PERGAMON

Neuropsychologia 40 (2002) 917–930

NEUROPSYCHOLOGIA

www.elsevier.com/locate/neuropsychologia

Mental rotation versus invariant features in object perception from different viewpoints: an fMRI study

Jan Vanrie [a], Erik Béatse [b], Johan Wagemans [a,*], Stefan Sunaert [b], Paul Van Hecke [b]

[a] *Department of Psychology, University of Leuven, Tiensestraat 102, B-3000 Leuven, Belgium*
[b] *Department of Radiology, University Hospital Gasthuisberg, University of Leuven, Leuven, Belgium*

Received 8 January 2001; accepted 25 July 2001

Abstract

It has been proposed that object perception can proceed through different routes, which can be situated on a continuum ranging from complete viewpoint-dependency to complete viewpoint-independency, depending on the objects and the task at hand. Although these different routes have been extensively demonstrated on the behavioral level, the corresponding distinction in the underlying neural substrate has not received the same attention. Our goal was to disentangle, on the behavioral and the neurofunctional level, a process associated with extreme viewpoint-dependency, i.e. mental rotation, and a process associated with extreme viewpoint-independency, i.e. the use of viewpoint-invariant, diagnostic features. Two sets of 3-D block figures were created that either differed in handedness (original versus mirrored) or in the angles joining the block components (orthogonal versus skewed). Behavioral measures on a same–different judgment task were predicted to be dependent on viewpoint in the rotation condition (same versus mirrored), but not in the invariance condition (same angles versus different angles). Six subjects participated in an fMRI experiment while presented with both conditions in alternating blocks. Both reaction times and accuracy confirmed the predicted dissociation between the two conditions. Neurofunctional results indicate that all cortical areas activated in the invariance condition were also activated in the rotation condition. Parietal areas were more activated than occipito-temporal areas in the rotation condition, while this pattern was reversed in the invariance condition. Furthermore, some areas were activated uniquely by the rotation condition, probably reflecting the additional processes apparent in the behavioral response patterns.

Keywords: Object recognition; Orientation specificity; Neuroimaging

1. Introduction

An intriguing feature of our visual system is the efficiency with which it is able to recognize objects, even when seen from different viewpoints. The issue of object recognition has been investigated extensively and several theoretical models have been proposed to account for this capacity. Over the years, the theoretical accounts as well as the way in which they are tested have evolved significantly. However, if one reviews the recent literature on the topic, two trends are especially prevalent.

The first trend can be described as a shift in the focus of discussion. In previous years, the debate was primarily concentrated on whether object recognition was "viewpoint-dependent" (e.g. [43]) or "viewpoint-independent" (e.g. [4]). In this debate, the two "sides" have used a variety of stimuli, experimental paradigms and techniques to test the predictions generated by their models and both, in fact, provided substantial empirical evidence in their favor. Given this growing amount of conflicting results, it became increasingly difficult to hold the view of a single, unitary system, which operates in a fixed manner. For this reason, notions such as "multiple routes to object recognition" are adopted more and more frequently in vision science in recent years (e.g. [28]). More specifically, there seems to be a growing consensus that there is a continuum from viewpoint-independency to viewpoint-dependency, which is in part influenced by stimulus discriminability and the task at hand (e.g. [3,12,19,20,47]). In short, the key question is no longer *if* object recognition is viewpoint-dependent or viewpoint-independent, but rather *when*, i.e. under which circumstances.

The second important trend is more general and consists of the increasing attention for the neurofunctional correlates of mental processes, including processes underlying visual perception. This line of research has obviously been strongly stimulated by the advent of neuroimaging techniques like PET and fMRI (for reviews see [10,15,44,54]).

* Corresponding author. Tel.: +32-16-325969; fax: +32-16-326099.
E-mail address: johan.wagemans@psy.kuleuven.ac.be (J. Wagemans).

0028-3932/02/$ – see front matter
PII: S0028-3932(01)00161-0

If we take these two trends into account, it is somewhat surprising that relatively little research has been done on the cortical areas associated with different routes to object recognition. This is precisely the main goal of the present investigation. Although some neuroimaging studies have addressed the issue of perceiving objects from different viewpoints (e.g. [18,26,38]), there have been—to our knowledge—no studies that explicitly tried to contrast these routes on the neurofunctional level.[1] Nevertheless, differences in the underlying cortical activations would provide strong convergent evidence that there are indeed distinct processes involved.

