Perception, 1999, volume 28, pages 483 – 487

DOI:10.1068/p2901

Viewpoint-dependent recognition of familiar faces

Nikolaus F Troje¶
Max-Planck Institut für biologische Kybernetik, Spemannstrasse 38, 72076 Tübingen, Germany
Daniel Kersten
Department of Psychology, University of Minnesota, Minneapolis, MN 55455, USA
Received 10 September 1998, in revised form 3 February 1999

Abstract. The question whether object representations in the human brain are object-centered or viewer-centered has motivated a variety of experiments with divergent results. A key issue concerns the visual recognition of objects seen from novel views. If recognition performance depends on whether a particular view has been seen before, it can be interpreted as evidence for a viewer-centered representation. Earlier experiments used unfamiliar objects to provide the experimenter with complete control over the observer's previous experience with the object. In this study, we tested whether human recognition shows viewpoint dependence for the highly familiar faces of well-known colleagues and for the observer's own face. We found that observers are poorer at recognizing their own profile, whereas there is no difference in response time between frontal and profile views of other faces. This result shows that extensive experience and familiarity with one's own face is not sufficient to produce viewpoint invariance. Our result provides strong evidence for viewer-centered representations in human visual recognition even for highly familiar objects.

1 Introduction

Viewpoint-dependent recognition has been reported for several classes of novel objects by researchers using a number of experimental paradigms (Bülthoff and Edelman 1992; Rock 1973; Tarr 1995). Recognition of unfamiliar faces also is clearly viewpoint dependent (Hill and Bruce 1996; Hill et al 1997; Troje and Bülthoff 1996). Other authors, in contrast, found viewpoint-invariant recognition in studies using both novel and familiar objects (Biederman and Gerhardstein 1993). Viewpoint dependence is often taken as evidence for viewer-centered as opposed to object-centered mental representations (Marr and Nishihara 1978; for a review, see also Ullman 1996). However, even for object-centered representations such as feature spaces or three-dimensional structural descriptions, one can expect to find some viewpoint dependence because particular views may provide more information about an object than others. This is in particular the case for viewpoint dependence measured in terms of view canonicality (Palmer et al 1981). Biederman and Gerhardstein (1993) used a priming paradigm to test for viewpoint dependence for familiar objects and found perfect viewpoint invariance. However, like viewpoint canonicality, this paradigm is not suitable for exploring the nature of long-term representations of familiar objects because it does not allow control over the information available to form the representation.

Bruce et al (1987) reported viewpoint dependence for familiar faces. Response times to profile views were slightly longer than to frontal views. The reason might be that the attention towards another person's face is triggered if this person is facing towards us, resulting in increased exposure to frontal views. If interpreted thus, the finding of Bruce et al would support a viewer-centered representation.

To explicitly test for the dependence of viewpoint on object representation, experiments have to be conducted that allow the experimenter to control which views of an object

¶ Author's current address to which all correspondence should be sent: Department of Psychology, Queen's University, Kingston, Ontario K7L 3N6, Canada; e-mail: niko@psyc.queensu.ca

an observer has seen prior to testing. For this reason, the question whether object representations are viewer centered or object centered seems to be approachable only with novel, unfamiliar objects. Unfamiliar objects, however, that are experienced for the first time in the course of the experiment may be stored and represented in a different manner than familiar objects that have been known for a long time. Recognizing unfamiliar, novel faces employs episodic memory, whereas recognizing familiar faces is a matter of semantic memory. Episodic memory is known to have very different properties from semantic memory (Tulving 1985), and it is therefore problematic to use results obtained from experiments with unfamiliar objects to draw conclusions about the representation of familiar objects.

To study the mental representation of familiar objects, one would like to be able to directly test the ability to generalize from familiar to novel views for familiar objects. The conundrum is how to find objects that have been identified many times in everyday life, but only from a restricted range of viewpoints. There are relatively few objects in the world that we are highly familiar with from certain views but not from others. One of them is our own face, which we experience from the daily glance into the mirror.

Typically, one has a disproportionately high exposure to near-frontal views of one's own face as seen in a mirror. The range of viewpoints is restricted by our oculomotor system (ca $\pm 40°$; Robinson 1981), and apart from relatively rare situations (eg photographs and mirror arrangements at the barbershop), we do not see our own face in profile view. Nevertheless, the range of possible viewpoints is large enough in principle to provide enough information to reconstruct the full three-dimensional structure of the head and thus all possible views (Koenderink and van Doorn 1991; Ullman and Basri 1991). If the visual system is able to use this information to establish an object-centered representation for familiar faces, the difference between the distribution of views seen from one's own face compared with the distribution of views seen from other familiar faces should not influence recognition performance. If there is a difference for recognizing frontal and profile views of familiar faces of other people, the same difference should be observed for recognizing one's own face. If there is no difference between recognition of frontal and profile views for familiar faces, this would also be expected for recognizing one's own face. A view-based representation, on the other hand, predicts poorer recognition performance for the profile view of one's own face.

