Vision Research 91 (2013) 36–44

Contents lists available at SciVerse ScienceDirect

Vision Research

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

Object recognition in congruent and incongruent natural scenes: A life-span study

F. Rémy [a,b,*,1], L. Saint-Aubert [a,b,c,1], N. Bacon-Macé [a,b], N. Vayssière [a,b], E. Barbeau [a,b], M. Fabre-Thorpe [a,b]

[a] Université de Toulouse, UPS, Centre de recherche Cerveau et Cognition, Toulouse, France
[b] CNRS, CerCo, Toulouse, France
[c] Inserm, Imagerie Cérébrale et Handicaps Neurologiques UMR 825, CHU Purpan, Place du Dr Baylac, F-31059 Toulouse Cedex 9, France

ARTICLE INFO

Article history:
Received 11 October 2012
Received in revised form 15 April 2013
Available online 25 July 2013

Keywords:
Aging
Visual object recognition
Natural scenes
Scene congruence

ABSTRACT

Efficient processing of our complex visual environment is essential and many daily visual tasks rely on accurate and fast object recognition. It is therefore important to evaluate how object recognition performance evolves during the course of adulthood. Surprisingly, this ability has not yet been investigated in the aged population, although several neuroimaging studies have reported altered activity in high-level visual ventral regions when elderly subjects process natural stimuli. In the present study, color photographs of various objects embedded in contextual scenes were used to assess object categorization performance in 97 participants aged from 20 to 91. Objects were either animals or pieces of furniture, embedded in either *congruent* or *incongruent* contexts. In every age group, subjects showed reduced categorization performance, both in terms of accuracy and speed, when objects were seen in incongruent vs. congruent contexts. In subjects over 60 years old, object categorization was greatly slowed down when compared to young and middle-aged subjects. Moreover, subjects over 75 years old evidenced a significant decrease in categorization accuracy when objects were seen in incongruent contexts. This indicates that incongruence of the scene may be particularly disturbing in late adulthood, therefore impairing object recognition. Our results suggest that daily visual processing of complex natural environments may be less efficient with age, which might impact performance in everyday visual tasks.

1. Introduction

Aging has long been known to affect visual perception (Weale, 1975). So far, most research has focused on low-level visual deficits in aging (for review see Owsley, 2011; Sekuler, Hutman, & Owsley, 1980; Werner, Peterzell, & Scheetz, 1990), with age-related alterations reported for acuity (Pitts, 1982; Weale, 1975), color perception (Hardy et al., 2005), dark adaptation (Jackson, Owsley, & McGwin, 1999), motion perception (Sekuler, Hutman, & Owsley, 1980; Wojciechowski, Trick, & Steinman, 1995), and contrast sensitivity (McKendrick et al., 2007; Owsley, Sekuler, & Siemsen, 1983). These low-level deficits may impact detection or recognition of natural stimuli encountered in daily life, such as faces, objects or scenes. For example, alterations in the processing of stimuli such as facial identity and emotional facial expressions have been shown from the age of 50 years old and increasing after 70 (Bartlett & Fulton, 1991; Boutet & Faubert, 2006; Crook & Larrabee, 1992; Daniel & Bentin, 2010; Grady, 2002; Isaacowitz et al., 2007). On the other hand, old age performance in face perception may depend upon more integrated high-level processes (Anstey et al., 2002). Moreover, in daily life, objects of interest are not shown in isolation, but embedded in rich visual backgrounds. The present life-span study aimed to investigate whether and how performance in recognition of natural common objects presented in contextual scenes is affected during the course of adulthood.

Studies in young subjects have shown that the human visual system is extremely fast and efficient at detecting and categorizing objects presented in their natural context (Fabre-Thorpe et al., 2001; Rousselet, Mace, & Fabre-Thorpe, 2003; Thorpe, Fize, & Marlot, 1996). The context can also influence object recognition performance (Biederman, 1972; Biederman, Mezzanotte, & Rabinowitz, 1982; Biederman et al., 1974; Palmer, 1975). When objects are embedded in a familiar context (e.g. a cow in a field), object recognition is both faster and more accurate than when objects are presented in an incongruent context, in which they are less likely to appear (e.g. a cow in an office) (Davenport, 2007; Fize, Cauchoix, & Fabre-Thorpe, 2011; Joubert et al., 2008). This influence of context on object recognition relies on the lifelong repetitive experience of the visual system with our visual surrounding world and

* Corresponding author. Address: Centre de Recherche Cerveau et Cognition, Université Toulouse 3, CNRS, Pavillon Baudot, CHU Purpan, BP 25202, 31052 Toulouse Cedex, France. Fax: +33 5 62 74 61 17.

E-mail address: florence.remy@cerco.ups-tlse.fr (F. Rémy).

[1] These authors contributed equally to this study.

0042-6989/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.visres.2013.07.006

its efficiency at extracting visual regularities. When faced with environmental scenes as in daily conditions, humans are likely to process simultaneously objects and context in the scene. Both processes may interact early and result in either facilitated (if congruent) or impaired (if incongruent) object recognition.

