Vision Research 68 (2012) 40–47

Contents lists available at SciVerse ScienceDirect

Vision Research

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

The transfer of object learning across exemplars and their orientation is related to perceptual similarity

Annelies Baeck [a,b,*,1], Ine Windey [a,1], Hans P. Op de Beeck [a]

[a] Laboratory of Biological Psychology, University of Leuven (KU Leuven), Tiensestraat 102, 3000 Leuven, Belgium
[b] Laboratory of Experimental Psychology, University of Leuven (KU Leuven), Tiensestraat 102, 3000 Leuven, Belgium

ARTICLE INFO

Article history:
Received 27 February 2012
Received in revised form 6 June 2012
Available online 20 July 2012

Keywords:
Perceptual learning
Object recognition
Backward masking
Specificity
Exemplar
Orientation

ABSTRACT

Recognition of objects improves after training. The exact characteristics of this visual learning process remain unclear. We examined to which extent object learning depends on the exact exemplar and orientation used during training. Participants were trained to name object pictures at as short a picture presentation time as possible. The required presentation time diminished over training. After training participants were tested with a completely new set of objects as well as with two variants of the trained object set, namely an orientation change and a change of the exact exemplar shown. Both manipulations led to a decrease in performance compared to the original picture set. Nevertheless, performance with the manipulated versions of the trained stimuli was better than performance with the completely new set, at least when only one manipulation was performed. Amount of transfer to new images of an object was related to perceptual similarity, but not to pixel overlap or to measurements of similarity in the different layers of a popular hierarchical object recognition model (HMAX). Thus, object learning generalizes only partially over changes in exemplars and orientation, which is consistent with the tuning properties of neurons in object-selective cortical regions and the role of perceptual similarity in these representations.

1. Introduction

Perceptual learning is a constant learning process in which the visual representations in the brain are altered (Fahle & Poggio, 2002). It is defined as an increase in the ability to extract information from the environment, as a result of experience and practice (Gibson, 1969).

Early studies documented the properties of this perceptual learning process, and have used these properties to derive the specific location of the underlying changes in the brain. For example, it was found that perceptual learning tends to be rather specific in studies using relatively simple stimuli known to primarily activate low-level regions in the visual processing hierarchy, such as gratings and short line segments. For example, no transfer of learning was found towards spatial frequency (Yu, Klein, & Levi, 2004), contrast (Yu, Klein, & Levi, 2004), distinct visual learning tasks (Fahle, 2004; Fahle & Morgan, 1996; Poggio, Fahle, & Edelman, 1992) nor towards different orientations (Ahissar & Hochstein, 1996; Crist et al., 1997; Fahle, 2004; Sigman & Gilbert, 2000; Yu, Klein, & Levi, 2004). This specificity is consistent with the hypothesis that perceptual learning for these stimuli involves changes in low-level visual regions. This assumption has been confirmed in electrophysiological studies (Schoups et al., 2001), although some controversy remains (e.g., Ghose, Yang, & Maunsell, 2002).

In contrast to properties of perceptual learning with simple stimuli, a different picture emerged about the expected and empirically verified specificity of perceptual learning with more complex stimuli such as pictures of objects or faces (Hussain, Sekuler, & Bennet, 2009b), here referred to as object learning. Given that such stimuli activate higher visual regions, and given that these regions are traditionally considered to contain representations of objects that are invariant for changes in many of the aforementioned manipulations (e.g., Booth & Rolls, 1998; Wallis & Rolls, 1997), one can expect more transfer across these dimensions if these representations are involved in the learning process. Furmanski and Engel (2000) made use of an object-naming task and found evidence that learning with objects was specific to the trained object but indeed generalized towards the trained objects shown at a different size. Other generalization effects with complex stimuli apart from manipulations of size (Furmanski & Engel, 2000; Lee, Matsumiya, & Wilson, 2006), include a transfer between distinct visual learning paradigms (Baeck & Op de Beeck, 2010) and a partial transfer across orientation using upright and inverted houses as stimuli (Husk, Bennet, & Sekuler, 2007). Also with other types of paradigms, such as adaptation and repetition priming (e.g., Vuilleumier et al.,

---

* Corresponding author at: Laboratory of Biological Psychology, University of Leuven (KU Leuven), Tiensestraat 102, 3000 Leuven, Belgium.

E-mail addresses: Annelies.baeck@ppw.kuleuven.be (A. Baeck), windey.ine@gmail.com (I. Windey), Hans.OpdeBeeck@ppw.kuleuven.be (H.P. Op de Beeck).

[1] These authors contributed equally.

0042-6989/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.visres.2012.06.023

2002), similar effects have been reported. For example, with perceptual priming experiments have found that priming occurs across changes in size (Fisher & Biederman, 2001) and object position (Biederman & Cooper, 1991). These results endorse theories claiming visual information in the higher visual areas are stored independent of momentary viewing parameters.

However, the rationale of using the degree and type of transfer as an index of where in the brain learning occurs has turned out to be simplistic. First of all, the degree of transfer can depend upon how and in which context stimuli are shown during training. For example, Zhang et al. (2010) showed that orientation specific perceptual learning could transfer completely to an orthogonal orientation when the observers were exposed to the orthogonal orientation in an irrelevant task.

