Research report

The occipital place area represents first-person perspective motion information through scenes

Frederik S. Kamps, Vishal Lall and Daniel D. Dilks[*]

Department of Psychology, Emory University, Atlanta, GA, United States

ARTICLE INFO

ABSTRACT

Neuroimaging studies have identified multiple scene-selective regions in human cortex, but the precise role each region plays in scene processing is not yet clear. It was recently hypothesized that two regions, the occipital place area (OPA) and the retrosplenial complex (RSC), play a direct role in navigation, while a third region, the parahippocampal place area (PPA), does not. Some evidence suggests a further division of labor even among regions involved in navigation: While RSC is thought to support navigation through the broader environment, OPA may be involved in navigation through the immediately visible environment, although this role for OPA has never been tested. Here we predict that OPA represents first-person perspective motion information through scenes, a critical cue for such "visually-guided navigation", consistent with the hypothesized role for OPA. Response magnitudes were measured in OPA (as well as RSC and PPA) to i) video clips of first-person perspective motion through scenes ("Dynamic Scenes"), and ii) static images taken from these same movies, rearranged such that first-person perspective motion could not be inferred ("Static Scenes"). As predicted, OPA responded significantly more to the Dynamic than Static Scenes, relative to both RSC and PPA. The selective response in OPA to Dynamic Scenes was not due to domain-general motion sensitivity or to low-level information inherited from early visual regions. Taken together, these findings suggest the novel hypothesis that OPA may support visually-guided navigation, insofar as first-person perspective motion information is useful for such navigation, while RSC and PPA support other aspects of navigation and scene recognition.

1. Introduction

Recognizing the visual environment, or "scene", and using that information to navigate is critical in our everyday lives. Given the ecological importance of scene recognition and navigation, it is perhaps not surprising then that we have dedicated neural machinery for scene processing: the occipital place area (OPA) (Dilks, Julian, Paunov, & Kanwisher, 2013), the retrosplenial complex (RSC) (Maguire, 2001), and the parahippocampal place area (PPA) (Epstein & Kanwisher, 1998). Beyond establishing the general involvement of these regions in scene processing, however, a fundamental and yet unanswered question remains: What is the precise function of each region in scene processing, and how do these regions

* *Corresponding author*. Department of Psychology, Emory University, 36 Eagle Row, Atlanta, GA 30322, United States.
 E-mail address: dilks@emory.edu (D.D. Dilks).

support our crucial ability to recognize and navigate our environment?

Growing evidence indicates that OPA, RSC, and PPA play distinct roles in scene processing. For example, OPA and RSC are sensitive to two essential kinds of information for navigation: sense (i.e., left vs right) and egocentric distance (i.e., near vs far from me) information (Dilks, Julian, Kubilius, Spelke, & Kanwisher, 2011; Persichetti & Dilks, 2016). By contrast, PPA is not sensitive to either sense or egocentric distance information. The discovery of such differential sensitivity to navigationally-relevant information across scene-selective cortex has lead to the hypothesis that OPA and RSC directly support navigation, while PPA does not (Dilks et al., 2011; Persichetti & Dilks, 2016). Further studies suggest that there may be a division of labor even among those regions involved in navigation, although this hypothesis has never been tested directly. In particular, RSC is thought to represent information about both the immediately visible scene and the broader spatial environment related to that scene (Epstein, 2008; Maguire, 2001), in order to support navigational processes such as landmark-based navigation (Auger, Mullally, & Maguire, 2012; Epstein & Vass, 2015), location and heading retrieval (Epstein, Parker, & Feiler, 2007; Marchette, Vass, Ryan, & Epstein, 2014; Vass & Epstein, 2013), and the formation of environmental survey knowledge (Auger, Zeidman, & Maguire, 2015; Wolbers & Buchel, 2005). By contrast, although little is known about OPA, it was recently proposed that OPA supports visually-guided navigation and obstacle avoidance in the immediately visible scene itself (Kamps, Julian, Kubilius, Kanwisher, & Dilks, 2016).

