Vision Research 133 (2017) 73–80

Contents lists available at ScienceDirect

Vision Research

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

Top-down preparation modulates visual categorization but not subjective awareness of objects presented in natural backgrounds

Mika Koivisto *, Ella Kahila

Department of Psychology, University of Turku, Assistentinkatu 7, 20014 Turku, Finland

ARTICLE INFO

Article history:
Received 18 August 2016
Received in revised form 12 January 2017
Accepted 14 January 2017

Keywords:
Attention
Awareness
Categorization
Masking

ABSTRACT

Top-down processes are widely assumed to be essential in visual awareness, subjective experience of seeing. However, previous studies have not tried to separate directly the roles of different types of top-down influences in visual awareness. We studied the effects of top-down preparation and object substitution masking (OSM) on visual awareness during categorization of objects presented in natural scene backgrounds. The results showed that preparation facilitated categorization but did not influence visual awareness. OSM reduced visual awareness and impaired categorization. The dissociations between the effects of preparation and OSM on visual awareness and on categorization imply that they influence at different stages of cognitive processing. We propose that preparation influences at the top of the visual hierarchy, whereas OSM interferes with processes occurring at lower levels of the hierarchy. These lower level processes play an essential role in visual awareness.

1. Introduction

In the classical hierarchical view of visual processing, neurons at low-level cortical areas (V1, V2) represent simple features such as orientation, color, and location (Hubel & Wiesel, 1968). They output to higher cortical levels (V3, V4, MT), which have lower spatial resolution and represent complex features. Object recognition and categorization occur when the activation reaches the highest levels (inferotemporal areas, prefrontal areas). Recent theories of visual processing and awareness emphasize that top-down (reentrant, recurrent, feedback) processes play an important role in visual perception, particularly in visual awareness (i.e., in subjective experience of seeing) (Bullier, 2001; Campana & Tallon-Baudry, 2013; Hochstein & Ahissar, 2002; Lamme, 2006; Lamme & Roelfsema, 2000). They argue that the feedforward flow of activation from low levels to the higher ones is not sufficient for vivid conscious perception, although it may enable unconsciously guided responding. The contribution of top-down feedback to the low-level areas is assumed to be either a general prerequisite for any kind of conscious perception to emerge (e.g., Lamme & Roelfsema, 2000; Lamme, 2006) or necessary for detailed conscious vision (Campana & Tallon-Baudry, 2013; Hochstein & Ahissar, 2002). For example, according to Reverse Hierarchy Theory (RHT) (Hochstein & Ahissar, 2002) the feedforward hierarchy acts nonconsciously and conscious perception emerges at high-level cortex, representing complex aspects or the gist of the scene. At a later stage of processing, conscious perception returns to lower areas via reentrant feedback connections, to integrate the detailed information into conscious vision. However, there has been little effort to dissociate behaviourally the effects of different types of top-down processes on visual awareness.

In contrast to theories of visual awareness, research on visual categorization of objects in natural scene backgrounds has stressed the power of bottom-up or feedforward processing. An influential study (Thorpe, Fize, & Marlot, 1996) showed that brain's electrical responses differentiated animal images from non-animal images already 150 ms after stimulus-onset. Later studies have confirmed the high-speed (Fabre-Thorpe, 2011), and measurements of saccadic latencies indicate that the fastest saccadic latencies toward target images may occur within 120 ms (Kirchner & Thorpe, 2006; Wu, Crouzet, Thorpe, & Fabre-Thorpe, 2015). In addition, heavy load on attention interferes only little with categorization (Fize, Fabre-Thorpe, Richard, Doyon, & Thorpe, 2005; Li, VanRullen, Koch, & Perona, 2002). The high speed of visual categorization and its independence from attentional load is consistent with the feedforward model of categorization.

On the other hand, categorization of objects in natural images at the basic level (e.g., dog vs. non-dog) is more demanding than at the superordinate level (e.g., animal vs. non-animal). Natural images are faster to categorize at the superordinate level ('animal') than at the basic level ('dog') (Macé, Joubert, Nespoulous, &

* Corresponding author.
E-mail address: mika.koivisto@utu.fi (M. Koivisto).

http://dx.doi.org/10.1016/j.visres.2017.01.005
0042-6989/

Fabre-Thorpe, 2009; Poncet, Reddy, & Fabre-Thorpe, 2012). Saccadic eye movements indicate that while superordinate level targets can be detected already 120 ms after the onset of the stimulus, basic level targets cannot be detected as quickly (Wu et al., 2015). Mack and Palmeri (2015) found the superordinate-level advantage compared to basic level with short stimulus duration but the effect reversed with longer durations into a basic-level advantage. However, a series of experiments (Poncet & Fabre-Thorpe, 2014) showed that the superordinate-level advantage is robust, and does not depend on the stimulus duration and thus is not a result of the coarser information quality associated with briefly flashed stimuli. Thus, the basic level categorization needs a longer information uptake time, but the kind of processing that is carried on during the extra time is not clear. Spatial attention is ruled out by the finding that basic-level categorization was possible during a concurrent task that put strong load on attentional resources (Poncet et al., 2012). One possibility is that conscious top-down processing is required for finer object representations that are required in discriminating different category members at basic level. Koivisto and Rientamo (2016) found that masked images produced priming for superordinate level (animal vs. non-animal) but not for basic level (dog vs. horse) categorization. Basic level priming occurred only when the prime images were not masked and thus consciously accessible. The (unconscious) feedforward flow of information might result in a coarse representation which allows discrimination between superordinate categories (Fabre-Thorpe, 2011; Koivisto, Kastrati, & Revonsuo, 2014) but is not detailed enough for making more fine-grained discriminations between basic categories.

