- Free of context-based guessing
- Allows signal detection analysis (d’)

Two experiments: 1) colored photos 2) gray scale photos
- Colored photos: Dot color chosen to contrast with local area
- Gray scale photos: Dot probe was always cyan

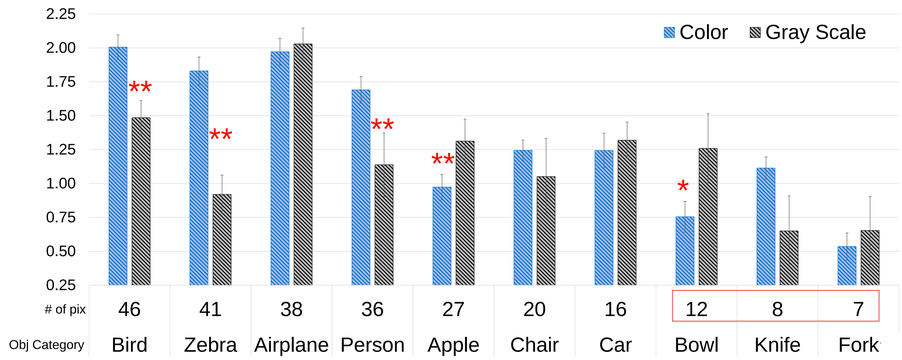
Participants viewed each photo once with either “on” or “off” dot

Two versions of each photo: one with “on” dot ; one with “off” dot
- Dot location balanced within & between participants

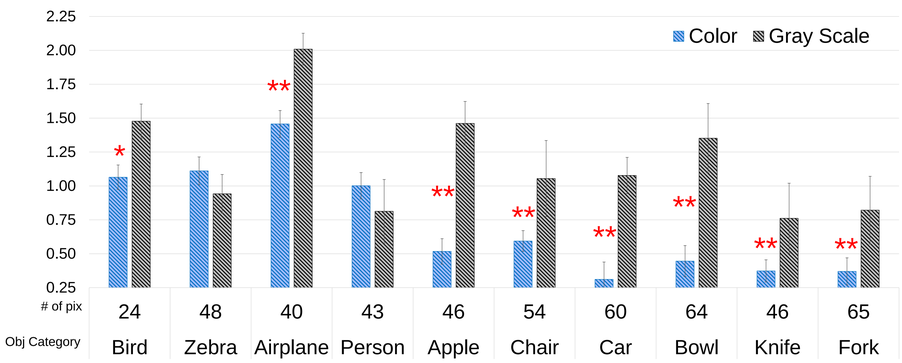
Trial Structure



d’: Central Region (Dot Probe ≤ 2º from Fixation) * $p < 0.05$ ** $p < 0.01$



d’: Peripheral Region (Dot Probe ≥ 2º from Fixation)



Dot probe task (“where”) provides different information than 2AFC (“what”)

Central dot probe: Between-category differences for gray scale vs. color

Peripheral dot probe: Poor performance on color photos
- May impede dot probe *and/or* scene processing

Low d’, High 2AFC

d’ = 0.41 2AFC Human: 96%

Perceptual organization processes (i.e., grouping, enclosure) affect processing of objects and backgrounds

Low 2AFC, High d’

d’ = 2.37 2AFC Human: 65%

Dot probe draws attention to the object, marks the relevant location within the scene

Summary

Dot probe task

Indexes “where” objects lie relative to a border

Free of context-based guessing
- Yet, affected by context: some colored photos better with central dots
  - Color can predict context; context can affect “where” response[4]

Limitations

Dot probe is too large for very small objects

Difficult to equate contrast of dot probes on color & gray scale photos
- Replicate with B/W striped dot probes on both types of photos

In occluded & crowded scenes, unclear which border is relevant

Future Directions

Integrate “what” and “where” tasks
- 1st Response: dot “on”/”off”; 2nd Response: object category (10AFC)

Assess the role of object and scene familiarity in dot probe task
- Inverted scenes: Reduce familiarity of both

References

[1] Clevenger, John, and Diane Beck. "How well do Deep Neural Networks model Human Vision?." *Journal of Vision* 16, no. 12 (2016): 176-176.

[2] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. (2014, September). Microsoft coco: Common objects in context. In *European conference on computer vision* (pp. 740-755). Springer, Cham.

[3] Portilla, J., & Simoncelli, E. P. (2000). A parametric texture model based on joint statistics of complex wavelet coefficients. *International journal of computer vision*, *40*, 49-70.

[4] Peterson, M. A., & Salvagio, E. (2008). Inhibitory competition in figure-ground perception: Context and convexity. *Journal of Vision*, *8*, 4-4.

Office of Naval Research Science & Technology

ONR N00014-14-1-067

Peterson Lab