Cognition 180 (2018) 118–122

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Brief article

# Typical visual-field locations facilitate access to awareness for everyday objects

Daniel Kaiser[a,*], Radoslaw M. Cichy[a,b,c]

[a] Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany
[b] Berlin School of Mind and Brain, Humboldt-Universität Berlin, Berlin, Germany
[c] Bernstein Center for Computational Neuroscience Berlin, Berlin, Germany

ARTICLE INFO

*Keywords:*
Object perception
Visual awareness
Continuous flash suppression
Real-world regularities
Scene statistics

ABSTRACT

In real-world vision, humans are constantly confronted with complex environments that contain a multitude of objects. These environments are spatially structured, so that objects have different likelihoods of appearing in specific parts of the visual space. Our massive experience with such positional regularities prompts the hypothesis that the processing of individual objects varies in efficiency across the visual field: when objects are encountered in their typical locations (e.g., we are used to seeing lamps in the upper visual field and carpets in the lower visual field), they should be more efficiently perceived than when they are encountered in atypical locations (e.g., a lamp in the lower visual field and a carpet in the upper visual field). Here, we provide evidence for this hypothesis by showing that typical positioning facilitates an object's access to awareness. In two continuous flash suppression experiments, objects more efficiently overcame inter-ocular suppression when they were presented in visual-field locations that matched their typical locations in the environment, as compared to non-typical locations. This finding suggests that through extensive experience the visual system has adapted to the statistics of the environment. This adaptation may be particularly useful for rapid object individuation in natural scenes.

## 1. Introduction

Human visual perception is tailored to the world around us: it is most efficient when the input matches commonly experienced patterns. This is evident from low-level vision, where previously experienced regularities determine perceptual interpretations of the input (Purves, Wojtach, & Lotto, 2011). Such influences of typical patterns are also observed for more complex stimuli, such as faces. Face perception is specifically tuned to the typical configuration of facial features (Maurer, Le Grand, & Mondloch, 2001), and a disruption of this configuration (e.g., through face inversion) drastically decreases perceptual performance (Valentine, 1988). Recent studies have suggested that not only the concerted presence of multiple features facilitates face perception, but that also individual facial features profit from typical positioning in the visual field (Chan, Kravitz, Truong, Arizpe, & Baker, 2010; de Haas et al., 2016; Moors, Wagemans, & de Wit, 2016): for example, it is easier to perceive an eye when it falls into the upper visual field (where it more often appears when looking at a face) than when it falls into the lower visual field (where it is not encountered so often).

Like faces, natural scenes are spatially structured. Scenes consist of arrangements of separable objects, which follow repeatedly experienced configurations (Bar, 2004): for instance, lamps appear above dining tables, and carpets tend to lie on the floor. Previous research has suggested that such typical configurations can facilitate multi-object processing (Draschkow & Võ, 2017; Gronau & Shachar, 2014; Kaiser, Stein, & Peelen, 2014, 2015). It has been proposed that just like in faces, spatial regularities in scenes may also impact the perception of individual objects (Kaiser & Haselhuhn, 2017). As we navigate around, the likelihood of encountering different objects varies across the visual field: for instance, lamps – unless directly fixated – are most often seen in the upper visual field and carpets most often appear in the lower visual field. Because of this repeated expose, typically positioned objects should be processed more efficiently than atypically positioned objects.

To test this hypothesis, we used a variant of continuous flash suppression (CFS; Tsuchiya & Koch, 2005). In breaking-CFS paradigms, a stimulus presented to one eye is temporarily rendered invisible by flashing a dynamic, high contrast mask to the other eye; suppression times, i.e. the time a stimulus needs to break inter-ocular suppression and reach visual awareness, are taken as a measure of processing

* Corresponding author at: Department of Education and Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany.
*E-mail address:* danielkaiser.net@gmail.com (D. Kaiser).

https://doi.org/10.1016/j.cognition.2018.07.009
Received 8 April 2018; Received in revised form 10 July 2018; Accepted 12 July 2018

