Neuropsychologia ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Neuropsychologia

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

# What is special about expertise? Visual expertise reveals the interactive nature of real-world object recognition ☆

Assaf Harel [a,b,*]

[a] Department of Psychology, Wright State University, Dayton, OH, United States
[b] Neuroscience Institute, Wright State University, Dayton, OH, United States

ARTICLE INFO

*Article history:*
Received 8 February 2015
Received in revised form
2 June 2015
Accepted 3 June 2015

*Keywords:*
Perceptual expertise
Vision
Object recognition
FFA
Top-down control
Attention
Review

ABSTRACT

Ever since Diamond and Carey (1986. *J. Exp. Psychol.: Gen.*, vol. 115, pp. 107–117) seminal work, the main model for studying expertise in visual object recognition ("visual expertise") has been face perception. The underlying assumption was that since faces may be considered the ultimate domain of visual expertise, any face-processing signature might actually be a general characteristic of visual expertise. However, while humans are clearly experts in face recognition, visual expertise is not restricted to faces and can be observed in a variety of domains. This raises the question of whether face recognition is in fact the right model to study visual expertise, and if not, what are the common cognitive and neural characteristics of visual expertise. The current perspective article addresses this question by revisiting past and recent neuroimaging and behavioural works on visual expertise. The view of visual expertise that emerges from these works is that expertise is a unique phenomenon, with distinctive neural and cognitive characteristics. Specifically, visual expertise is a controlled, interactive process that develops from the reciprocal interactions between the visual system and multiple top-down factors, including semantic knowledge, top-down attentional control, and task relevance. These interactions enable the ability to flexibly access domain-specific information at multiple scales and levels guided by multiple recognition goals. Extensive visual experience with a given object category culminates in the recruitment of these multiple systems, and is reflected in widespread neural activity, extending well beyond visual cortex, to include higher-level cortical areas.



## 1. Introduction

Human expertise in visual object recognition provides a unique opportunity to study how cognitive and neural representations change with experience. Ever since Diamond and Carey's seminal work (1986), the main impetus for the study of expertise in visual object recognition (referred henceforth as visual expertise) came from the debate regarding the modularity of face perception. Visual expertise has been invoked as an alternative to the idea that faces are special (i.e., processed in a distinct fashion and represented by distinct neural mechanisms). Under this view, faces were considered as an example of a category with which humans have gained expertise with, and thus, it was argued, any category of expertise sharing the same computational demands with faces, should elicit similar face-like signatures. This led to an extensive body of research studying face-processing mechanisms in visual expertise (for a review see, Sheinberg and Tarr, 2010). However, while there is no doubt that human observers are experts in face recognition, visual expertise is obviously not restricted to faces. Consequently, one may ask whether face recognition is actually the right model to study visual expertise, and what are the consequences of studying visual expertise based on the face recognition model. I will argue in this perspective article that to gain a full understanding of the cognitive and neural substrates of visual expertise, one should study expertise in and of itself, free of any constraints imposed by the face perception analogy. I will propose that while some similarities between face processing and visual expertise certainly exist, visual expertise is unique, and presents a special case of visual recognition. Visual expertise exemplifies an "enhanced", ecological form of visual object recognition that emerges from the reciprocal interactions between multiple top-down factors, such as semantic knowledge, attention, and task relevance, and the visual system. Extensive visual experience with an object category culminates in the recruitment of

☆Perspective article submitted to Neuropsychologia special issue on functional selectivity in neuronal cognitive and perceptual systems.

* Correspondence address: Department of Psychology, Wright State University, 3640 Colonel Glenn Highway, Dayton, OH 45435, United States.
*E-mail address:* assaf.harel@wright.edu

http://dx.doi.org/10.1016/j.neuropsychologia.2015.06.004
0028-3932/

these multiple systems, and manifests in widespread neural activity, extending well beyond visual cortex, to include multiple high-level areas across the cortex. Notably, studying the different networks that form the neural correlates of expertise may inform us of the diverse cognitive processes involved in particular domains of expertise, as these processes are often not consciously accessible for the experts themselves (Barton et al., 2009; Palmeri et al., 2004; Shen et al., 2014). I will start by describing how following the face model, visual expertise was suggested as a principle of cortical organisation in occipito-temporal cortex (OTC). I will then elaborate on the reasons why visual expertise should be studied independently of faces, and suggest a new framework for visual expertise, the interactive framework for expertise (Harel et al., 2013). I will follow this with a thorough discussion of neuroimaging and behavioural studies of real world visual expertise, which together provide strong support for the interactive view of expertise.