2 Material and methods

2.1 *Stimuli*

We took color pictures of frontal and profile views of 26 members of our laboratory. The subjects were allowed to smile if they wanted to and they wore glasses if they did so normally. The pictures were all taken in front of a neutral grey wall. Images were digitized. The views ranged between 8 and 10 cm on a computer screen corresponding to a visual angle of about 6 deg at the viewing distance of 85 cm.

2.2 *Subjects*

The same 26 people that served as models for the pictures also participated as observers in the experiment. All of them had worked in the laboratory for more than three months, saw each other at least once a week in the regular laboratory meeting (but usually much more often) and knew each other well by name.

2.3 *Procedure*

We used a naming paradigm, measuring the time between stimulus onset and the beginning of the subject's response. Before the experiment the procedure was explained in detail to the subjects while leaving them naïve to the purpose of the experiment. We also prepared them not to be upset if they had a momentary memory block for a friend's name, an instance which happened to a few subjects once or twice. Before the experiment, the subjects were shown a list of all occurring names, including their own, and they were asked to recall the corresponding people one by one. In addition, they were explicitly informed that the series of images would include images of themselves. Finally, they were prepared to call out the first name of the person shown in the image as quickly as possible. After starting the experimental run, all 54 images were shown, each separated by a blank screen (1000 ms) and a fixation cross (750 ms). Each image remained on the screen until a microphone attached to the system registered an answer, and response time was measured. The order of the images was randomized individually for each subject according to the following constraints: (i) two successive images should not show the same face; (ii) the subject's own face should not appear within the first 10 trials; (iii) from the total of 26 subjects, 13 randomly chosen subjects saw first their frontal view and then their profile view, whereas the other 13 subjects saw first their profile view and then their frontal view.

2.4 *Data analysis*

Before the analysis we excluded all trials with response times longer than 2 standard deviations away from the mean response time calculated for each individual subject. A 2×2 ANOVA was run on the other data with factors coding for own versus other face and frontal versus profile view.

3 Results

We measured response times for correct naming of frontal and profile views of 26 familiar faces including the subject's own face. Figure 1 shows that response times to one's own face are faster than to other faces ($F_{1,25} = 26.1$, $p < 0.01$), possibly owing to

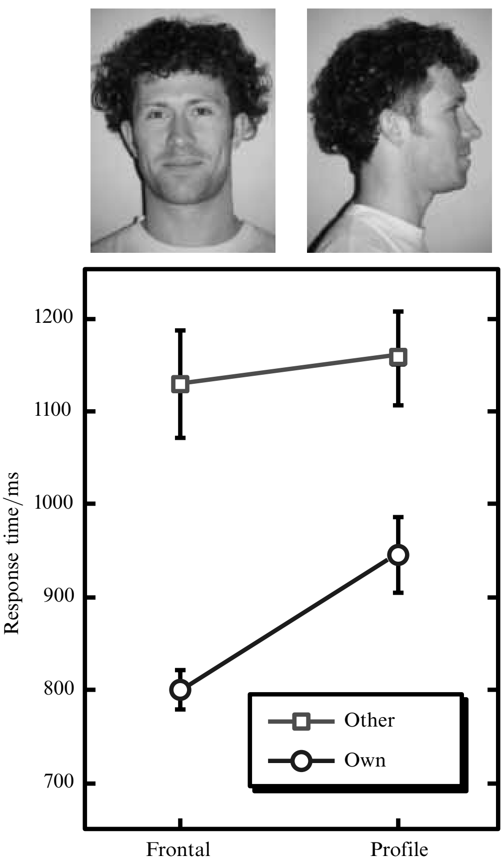

Figure 1. Mean response times for naming frontal and profile views of a subject's own face and for all the other faces.

greater familiarity with one's own name. There is also a main effect of face orientation ($F_{1,25} = 13.2$, $p < 0.01$). Response times to frontal views are faster than those to profile views. Most importantly, the results show an interaction between view and whether the image was that of the subject's own face or not ($F_{1,25} = 6.4$, $p < 0.05$). Subjects can name profile views just as fast as frontal views when dealing with the familiar faces of their colleagues ($t_{25} = 1.9$, $p > 0.05$), whereas they are significantly slower to recognize profiles than frontal views of themselves ($t_{25} = 3.3$, $p < 0.005$).

The procedure used to cope with the outliers excluded 53 data points. The outliers were mainly due to memory blocks; however, some of them were caused by not speaking loud enough to register a response. Not excluding the outliers yields longer response times to the faces of the other people but still no difference between frontal and profile views. Response errors were extremely infrequent (mean error rate: 0.7%) and were not analyzed.