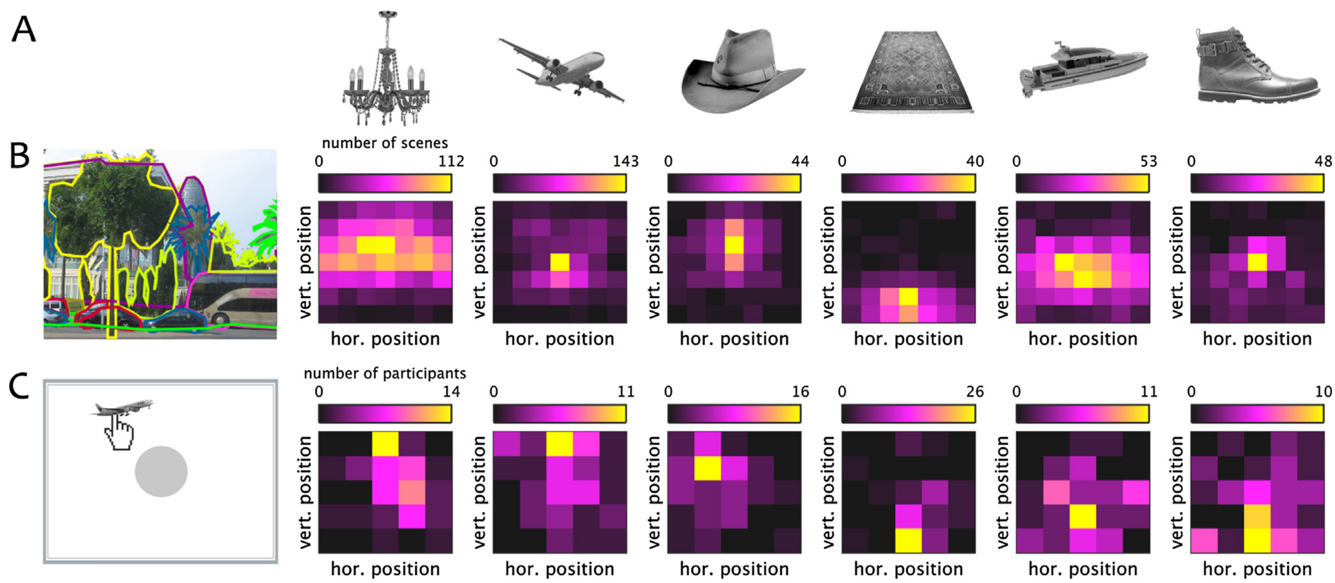


**Fig. 1.** The stimulus set consisted of six objects (10 exemplars each), of which three (lamp, airplane, hat) were associated with upper visual-field locations and three (carpet, boat, shoe) were associated with lower visual-field locations (A). The visual-field associations were validated by computing two measures (see Section 2 for details): First, we used a large set of labelled scenes (Russell et al., 2008) to extract typical within-scene positions for each object (B). Second, we asked a set of participants to freely place the object on the screen so that its position best matches its typical real-world position (C). Heatmaps reflect the distribution of locations across a scene photograph (B) or the computer screen (C).

efficiency (Stein, Hebart, & Sterzer, 2011). Previous studies using this method have shown that suppression times depend on spatial regularity patterns. For example, the typical configuration of faces and bodies facilitates their access to awareness (Jiang, Costello, & He, 2007; Stein, Sterzer, & Peelen, 2012). Similarly, breakthrough is facilitated for typically arranged multi-object configurations (Stein, Kaiser, & Peelen, 2015), demonstrating that the spatial regularities among different objects can facilitate processing under CFS.

To test whether such spatial regularities also impact the processing of individual objects we investigated whether typical retinotopic positioning facilitates an object's access to awareness. We used a stimulus set consisting of six everyday objects that were either associated with upper or lower visual-field locations (Fig. 1). In two CFS experiments, participants were shown individual exemplars of these objects in their typical or atypical locations onto one eye; a dynamic mask was flashed onto the other eye and temporarily rendered the object invisible (Fig. 2). Participants had to localize the object as fast as possible, irrespective of its identity. In Experiment 1, suppression times (i.e., times until successful localization) were significantly shorter for typically than for atypically positioned objects. In Experiment 2, we replicated this finding, while additionally controlling for potential response conflicts. These results demonstrate that objects appearing in typical visual-field locations gain preferential access to visual awareness, highlighting the influence of natural scene structure on individual object perception.

## 2. Material and methods

### 2.1. Participants

34 healthy adults participated in Experiment 1 (mean age 26.4 years, $SD = 4.7$, 26 female) and another 34 participated in Experiment 2 (mean age 22.9 years, $SD = 4.4$, 26 female). Participants were recruited from the online participant database of the Berlin School of Mind and Brain (Greiner, 2015). All participants had normal or corrected-to-normal vision, provided informed consent and received monetary reimbursement or course credits for participation. All procedures were approved by the local ethical committee and were in accordance with the Declaration of Helsinki.

Sample size was determined by an a-priori power calculation: assuming a hypothetical, medium-sized effect of $d = 0.5$, 34 participants are needed for a power of 80%[1].

### 2.2. Stimuli

The stimulus set consisted of six objects (Fig. 1A). Three of the objects were associated with upper visual-field locations (lamp, airplane, and hat) and three were associated with lower visual-field locations (carpet, boat, and shoe). For each object, we collected ten exemplars. The objects were matched for their categorical content (two furniture items, two transportation items, and two clothing items) to match high-level properties (e.g., the objects' size, manipulability and semantic associations) across upper and lower visual-field objects. To control for low-level confounds, stimulus images were gray-scaled and matched for overall luminance (Willenbockel et al., 2010). Additionally, we checked whether there was a consistent low-level difference across objects associated with upper and lower visual-field locations. For this, we computed pair-wise pixel correlations for all conditions, and compared results for objects associated with the same visual-field locations versus objects associated with different visual-field locations. This test was not significant, $t(1498) = 0.50$, $p = 0.62$, suggesting that there was no consistent low-level difference across upper and lower visual-field objects.

To validate the objects' associations with specific locations, we used two complementary approaches. First, we automatically queried a large database (> 10,000 images) of labelled scene photographs (LabelMe; Russell, Torralba, Murphy, & Freeman, 2008). We assumed that the distribution of objects across a larger number of photographs approximates their distribution under natural viewing conditions. For each scene that contained one of the six objects, we extracted the within-scene location (the mean coordinate of the labelled area) of the object (Fig. 1B). Second, we explicitly asked a set of participants to place each object on a computer screen such that its on-screen position mirrored its most probable real-world positioning (Fig. 1C). For both validation approaches, vertical locations were significantly higher for upper than

[1] A power analysis based on the effect obtained in Experiment 1 ($d = 0.59$) revealed a power of 92% for a sample size of 34 in Experiment 2.