One critical source of information for such visually-guided navigation is the first-person perspective motion information experienced during locomotion (Gibson, 1950). Thus, here we investigated how OPA represents first-person perspective motion information through scenes. Responses in the OPA (as well as RSC and PPA) were measured using fMRI while participants viewed i) 3-sec video clips of first-person perspective motion through a scene ("Dynamic Scenes"), mimicking the actual visual experience of locomotion, and ii) 3, 1-sec still images taken from these same video clips, rearranged such that first-person perspective motion could not be inferred ("Static Scenes"). We predicted that OPA would respond more to the Dynamic Scenes than the Static Scenes, relative to both RSC and PPA, consistent with the hypothesis that OPA supports visually-guided navigation, since first-person perspective motion information is undoubtedly useful for such navigation, while RSC and PPA are involved in other aspects of navigation and scene recognition.

2. Method

2.1. Participants

Sixteen healthy university students (ages 20—35; mean age = 25.9; SD = 4.3; 7 females) were recruited for this experiment. All participants gave informed consent. All had normal or corrected to normal vision; were right handed (one reported being ambidextrous), as measured by the Edinburgh Handedness Inventory (mean = .74; SD = .31, where +1 is considered a "pure right hander" and −1 is a "pure left hander") (Oldfield, 1971); and had no history of neurological or psychiatric conditions. All procedures were approved by the Emory University Institutional Review Board.

2.2. Design

For our primary analysis, we used a region of interest (ROI) approach in which we used one set of runs (Localizer runs, described below) to define the three scene-selective regions (as described previously; Epstein & Kanwisher, 1998), and then used a second, independent set of runs (Experimental runs, described below) to investigate the responses of these regions to Dynamic Scenes and Static Scenes, as well as two control conditions: Dynamic Faces and Static Faces. As a secondary analysis, we performed a group-level analysis exploring responses to the Experimental runs across the entire slice prescription (for a detailed description of this analysis see Data analysis section).

For the Localizer runs, we used a standard method used in many previous studies to identify ROIs (Epstein & Kanwisher, 1998; Kamps et al., 2016; Kanwisher & Dilks, in press; Park, Brady, Greene, & Oliva, 2011; Walther, Caddigan, Fei-Fei, & Beck, 2009). Specifically, a blocked design was used in which participants viewed static images of scenes, faces, objects, and scrambled objects. We defined scene-selective ROIs using static images, rather than dynamic movies for two reasons. First, using the standard method of defining scene-selective ROIs with static images allowed us to ensure that we were investigating the same ROIs investigated in many previous studies of cortical scene processing, facilitating the comparison of our results with previous work. Second, the use of dynamic movies to define scene-selective ROIs could potentially bias responses in regions that are selective to dynamic information in scenes, inflating the size of the "dynamic" effect. The same argument, of course, could be used for the static images (i.e., potentially biasing responses in regions that are selective to static information in scenes, again inflating the size of the "dynamic" effect). However, note that in either case, the choice of dynamic or static stimuli to define scene-selective ROIs would result in a main effect of motion (i.e., a greater response to Dynamic Scenes than Static Scenes in all three scene-selective regions, or vice versa), not an interaction of motion by ROI (i.e., a greater response in OPA to Dynamic Scenes than Static Scenes, relative to PPA and RSC), as predicted. Each participant completed 3 runs, with the exception of two participants who only completed 2 runs due to time constraints. Each run was 336 sec long and consisted of 4 blocks per stimulus category. For each run, the order of the first eight blocks was pseudorandomized (e.g., faces, faces, objects, scenes, objects, scrambled objects, scenes, scrambled objects), and the order of the remaining eight blocks was the palindrome of the first eight (e.g., scrambled objects, scenes, scrambled objects, objects, scenes, objects, faces, faces). Each block contained 20 images from the same category for a total of 16 sec blocks. Each image was presented for 300 msec, followed by a 500 msec interstimulus interval (ISI), and subtended 8 × 8 degrees of visual angle. We also included five