4 Discussion

The pronounced interaction between the two factors of interest is caused by a significantly longer time for naming one's own profile view compared to one's own frontal view. This result indicates that we are not able to use the range of normally available views of our own face to construct a viewpoint-independent representation. More generally, the results suggest that evidence for viewer-centered representations that comes from several experiments with unfamiliar objects and faces can be generalized to highly familiar objects like one's own face.

Of course, the observed viewpoint dependence does not compellingly exclude the possibility of object-centered representations. Depending on the information available, an object-centered representation might be incomplete and thus would not provide the information needed to recognize the object from every possible view. However, considering an object-centered representation that is only based on the single views the system has been exposed to without making use of integrating them to achieve three-dimensional structural information, would blur the difference between the concepts of viewer-centered versus object-centered representations to an extent that questions their general usefulness.

The data show two other outcomes that we want to mention. First, we observe that the mean response time for other faces is significantly longer than the one for one's own face. The task that subjects had to solve in this experiment contains actually two conceptually different parts: First, the face has to be recognized and subsequently a name has to be assigned to it. The difference in mean response times between other faces and one's own face is most likely due to greater familiarity with one's own name. Sometimes, access to a name even of a good friend may be blocked, but never to one's own name. In fact, all the trials excluded as outliers were trials in which other faces were shown.

There is also the possibility that the observed bias is of representational nature. In particular, when considering view-based representations, the question arises whether the name of a person is linked equally to all of the stored views or whether it is predominantly linked to particular views. Naming someone explicitly usually involves social interaction and thus an increased probability for seeing the person in frontal view. Although the connection between naming and social interaction is not true for one's own face, a general association of a person's name with his or her frontal view is not completely implausible, and the faster naming of one's own face would then be based on the fact that this particular face is mainly represented in terms of its frontal view.

The second point we want to mention refers to the slight trend suggesting that even with other familiar faces there is a small advantage for frontal views. This trend is not significant in our data, but it is in accordance with findings by Bruce et al (1987) as discussed above, and it might reflect that even for other faces the distribution of views we are exposed to is not completely homogeneous.

There is still an ongoing discussion in the recent literature about the level on which differences between the processing of faces and the processing of other objects have to be described. We are aware that we have to be careful about extrapolating the conclusions of our results towards other object classes.

Acknowledgements. We thank Jeffrey Liter and Michael Tarr for useful comments on the manuscript and U Siebeck for taking the pictures. We acknowledge the participation of all members of the psychophysics subsection of the MPI for Biological Cybernetics in Tübingen in the experiment.

References

Biederman I, Gerhardstein P C, 1993 "Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance" *Journal of Experimental Psychology: Human Perception and Performance* **19** 1162 – 1182

Bruce V, Valentine T, Baddeley A D, 1987 "The basis of the 3/4 view advantage in face recognition" *Applied Cognitive Psychology* **1** 109 – 120

Bülthoff H H, Edelman S, 1992 "Psychophysical support for a two-dimensional view interpolation theory of object recognition" *Proceedings of the National Academy of Sciences of the USA* **89** 60 – 64

Hill H, Bruce V, 1996 "Effects of lighting on the perception of facial surfaces" *Journal of Experimental Psychology: Human Perception and Performance* **22** 986 – 1004

Hill H, Schyns P, Shigeru A, 1997 "Information and viewpoint dependence in face recognition" *Cognition* **62** 201 – 222

Koenderink J J, Doorn A van, 1991 "Affine structure from motion" *Journal of the Optical Society of America A* **8** 377 – 385

Marr D, Nishihara H K, 1978 "Representation and recognition of the spatial organization of three dimensional shapes" *Proceedings of the Royal Society of London, Series B* **200** 269 – 294

Palmer S E, Rosch E, Chase P, 1981 "Canonical perspective and the perception of objects", in *Attention and Performance IX* Eds J Long, A Baddeley (Hillsdale, NJ: Lawrence Erlbaum Associates) pp 135 – 151

Robinson D A, 1981 "Control of eye movements", in *Handbook of Physiology* volume 2, part 2, Ed.V B Brooks (Baltimore, MD: Williams and Wilkins) pp 1275 – 1320

Rock I, 1973 *Orientation and Form* (New York: Academic Press)

Tarr M J, 1995 "Rotating objects to recognize them: A case study on the role of viewpoint dependency in the recognition of three-dimensional objects" *Psychonomic Bulletin & Review* **2** 55 – 82

Troje N F, Bülthoff H H, 1996 "Face recognition under varying poses: the role of texture and shape" *Vision Research* **36** 1761 – 1771

Tulving E, 1985 *Elements of Episodic Memory* (Oxford: Clarendon Press)

Ullman S, 1996 *High Level Vision* (Cambridge, MA: MIT Press)

Ullman S, Basri R, 1991 "Recognition by linear combinations of models" *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13** 992 – 1005

© 1999 a Pion publication