Object recognition can be viewpoint dependent or invariant – it's just a matter of time and task

Branka Milivojevic[1,2]*

[1] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, Netherlands
[2] Experimental and Developmental Psychology, Utrecht University, Utrecht, Netherlands
*Correspondence: branka.mili@gmail.com

As we move through our environment, we encounter familiar objects from various viewpoints. Despite the ensuing variability of the images projected onto the retina, we have seemingly little difficulty when it comes to recognizing objects we encounter. We can, however, see how the objects are oriented, suggesting that object recognition is to a certain degree dissociable from perception of other object "features" such as orientation. Changes in orientation of objects, particularly inversion, can also affect how we perceive the objects. A particularly illustrative example (shown in **Figure 1**) is that of the Thatcher illusion (Thompson, 1980), where the grotesque appearance of a face with its inverted eyes and mouth is "hidden" when the whole face is also inverted. The percept itself, therefore, is affected by the change in orientation. In addition, there are also subtle effects of viewpoint changes on object recognition itself. For example, identifying rotated objects is more difficult when they are briefly presented than when viewing time is unlimited (Lawson and Jolicoeur, 2003), and identifying a face is considerably more difficult the face has been inverted (Yin, 1969), as is discrimination between characters "b" and "d," or "p" and "q" which requires (physical or mental) rotation of the characters to upright, before we can be certain which letter we are looking at (Corballis and McLaren, 1984).

These subtle, yet persistent, effects of viewpoint changes on perception and recognition arise as a consequence of how visual object processing is handled by the brain. Here, I discuss how neural mechanisms underlying visual processing give rise to perception and recognition which can be both viewpoint dependent and viewpoint invariant depending on the timing of those processes, as well as specific task demands or current "perceptual goals" of an individual. To do so, I will firstly explain how temporal dynamics of low-level visual processing may give rise to impaired recognition at short viewing latencies and suggest that this may also relate to effects of viewpoint changes on perceptual experience. I will then discuss how the perceptual goals of an individual determines whether recognition is accomplished in viewpoint invariant or dependent manner with a particular focus on cognitive operations thought to be subserved by ventral and dorsal visual streams, namely object recognition and mental rotation, respectively.

PERCEPTION IS AFFECTED BY POINT OF VIEW

Change in orientation must affect processing of visual information. For example, as our viewpoint changes, so does the shape of the image that falls on the retina. In the case of picture-plane rotations, the orientation of the edges of that shape will also change and thus stimulate different populations of orientation-tuned visually responsive neurons in primary visual cortex. However, these initial effects of orientation-changes on neural processing probably do not give rise to altered perceptual experience such as those associated with inversion of a Thatcherized face.

Inversion affects how we perceive the spatial relations between objects' features and may, as James (1890) suggested, depend on perceptual experience with an object at a given orientation. This could explain why recognition of faces is particularly impaired by inversion: faces are most frequently seen the right way up, and are thought to be recognized using information about the configuration of the constituent features. As mirror reversal is also a special case of a configural change where the relative configuration of object's features remains the same but reverses in its left–right orientation, this could also explain why mirror–images are difficult to tell apart when they are rotated away from a canonical viewpoint, and which is why we must rotate objects into alignment with our egocentric reference frames before we can distinguish between parity-defined characters such as "b" and "d" (Corballis and McLaren, 1984). Interestingly, neural responses to unaltered and thatcherized images also follow the perceptual illusion and disappear as the face is rotated away from upright (Milivojevic et al., 2003a).

On neural level, large changes in the viewpoint of an object, such as inversion of faces (Rossion et al., 2000) and alphanumeric characters (Milivojevic et al., 2008), result in delays of the N170 component. The N170 is thought to reflect object classification, and inversion-related delays of N170 possibly reflect increases in time required to accumulate sufficient neural activity to reach a threshold at which recognition can occur (Perrett et al., 1998; Heekeren et al., 2008). If changes in viewpoint delay visual object encoding, this could explain why accurate recognition of rotated objects requires longer viewing times than recognition of canonically oriented objects (Jolicoeur and Landau, 1984; Lawson and Jolicoeur, 2003; Mack and Palmeri, 2011).

VIEWPOINT MATTERS ONLY FOR SOME PERCEPTUAL GOALS

Task-dependent effect of viewpoint changes on neural processing are only observed around 250 ms after stimulus onset and coincide with the P2 component of the ERP. For example, if the observers need to determine whether a rotated alphanumeric character is normal or mirror-reversed, they will mentally rotate it to upright before making the decision. Although the beginning of mental rotation is later than the P2, parity decisions are associated with linear increases of P2 amplitudes while this is not the case for categorization of alphanumeric characters which does not require mental rotation (Milivojevic et al., 2011). Interestingly, similar increases in P2 amplitudes can be observed as a consequence of stimulus degradation, either by addition of noise (Banko et al., 2011) or by occlusion

FIGURE 1 | Unaltered and "thatcherized" version of Margaret Thatcher's face. The grotesque appearance of the face when its eyes and mouth are inverted is hidden by the inversion of the whole image. Rotating the pictures to upright makes discrimination between the two versions of the face easier.