## 2. Category selectivity and face specificity

In order to understand the reason that visual expertise research was guided for many years by the face perception analogy, one must turn to the question of functional selectivity and modularity of the mind. One of the most prominent findings in the cognitive neuroscience of vision has been the presence of category-selective regions in OTC. Many regions in OTC exhibit enhanced neural response to the visual presentation of certain object categories relative to other object categories. This robust finding has been demonstrated across a variety of experimental approaches, including single-cell recording in nonhuman primates (Baker et al., 2002; Freedman et al., 2003; Suzuki et al., 2006; Tsunoda et al., 2001), functional Magnetic Resonance Imaging (fMRI) in human – (Grill-Spector and Malach, 2004; Grill-Spector, 2003; Martin, 2007) and nonhuman- primates (Bell et al., 2009; Orban et al., 2004; Tsao et al., 2003), transcranial Magnetic Stimulation (Pitcher et al., 2009; Urgesi et al., 2004), as well as intracranial (Liu et al., 2009; Privman et al., 2007; Vidal et al., 2010) and scalp recordings in humans (Kiefer, 2001; Rossion et al., 2003; Rousselet et al., 2008). Selective lesions to OTC result in deficits in the recognition of specific object categories, implicating the causal role of OTC in representing information about objects (Gainotti, 2000; Mahon and Caramazza, 2009). One of the prime examples of category selectivity is face perception. Faces have been found to be distinguished from other object categories using many of the methods outlined above (Calder et al., 2011). For example, faces engage a dedicated network of cortical regions (Haxby et al., 2000; Ishai et al., 2005; Nestor et al., 2011), they elicit a unique electrophysiological response at least as early as 170 ms after stimulus onset (Bentin et al., 1996), and specific lesions to posterior fusiform gyrus result in the loss of the ability to recognise faces (De Renzi et al., 1991; Hecaen and Angelergues, 1962). Lastly, faces are processed more holistically compared to non-face objects, which are usually processed in a piecemeal fashion (Maurer et al., 2002).[1] Accordingly, face perception came to be thought of as exemplifying the idea that visual categories are uniquely represented by distinct neural and cognitive mechanisms (Malach et al., 2002; Reddy and Kanwisher, 2006), in line with modular domain-specific accounts of cognition (Kanwisher, 2010; Pylyshyn, 1999).

## 3. Expertise as a principle of cortical organisation

Other principles besides the conceptual category of the object, however, have been suggested to account for the large-scale organisation of OTC, including the unique manifestations of face processing. According to such domain-general accounts, faces do not have a specialized status as a category, and any difference they exhibit relative to other categories simply reflects their variation along other independent dimensions (i.e. principles of OTC organisation). Among the large-scale principles of cortical organisation of OTC that have been proposed are object attributes such as animacy and manipulability (Chao et al., 1999), real-world size (Konkle and Oliva, 2012), object shape (Op de Beeck et al., 2008), visual field eccentricity biases (Levy et al., 2001), and visual experience (Gauthier, 2000; Tarr and Gauthier, 2000).[2]

The latter account, also known as the process or expertise account, is particularly relevant for the current discussion, as it provides the backdrop for many of the works that investigated expertise in visual object recognition in the last two decades. The general idea behind this account, is that through experience, people learn to associate particular types of visual information distinctive to object categories with task-specific recognition strategies which automatically recruit components of a neural 'process map' (Bukach et al., 2006; Gauthier, 2000; Tarr and Gauthier, 2000; Wong et al., 2012). From this perspective, visual experience, and specifically, visual expertise, plays a critical role in the shaping of cortical object representations in high-level visual cortex by 'gluing together' stimulus information on the one hand, and characteristic task demands on the other. According to the process/expertise hypothesis, specifying the mechanisms (both task and visual information) underlying expertise in object recognition will ultimately uncover the general organisation of object representations in OTC. In this context, face perception seems the ideal domain to study these ideas. Indeed, in what other category do humans have more expertise than with faces? Humans are extremely adept at discriminating and recognising individual faces (Moscovitch et al., 1997; Tanaka, 2001), even though this should be a very difficult perceptual task (a set of highly homogeneous stimuli with a very similar configuration of parts). According to the expertise hypothesis, in order to overcome this inherent difficulty, face-dedicated regions in OTC became specialized during development for the processes underlying expert (face) perception, key among which is individuation (Bukach et al., 2006). Over time, face individuation becomes automatic and effortless as people gain experience in utilising the holistic information contained in face images (Gauthier and Nelson, 2001; Mondloch et al., 2002, 2006). An essential assertion of the expertise hypothesis is that these areas in OTC developed for the general task of within-category discrimination, and not solely for face individuation. The corollary of this view is that any expert category that requires the same type of visual information and shares the same task demands with faces (i.e. any category of objects with a prototypical configuration of parts that through experience its recognition becomes holistic) should engage "face-selective" areas in OTC (Gauthier and Tarr, 2002; McGugin et al., 2012).

Following this strong prediction, numerous works examined

[1] Broadly defined, holistic processing refers to the calculation of the relations between the parts of the object rather than the piecemeal processing of individual object parts (for a review see Maurer et al., 2002). The term holistic is notorious in the face perception literature for its many definitions and associations (Gauthier and Tarr, 2002). In the present article, I use the term holistic in its most general, inclusive sense subsuming first- and second-order configural representations as well as holistic (integral) processing.

[2] Note that these principles of organisation are not necessarily mutually exclusive and may interact with one another. Thus, recent work shows how both object animacy and object size are represented in OTC (Konkle and Caramazza, 2013) and eccentricity biases may develop following particular visual experiences (Malach et al., 2002).