In summary, the observations seem to converge on the conclusion that categorization at the superordinate level (e.g., animal or non-animal) may occur in feedforward manner, and possibly without the contribution of awareness (Koivisto & Rientamo, 2016; Koivisto et al., 2014), whereas categorization at the basic level (e.g., dog or non-dog) requires more time, perhaps an additional phase of top-down, feedback processing. However, it still remains possible that top-down processes do play some role in visual categorization at the superordinate level. The procedure in standard experiments on categorization is such that attention can be prepared in advance for detecting the target images. In each stimulus block, the participants categorize the target image according to whether it represents the category that is defined at the beginning of the task block (e.g., "press GO when there is an animal"). Such preparation or expectancy may bias processing in favour of the searched-for category (Peelen, Fei-Fei, & Kastner, 2009; Peelen & Kastner, 2011). In Experiment 1 we tested directly whether top-down preparation plays a role in visual categorization at superordinate or basic level as well as in visual awareness by manipulating the experimental task such that in the blocked condition, in which preparation was possible, the procedure followed the typical categorization protocol in which each target image was categorized according to whether or not it represented the category that was specified before the task block began. In the non-blocked condition, the category varied randomly from trial to trial and it was specified only after the offset of the target image. Thus, top-down preparation was possible in the blocked condition but not in the non-blocked condition. In Experiment 2, preparation was manipulated by presenting the relevant category name either only after the target image had been presented (post-cue) or also before the target image was presented (pre-cue).

In addition, both experiments used object substitution masking (OSM) (Enns & Di Lollo, 1997) to manipulate another type of process that is known to influence visual awareness of natural images (Koivisto et al., 2014). OSM is widely assumed to selectively influence the reentrant stage of processing (Di Lollo, Enns, & Rensink, 2000, but see Francis & Hermens, 2002; Põder, 2013) while leaving feedforward processes intact (e.g., Goodhew, Visser, Lipp, & Dux, 2011; Koivisto, 2012). In a typical OSM experiment, a target stimulus and a mask (e.g., dots surrounding the target) appear simultaneously in a display containing 1 to 16 stimuli but the offset of the mask is delayed so that it persists to be visible after the offset of the target. Compared with the simultaneous offset of the target and mask, the delayed offset of the mask impairs the visibility of the target. The object substitution theory (Di Lollo et al., 2000) explains this masking effect by assuming that after initial encoding of the stimulus display at a low level, feedforward activation proceeds to higher levels where a tentative, low resolution representation (a perceptual hypothesis) is formed. This representation involves ambiguity which can be resolved on the second and later iterations by comparing the high level coarse representation with the initial pattern of activity at the lower level. When the offset of the mask is delayed, the tentative representation (target + mask) does not match with information at the lower level (mask). Therefore, the representation (target + mask) is replaced or updated (Moore & Lleras, 2005) with that of the trailing mask and the observer perceives only the mask.

We studied the roles of the two types of top-down processes (attentional preparation and OSM-dependent processes) in visual categorization and awareness by manipulating both at the same time. The hypothesis that visual awareness depends on top-down processing (Hochstein & Ahissar, 2002; Lamme, 2006), whereas categorization at superordinate level may succeed on the basis of the feedforward sweep reaching the higher levels in hierarchy (Fabre-Thorpe, 2011; Thorpe et al., 1996), predicts that visual awareness, but not superordinate level categorization, should depend on manipulations of top-down processing. On the other hand, if more fine-grained discriminations between category instances at the basic level cannot be made on the basis of the feedforward sweep but require top-down processing (Koivisto & Rientamo, 2016), basic level categorization should be influenced by the top-down manipulations. The manipulation of preparation and OSM allowed us to study in more detail their relationship during categorization and awareness as it is not clear whether they influence the same or different stages of visual processing.

2. Experiment 1

2.1. Methods

2.1.1. Participants

Thirty-two healthy students (8 male; age 19–30 years) with normal or corrected-to-normal vision participated. The experiments were conducted in accordance with the Declaration of Helsinki and with the understanding and written consent of each participant. Experiment 1 was run in two different laboratories (16 participants in each laboratory) with two different experimenters but with the same equipment.

2.1.2. Stimuli

All visual stimuli were color photographs of object in natural scene backgrounds from the study by Koivisto and Rientamo (2016, Experiment 1). The images of animals and non-animals varied in luminance, color, and spatial frequency, and represented a mixture of general views and close-ups so that the categorization tasks could not be performed on the basis of low level visual features. The participants had not seen the photographs before. The positive stimuli (i.e. those requiring 'yes' responses) in the superordinate (animal/non-animal) categorization condition represented horses (n = 64), dogs (n = 64), and birds (n = 64) (total n = 192). The negative stimuli (total n = 96), requiring 'no' responses, represented vehicles (n = 32), buildings (n = 32) and a mixed category of