Second, the distinction between low-level and high-level representations in terms of invariance to image transformations is not clear-cut. Recent studies have suggested that high-level visual representations are much more sensitive to a wide range of object transformations than suggested by the high degree of invariance in behavior. Experimental findings have indeed shown a surprising degree of position information in these representations, both in monkeys (DiCarlo & Maunsell, 2003; Op de Beeck & Vogels, 2000) and in humans (Kravitz, Kriegeskorte, & Baker, 2010; Schwarzlose et al., 2008). Neural responses in these brain regions are also affected by changes in viewpoint (Freiwald & Tsao, 2010; Grill-Spector et al., 1999; Logothetis, Pauls, & Poggio, 1995), and the exact exemplar shown of a particular object type (Vogels, 1999). Computational models have suggested that the discrepancy between the tuning properties of single neurons and invariance in behavior can be explained by the fact that (i) behavior depends on the pattern of activity across a whole population of neurons (Hung et al., 2005; Zoccolan et al., 2007), and (ii) objects are typically seen across multiple transformations (Goris & Op de Beeck, 2010).

Given the non-negligible sensitivity of high-level visual neurons for a wide range of image transformations, it is no longer a straightforward prediction that object learning would transfer across such transformations. At least a partial specificity should be found. A few studies have already confirmed that learning about objects can be specific to viewpoint (Lee, Matsumiya, & Wilson, 2006) and to retinal position (Kravitz, Vinson, & Baker, 2008). The latter review stressed however that the specificity of learning might be dependent on the exact paradigm used.

Here we further test the specificity of object learning using the paradigm of Furmanski and Engel that suggested that object learning generalizes across size (Furmanski & Engel, 2000) and across the type of noise added to the object images (Baeck & Op de Beeck, 2010). The present experiment included two new manipulations, object exemplar and orientation (in the image plane). These two manipulations were chosen amongst others because they are very different in nature. Orientation manipulations are widely tested with simple stimuli such as gratings (e.g., Ahissar & Hochstein, 1996; Crist et al., 1997; Fahle, 2004; Ghose, Yang, & Maunsell, 2002; Logothetis, Pauls, & Poggio, 1995; Sigman & Gilbert, 2000; Yu, Klein, & Levi, 2004), but not often with complex, everyday objects (but see Husk, Bennet, & Sekuler, 2007; Hussain, Sekuler, & Bennet, 2009a). This kind of manipulation changes the physical appearance of the object, but the identity remains the same. Other examples of manipulations that preserve the identity (with complex objects) are a position change (Stringer & Rolls, 2000) and changes in viewpoint (Stone, 1999). On the contrary, manipulations in object exemplar change both the physical appearance and the identity of the presented object.

These two types of changes are treated very differently in theories of object recognition (e.g., Biederman, 1987; Riesenhuber & Poggio, 1999), as it is assumed that invariance is built up for identity-preserving transformations (here represented by orientation) whereas selectivity is preserved or even enhanced for identity changes. This distinction is for example very explicit in the models of Riesenhuber and Poggio (1999) and Poggio and Edelman (1990), which were in large part validated by the paperclip identification experiments of Logothetis, Pauls, and Poggio (1995). This distinction is also in line with experimental findings of neurons with high selectivity for individual objects combined with high invariance in human cortex (e.g., the famous Jennifer Aniston neuron described by Quiroga et al. (2005)). Recent theoretical and methodological (e.g., pattern classification) developments suggest that this dichotomy might not be so strict because perfect invariance is not a goal (e.g., DiCarlo & Cox, 2007). Our choice of transformations, although still limited in extent, allows a first comparison of an identity-preserving transformation with an identity change. As indicated above, high-level neurons are sensitive to differences between exemplars and changes in orientation, but, if asked for, humans can easily generalize across exemplars and orientation in a wide range of behavioral tasks (Ashby & Maddox, 2005).

Participants were trained to name briefly presented object images in a backward-masking paradigm, with 5 days of practice with the same stimulus set during which the time of presentation was gradually decreased in an adaptive manner. After the training, the performance was tested with four different stimulus sets: (1) a new object set, (2) the original objects presented in a different, untrained orientation, (3) untrained exemplars from the original object set and (4) a combination of the two last manipulations. We replicated the object-specific training effect from earlier studies (Baeck & Op de Beeck, 2010; Furmanski & Engel, 2000). In addition, the training-induced improvement generalized to untrained exemplars and untrained orientations, but only partially. Finally, we determined that this degree of transfer was related to the perceived similarity among the trained stimuli and the transformed stimuli, in contrast to other measures of similarity such as physical pixel-based similarity and similarity according to the computational model of Riesenhuber and Poggio (1999).

2. Method

2.1. Participants

Sixteen students of the University of Leuven (KU Leuven) participated in the main experiment. Among them were 5 men and 11 women (ages between 19 and 23) who were naïve with respect to the aim of the study participated in this study as paid volunteers. A separate group of eight participants (2 male, ages between 22 and 33) participated in the subjective rating experiment. All participants had a normal or corrected-to-normal sight. The experiments were approved by the ethical committee of the Faculty of Psychology and Educational Sciences (KU Leuven) and participants signed an informed consent at the start of the first session.

2.2. Apparatus

Stimuli were presented by a Toshiba laptop using Matlab 6.0 (Psychtoolbox 2.54) in a darkened room. The 100 Hz 22 in. screen was gamma corrected. The viewing distance was fixed at 94 cm by a chin support device.

2.3. Stimuli

Forty objects were selected. Criteria for the selection of objects were an easy recognition and few available synonyms for the same object. Object images were converted to gray-scale. For each object two exemplars in the same orientation were included. Each exemplar was also rotated, either 90° (for the 20 objects that had the