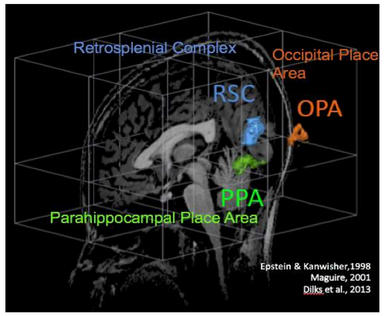
# Uncovering a scene-defining feature using converging stimuli-based, behavioral and neural approaches

Annie Cheng*, Jeanette Wong, & Daniel D. Dilks
*Psychology, Emory University*
* corresponding author: rcheng6@emory.edu

## Introduction

Neuroimaging studies reveal distinct cortical regions that respond selectively to scenes.

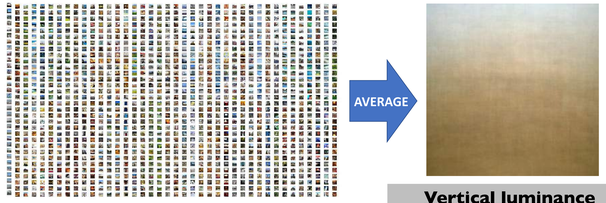



## Research Question

**Is there a visual feature that enables the human brain to differentiate scene from non-scene stimuli?**

## Part I: What is a visual feature common across most scenes?

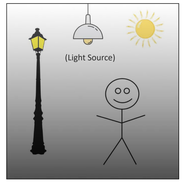
**A stimuli-driven approach:**
Average across highly variable scene images to wash out the unique features of each image and uncover the consistent, recurring visual features.




- Adopted from Konkle et al. (2010)
- 1280 scene images from 80 scene categories, 16 different images per category

**Vertical luminance gradient (VLG):**
Brighter in the upper versus lower half of the image p<.0001.

**Intuitively, VLG makes sense because in our everyday lives, light always comes from above.**




*But is VLG specific to scenes, or common across all visual inputs?*

1. Is there a VLG in the stimuli commonly used for studying object recognition?

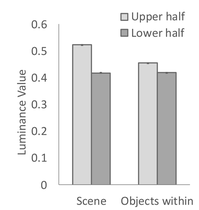
n.s. in the luminance of the upper vs. lower halves of objects


- Adopted from Brady et al. (2008)
- 2235 total object images from 198 objects categories, 15 different images per category

2. Is there a VLG in the objects within naturalistic scene images?

A greater VLG in scenes than in the objects within a scene.

- Adopted from ADE20K database (Zhou et al., 2017)
- 2220 scene images, all objects within each scene are individually segmented
- At least 5 objects per scene image
- Compare the VLG of the scenes vs. the objects within

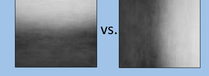




## Hypothesis:
## VLG is a scene-defining feature.

Predictions:

1. If VLG is a scene-defining feature, then scene processing should be impaired when VLG is disrupted (e.g., by rotating the image 90°).

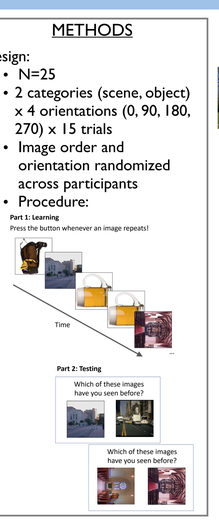

2. If VLG is a scene-defining feature, then even a simple image with only VLG should be sufficient to be categorized as a scene (i.e., scene-selective regions will respond significantly more to an upright vs. rotated "impoverished" scene image).

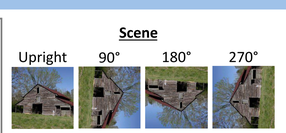



## Part II: Behavioral Experiment

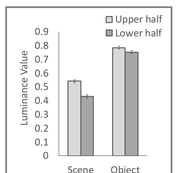
METHODS

Design:
- N=25
- 2 categories (scene, object) x 4 orientations (0, 90, 180, 270) x 15 trials
- Image order and orientation randomized across participants
- Procedure:

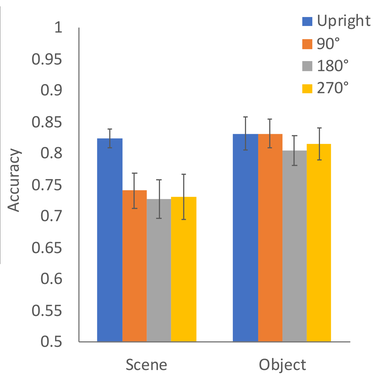







**VLG is stronger in the upright scene images than object images.

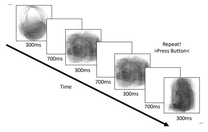



Rotating an image, thus disrupting VLG in scenes, impaired scene processing, but not object processing.

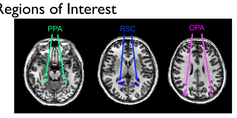
## Part III: fMRI Experiment

METHODS

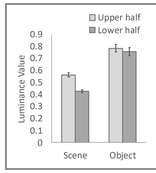
Design:
- N=3
- Block design (12s, 12 images per block)
- Image order and block order randomized across participants
- One-back repetition detection task




Regions of Interest

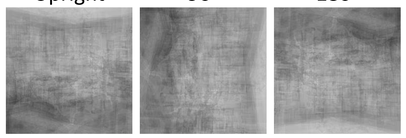



Independent functional localizers:
- PPA, RSC & OPA: Scenes – Objects
- LOC: Objects – Scrambled Objects
- FC: Scrambled Objects-Objects

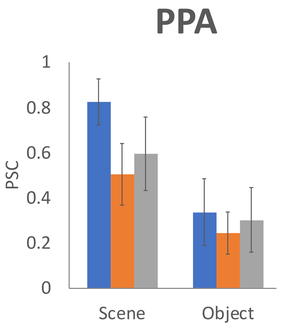


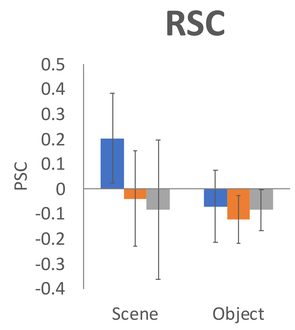

**VLG is only present in the impoverished upright scene images, but not the rotated scene images and the object images.

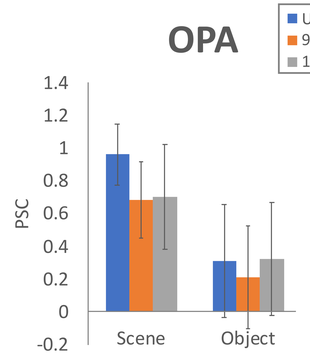
**Scene**

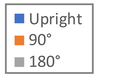


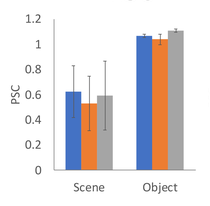

**Object**

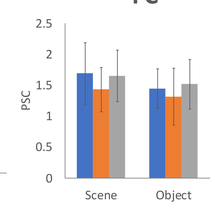




All scene-selective regions responded more to an upright than rotated "impoverished" scene image (with disrupted VLG), whereas LOC and FC do not.