In order to investigate neurofunctional correlates of different routes to object perception, we first need to create circumstances in which different processes will be invoked.

Fortunately, we know of such circumstances in which recognition is undoubtedly viewpoint-dependent or viewpoint-independent. An extreme example of the former is the case of mental rotation, while the latter is clearly obtained when the visual system makes use of locally diagnostic, viewpoint-invariant features.

1.1. Mental rotation

A special case in object recognition is the situation in which the visual system needs to distinguish an object from its mirrored counterpart. It has been shown that in this case, the visual system engages in a process of mental rotation. This visuo-spatial process, first described by Shepard and Metzler [37], yields very robust viewpoint-dependent effects:[2] response times in a same–different judgment increased linearly with increasing angular difference (AD) between two comparison objects. Although making such a distinction might not be a typical object recognition task (see also [42,58]), it nevertheless, "involves extensive recognition and comparison of objects and object parts" ([5], p. 11, on the Shepard and Metzler paradigm).[3] Because a detailed computational model of the mental rotation paradigm is beyond the scope of the present study (see [5,22]) we will focus on one aspect relevant to the issue at hand: the use of a viewer-centered reference frame. Obviously, advocates of viewpoint-dependent theories are likely to endorse the use of such a reference frame. Since handedness is not encoded in object-centered representations [4,40], the use of a viewer-centered frame of reference becomes necessary in case of handedness discrimination, even according to viewpoint-independent theories. More specifically, the computation of the spatial coordinates of the object being mentally rotated will proceed from the specific viewpoint in which the observer perceives the objects. This makes the mental rotation process a highly appropriate research subject in the present discussion.

1.2. Use of invariant features

If the mental rotation process is an extreme case of viewpoint-dependency, the process in which observers make use of locally diagnostic, view-invariant features to determine the identity of an object (e.g. [13]), can be seen as an equally extreme case of viewpoint-independent object processing. Again, this process is not the most typical for everyday object recognition, but the operation of detecting and exploiting image invariances is clearly an important and relevant topic (e.g. [53]). Moreover, in the present experiment, top-down influences are likely to come into play, because the observers know what kind of feature is crucial to solve the task. That is, it is advantageous for task performance to direct attention to those locations in which the feature is present. In this sense, making use of diagnostic viewpoint-invariant features cannot be reduced to simple shape discrimination.

By including these rather extreme processes, mental rotation and the use of invariant, diagnostic features, we can expect clear and unambiguous behavioral results. In addition, neuropsychological investigations also demonstrated that the two processes can be selectively impaired. This constitutes a strong indication that—at least in brain damaged patients—they are indeed associated with different neural systems. For example, patient R.T., described by Farah and Hammond [16], was still able to recognize misoriented objects, despite his inability to perform mental rotation (see also [8,11,24]). Turnbull and McCarthy [46], on the other hand, reported the opposite pattern of a patient who succeeded on mental rotation tasks, but failed to recognize disoriented objects. Moreover, the double dissociation between "recognition by axes" and "recognition by features" (as described by Humphreys and Riddoch [21]) demonstrates the functional independence of a recognition process based on the detection and use of a distinctive feature when subjects are matching objects from different viewpoints. Four of the patients in the study by Humphreys and Riddoch [21] were able to match objects that were rotated in the plane even though the saliency of their primary distinctive feature was reduced. On the other hand, when shown the objects with their principal axis foreshortened, leaving the visibility of the distinctive feature unaltered, they failed to make the match. The reverse pattern was found in a fifth patient.

[1] To illustrate this point, while Vecera [49] has recently reviewed the literature on object representation and invariances in the context of a special issue on neuroimaging, he mentioned only Kosslyn et al. [26] as a neuroimaging study directly relevant to the viewpoint- versus object-centered debate.

[2] Although the coupling of viewpoint-dependency and increased response times on the one hand, and viewpoint-independency and no effect on response times on the other hand is widely accepted, some evidence suggests that this assumption should be treated with caution [52].

[3] It is worth stressing that we do not assume that our "mental rotation" condition recruits all forms of viewpoint-dependent process; nor do we assert that our "invariance" condition taps into all viewpoint-independent recognition processes. What we do claim, however, is that both conditions encompass a number of processes relevant to the issue of perceiving objects from different perspectives, and that the two conditions are representative of viewpoint-dependent and viewpoint-independent routes to object recognition, as distinguished in the literature.