How is visual object recognition performance affected with age? What is the influence of the surrounding context on object recognition performance? How strong is this influence in elderly relatively to young and middle-aged subjects? Two previous imaging studies have reported age-related changes in activity during shallow encoding (Park et al., 2004) or passive viewing (Chee et al., 2006) of pictures of scenes and objects. Elderly subjects evidence reduced selectivity in temporal and occipital areas of the visual ventral pathway, suggesting that fast bottom-up processing of natural stimuli may be less efficient with age (Park et al., 2004). Moreover, whereas the activity of the parahippocampal cortex related to context processing, appears preserved with age, the activity in the lateral occipital region related to object processing shows age-related changes, suggesting altered object processing in elderly subjects (Chee et al., 2006). Accordingly, parallel bottom-up processing of both object and context in the scene may be deficient in older adults, which may result in impaired contextual binding (Chee et al., 2006). We hypothesize that, when engaged in rapid categorization of natural objects embedded in contextual scenes, elderly subjects' performance would be impaired when compared to younger subjects. We further hypothesized that the influence of context on object recognition performance would increase with age. On one hand, older subjects may rely on a lifetime-based functional shaping of their visual system by repetitive experience, ensuring optimal processing of familiar (congruent) contextual presentations of objects. On the other hand, incongruent object–context associations, which have not been previously experienced and registered in the visual system, may result in deficient contextual binding and impaired object processing. Knowledge of these deficits is crucial because of their possible impact on everyday life. In particular, deficits in recognition of unexpected (incongruent) objects in the environment could have negative implications in daily life, e.g. during vehicle driving.

2. Materials and methods

2.1. Participants

Ninety-seven subjects aged from 20 to 91 were recruited for the study. Four age groups (young, middle-aged, old and very old) were considered as follows:

- group 1 (G1) included 20- to 30-years-old participants ($n = 22$),
- group 2 (G2) included 45- to 55-years-old participants ($n = 26$),
- group 3 (G3) included 60- to 75-years-old participants ($n = 23$),
- group 4 (G4) included participants over 75 years old ($n = 26$).

Information on subjects is shown in Table 1. All subjects were free from any neurological disease. They all reported having normal or corrected-to-normal vision. Two visuo-motor tasks were used as preliminary tests to check the subjects' ability to perform the main visual task (see below). The study was approved by the Ethics Committee of Lille, France (CPP no. 09/68 Nord-Ouest IV) and all subjects gave their written informed consent.

2.2. Stimuli

Two hundred stimuli (real-life color pictures, 778 × 518 pixels) were created for the study. Pictures were in 24-bit bitmap format (16 million colors). All stimuli consisted in a target-object embedded in a real-life background. Object pictures were selected from the Hemera Photo Objects library. Objects were either animals (excluding human and common domestic animals, e.g. rabbit, dog, cat, etc.) or pieces of furniture (e.g. sofa, table, dresser, etc.). Background pictures were selected from a large commercial CD-ROM library (Corel Stock Photo Libraries) or from the Internet. Backgrounds were either "natural" or "man-made" scenes. "Natural" scenes included seascapes and landscapes (mountain, desert, beach, iceberg, forest and fields) without any buildings, while "man-made" scenes were indoor scenes (room, kitchen, corridor, church, terrace, etc.) without any foreground objects.

To control for low-level visual differences between stimuli due to the use of real-life pictures, fifty sets of 4 images (object pasted in a background context) were created with a home-made software (Fize, Cauchoix, & Fabre-Thorpe, 2011) and using the following procedure. Each set of four stimuli was alternatively combining two different backgrounds (natural and man-made) with two different objects (animal and furniture). In each set, care was taken to select natural and man-made backgrounds in the same range of colors, with similar layout and complexity. Also, animal and furniture objects were chosen with comparable colors and shapes. Backgrounds were equalized in luminance and contrast. Object sizes were equalized in number of pixels (i.e. object area). Objects were pasted at identical positions in each of the backgrounds, and respecting rules for interposition, support and size (Biederman, Mezzanotte, & Rabinowitz, 1982). Progressive transparency (2 pixels wide) was applied on the object contours in order to avoid sharp edges. The 4 images resulted in two *congruent* object–context associations (an animal in a natural context and a piece of furniture in an indoor man-made context) and two *incongruent* object–context associations (an animal in an indoor man-made context and a piece of furniture in a natural context) that were finally equalized again in contrast and luminance. Examples of sets of 4 stimuli are shown in Fig. 1 (all sets can be seen at http://cerco.ups-tlse.fr/StimuliSaintAubert/).

Considering all 200 stimuli, mean luminance was of 124.1 ± 0.3 (on a linear scale from 0 (black) to 255 (white)), with a variance of 57.1 ± 0.7. Average object size was rather large (12.7 ± 4.7% of the image) so that stimuli were suitable for older subjects. Object sizes ranged from 1.8° × 3.6° for the smallest to 10.5° × 6.4° for the largest. Moreover, the power spectral signatures of the stimulus categories was computed (Torralba & Oliva, 2003) and, consistent with the literature, they revealed classical features for both natural and man-made contexts (Fig. 1) regardless of the object embedded in the contexts. Thus, unlike reported for non-manipulated stimuli

Table 1
Description of the four groups of age. Age ranges are mentioned in brackets. Groups' performance on the preliminary tasks is indicated. Differences between groups were tested when appropriate. Values are mean ± SD.

	Group 1 (20–30)	Group 2 (45–55)	Group 3 (60–75)	Group 4 (>75)	Significance
Mean age (years)	24.0 ± 3.0	50.5 ± 2.5	65.1 ± 4.8	82.1 ± 6.3	
Gender (females/males)	12/10	15/11	11/12	12/9	$p = 0.9$
Accuracy on basic color categorization task (%)	96.8 ± 3.6	96.2 ± 2.7	96.4 ± 2.3	94.0 ± 7.7	$p = 0.37$
Accuracy on object categorization training task (%)	94.6 ± 4.3	96.4 ± 3.4	93.2 ± 9.3	92.0 ± 5.7	$p = 0.07$
Mean RT on object categorization training task (ms)	501 ± 71	543 ± 56	588 ± 94	617 ± 99	$p < 0.001$