



HAL
open science

Image Characteristics and Task Influence Interobserver Consistency

Sophie Heer, Marek A Pedziwiatr, Kristina Zeljic, Peter Bex, Antoine Coutrot, Isabelle Mareschal

► **To cite this version:**

Sophie Heer, Marek A Pedziwiatr, Kristina Zeljic, Peter Bex, Antoine Coutrot, et al.. Image Characteristics and Task Influence Interobserver Consistency. Vision Science Society, May 2024, St Pete Beach, Florida, United States. Vision Sciences Society Annual Meeting Abstract, 24 (867), 2024, 10.1167/jov.24.10.867 . hal-04717020

HAL Id: hal-04717020

<https://hal.science/hal-04717020v1>

Submitted on 1 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Image Characteristics and Task Influence Interobserver Consistency

Background

Interobserver consistency (IOC) refers to the degree of similarity in the gaze patterns of different observers when viewing the same image. Here, we explored how image characteristics and observers' viewing task influence IOC.

Our Study

- Participants ($N = 96$) viewed 192 images of scenes taken from the ADE20K dataset (Zhou et al., 2017).
- Participants either free-viewed the images or performed a memory task (between-subjects design).
- We recorded their eye movements.

Image Characteristics

We focused on two measures: visual clutter and semantic complexity.

Visual Clutter

- We used the Feature Congestion measure of visual clutter (Rosenholtz et al., 2007) which analyses how features (e.g., colour and luminance contrast) interact across an image.
- Feature Maps \rightarrow Pooling \rightarrow Conspicuity Maps \rightarrow Normalisation \rightarrow Feature Congestion Value.
- A higher Feature Congestion value indicates greater clutter.

Semantic Complexity

- We used GloVe (Pennington et al., 2014) to capture the complexity of the semantic content of each image.
- To do this, we selected images from the ADE20K dataset of labelled images (Zhou et al., 2017).
- We mapped object labels to vectors in the GloVe word embedding and calculated the pairwise semantic dissimilarity between objects within each image.
- Then, we took the average to obtain a 'semantic complexity' value per image, where higher scores indicate greater complexity.

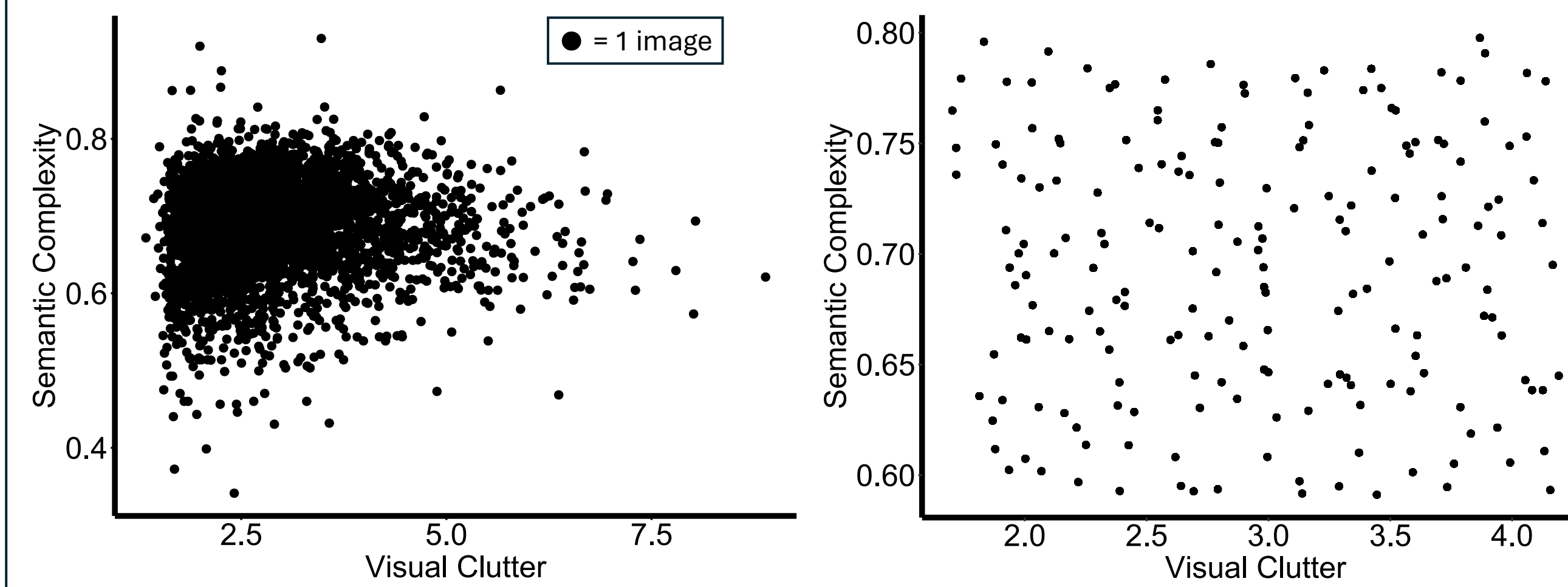


Semantic Complexity = 0.67

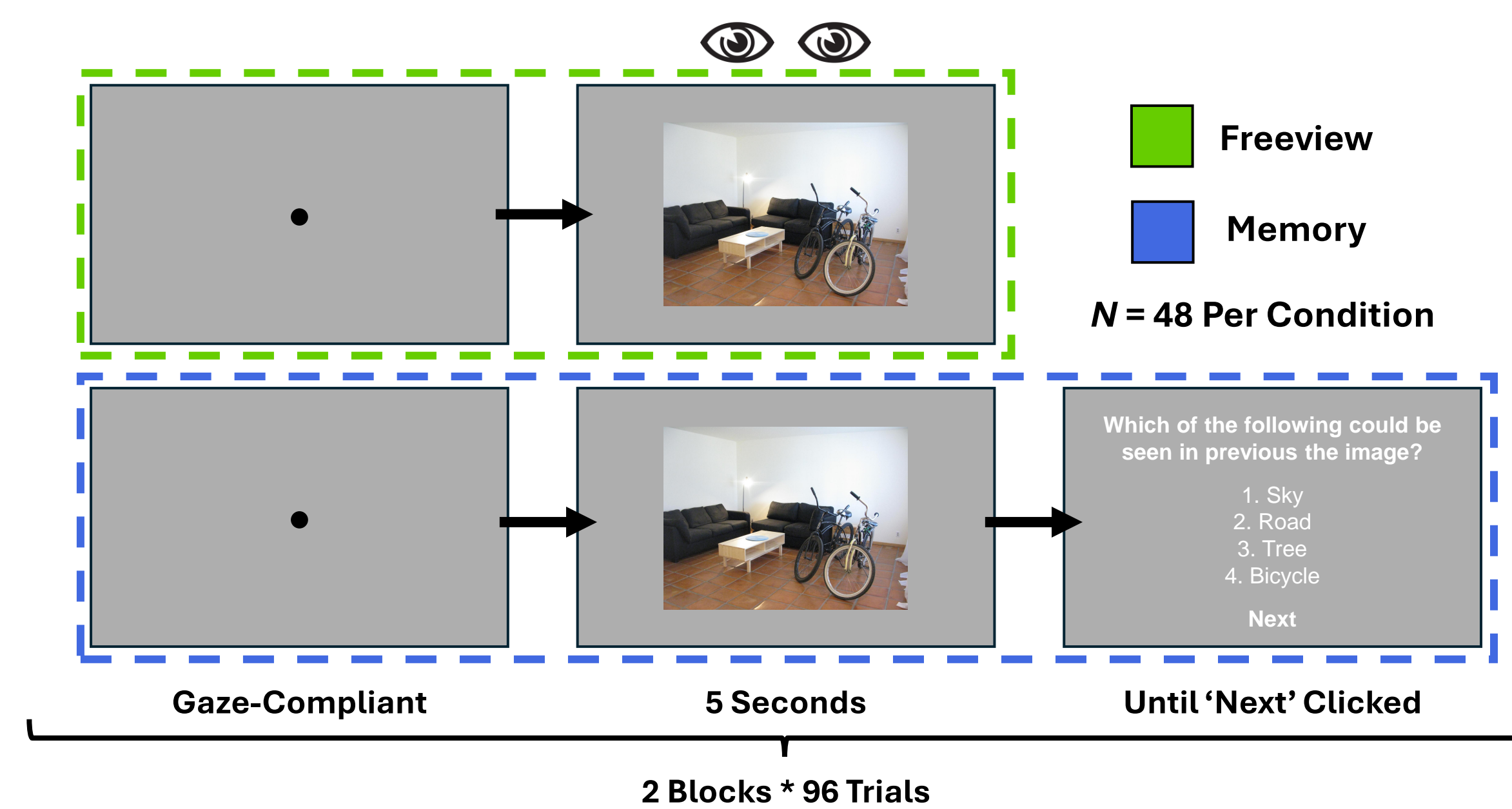
bicycle	0	0.86	0.71	0.77	0.8	0.86	0.72	0.7	0.89	0.81	0.79	0.71	0.73	0.9
ceiling	0.86	0	0.75	0.44	0.34	0.79	0.45	0.54	0.71	0.57	0.57	0.74	0.31	0.8
coffee	0.71	0.75	0	0.8	0.67	0.83	0.74	0.69	0.88	0.57	0.56	0.85	0.74	0.7
curtain	0.77	0.44	0.8	0	0.53	0.8	0.53	0.57	0.79	0.57	0.64	0.71	0.44	0.7
floor	0.8	0.34	0.67	0.53	0	0.82	0.65	0.62	0.77	0.48	0.51	0.74	0.37	0.6
highlight	0.86	0.79	0.83	0.8	0.82	0	0.85	0.88	0.97	0.92	0.76	0.92	0.78	0.6
lamp	0.72	0.45	0.74	0.53	0.65	0.85	0	0.33	0.59	0.6	0.57	0.61	0.5	0.5
light	0.7	0.54	0.69	0.57	0.62	0.68	0.33	0	0.72	0.76	0.69	0.63	0.56	0.5
socket	0.89	0.71	0.88	0.79	0.77	0.97	0.59	0.72	0	0.83	0.74	0.63	0.69	0.6
sofa	0.81	0.57	0.57	0.57	0.48	0.92	0.6	0.76	0.83	0	0.46	0.79	0.61	0.5
table	0.79	0.57	0.56	0.64	0.51	0.76	0.57	0.69	0.74	0.46	0	0.83	0.57	0.5
tube	0.71	0.74	0.85	0.71	0.74	0.92	0.61	0.63	0.63	0.79	0.83	0	0.56	0.5
wall	0.73	0.31	0.74	0.44	0.37	0.78	0.5	0.56	0.69	0.61	0.57	0.56	0	0.5
bicycle														0.9
ceiling														0.8
coffee														0.7
curtain														0.6
floor														0.5
highlight														0.4
lamp														0.3
light														0.2
socket														0.1
sofa														0
table														0
tube														0
wall														0

Image Selection

- We wanted a diverse set of images covering the range of visual clutter and semantic complexity scores, with these metrics uncorrelated.
- To achieve this, we computed metrics for all images and selected a final stimulus set meeting this criteria.



Tasks: Freeviewing & Memory



Interobserver Consistency

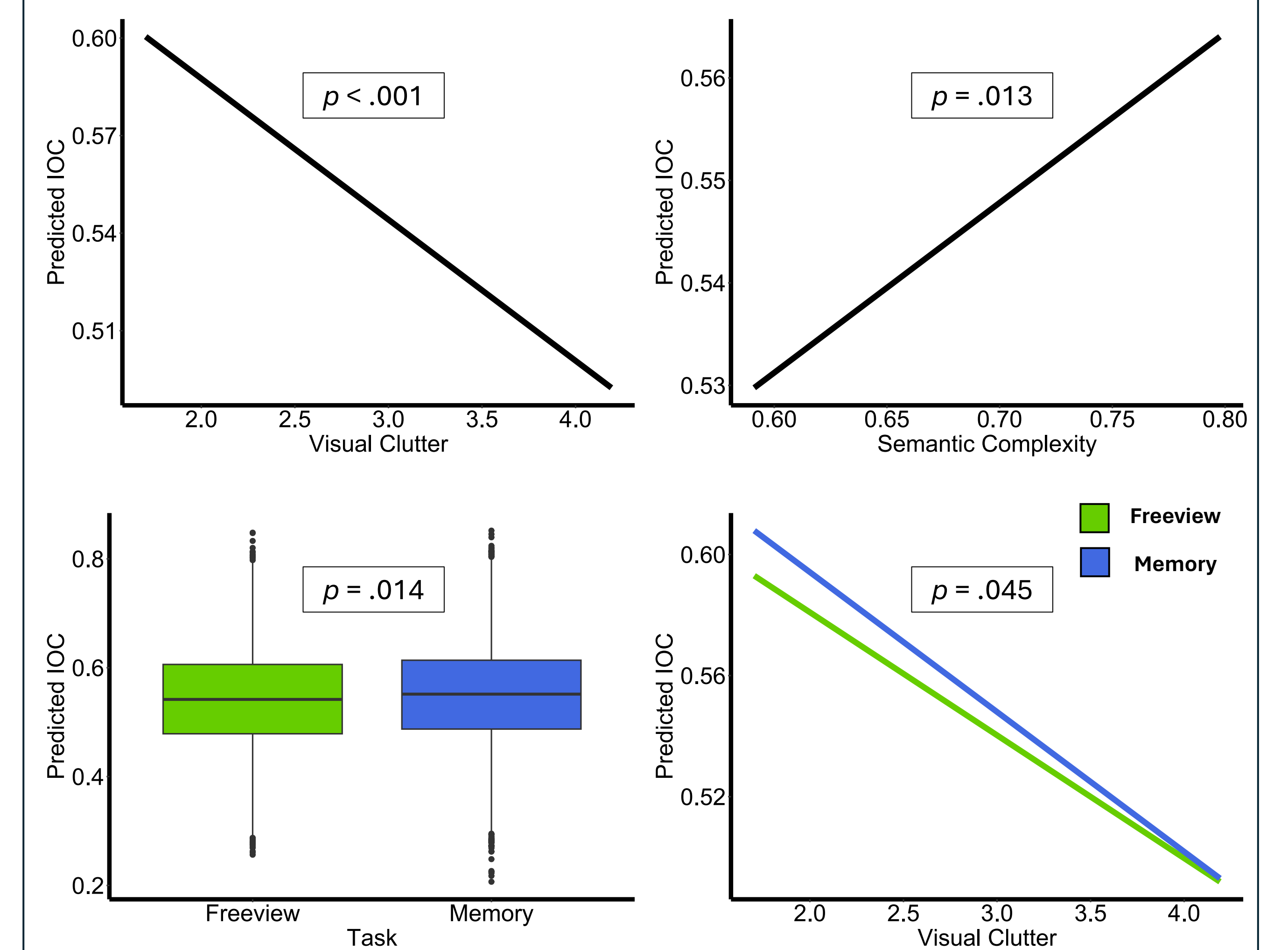
How well the fixation heatmap of $N-1$ observers predicts the fixation heatmap of the remaining observer using Pearson's correlation coefficient.



Linear Mixed Effects Model

$IOC \sim \text{Visual Clutter} * \text{Task} + \text{Semantic Complexity} * \text{Task} + \text{Number of Labels} + (1|\text{Image}) + (1|\text{Participant})$

Results



Conclusions

- Image characteristics influence IOC, with IOC decreasing as visual clutter increases and IOC increasing with semantic complexity.
- Task also impacts IOC, being higher during the memory task compared to freeviewing. However, a more defined task may enhance this effect.
- Clutter exerts a greater influence on IOC during the memory task compared to freeviewing.

References

- Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global vectors for word representation. In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP) (pp. 1532-1543).
- Rosenholtz, R., Li, Y., & Nakano, L. (2007). Measuring visual clutter. Journal of Vision, 7(2), 17-17.
- Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A., & Torralba, A. (2017). Scene parsing through ADE20K dataset. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 633-641).