(Doniger et al., 2000), but not size transformation (Muthukumaraswamy et al., 2003), suggesting that changes in orientation degrade certain types of perceptual information which may be required for task-specific decision making, and may be, thus, associated with some form of perceptual decision making (Heekeren et al., 2008; Schendan and Lucia, 2009, 2010), such as whether sufficient information is available for the perceptual goal to be achieved. This decision would then trigger other visuospatial cognitive operations, such as mental rotation or more detailed inspection of individual features of an object. Those cognitive operations would lead to acquisition of additional information about the object which would, in turn, enable a more accurate completion of the perceptual task at hand. For the purpose of illustration, two types of "perceptual goals" that depend on object orientation will be described: object identification and parity-based recognition.

IDENTIFICATION IS VIEWPOINT DEPENDENT BUT CATEGORISATION IS NOT

As already mentioned, face recognition is worse when faces are inverted (Yin, 1969), both in terms of reduced recognition accuracy and increased reaction times. This seems to be the case both for familiar and unfamiliar faces, and may be a consequence of disrupted neural processing underlying object classification although a causal relationship has not been firmly established. It should be noted here that faces are nevertheless recog-

nized as faces, what seems to be disrupted is the *identification* of the face as belonging to a particular person or *identification* of an emotional expression, while differentiation between categories of "face" and "non-face" objects is largely unimpaired by inversion.

The difference in viewpoint-sensitivity of identification and categorization has also been established for other classes of objects. For example, identifying letters of the alphabet is affected by character orientation while the same is not the case for between-category decisions such as letter–digit categorization (Corballis et al., 1978). In a sense, categorization may relate to recognition at a basic or entry level described by Roch (Rosch et al., 1976), while identification may be more closely related subordinate-level recognition. Object recognition at basic level (e.g., deciding a shape is a dog) are not affected by changes in viewpoint, while subordinate-level decisions (e.g., identifying a dog as a poodle) are affected by viewpoint changes in terms of reaction times and accuracy (Hamm and McMullen, 1998).

Studies which have directly compared identification and categorization of objects using neuroimaging methods are scarce. Nevertheless, studies investigating neural correlates of rotated-object categorization show little evidence of orientation-dependence at visual processing stages beyond the initial encoding of the objects (see above). In contrast, studies investigating rotated-object recognition either as identity-matching or in terms of explicit identification show that there is

an increase in activity in areas involved in object recognition within the inferior temporal cortex for various object classes such as faces (Haxby et al., 1999), bodies (Brandman and Yovel, 2010), landscapes (Epstein et al., 2006). Some authors have suggested that this increase in activity may reflect a shift in recognition strategy from one that is based on the whole shape to one that is based on the analysis of individual object features (i.e., details Jolicoeur, 1990).

RECOGNIZING PARITY-DEFINED SHAPES REQUIRES MENTAL ROTATION

Decisions regarding the direction of the left–right axis of an object, or its *handedness*, require alignment between the object and our own egocentric frame of reference. For example, deciding whether a shoe is the left or the right one requires either physical or *mental* rotation of the shoe into alignment with our feet, or the feet with the shoe. The same holds for any object class that has a well-defined left–right orientation, such as alphanumeric characters, which can be readily recognized as "backward" if they have been mirror-reversed (Cooper and Shepard, 1973) – but only if they are presented at upright. Rotated characters require rotation to their canonical upright before we can notice if they are normal or backward, particularly if they are rotated by a large degree (Kung and Hamm, 2010). When the identity of an object depends on its left–right parity, as is the case with lower-case letters "b" and "d" or "p" and "q," then the discrimination of such characters also requires rotation to upright before it can be successfully recognized (Corballis and McLaren, 1984).

This suggests that information regarding the identity of the object must be extracted before information about the handedness of an object can be determined. Although generally we need to recognize an object before mental rotation begins (Heil et al., 1996; Schendan and Lucia, 2009), this cannot be the case for objects whose identity depends on their handedness, such as "b" and "d" or "p" and "q." With the exception of alphanumeric characters, there are not many commonly encountered objects whose identity is defined by parity (i.e., a hand is a hand irrespective of whether it is a left one or a right one) and those objects can be seen as special case whose identity cannot be determined at all orientations. For these objects, identification from a feature-based

descriptor such as "a semi-circle attached at an end of a long stem" could lead to selection of possible four candidates, and the remaining possibilities would need to be resolved with mental rotation.

Mental rotation has been associated with linear increases in centro-parietal negativity between ~400 and 800 ms after stimulus onset (e.g., Milivojevic et al., 2009b) which last somewhat longer for larger angular departures from upright (Milivojevic et al., 2003b; Hamm et al., 2004). The ERP correlates of mental rotation are probably generated by a distributed network of sources localized (Milivojevic et al., 2009b) within a network of prefrontal and posterior parietal areas which has been identified using fMRI (e.g., Milivojevic et al., 2009a). Whether these areas also subserve recognition of rotated parity-defined objects is still unclear as this particular question has not been investigated using neuroimaging.

SUMMARY AND CONCLUSION

Although changes in viewpoint rarely interfere with common perceptual goals, such as categorizing objects into basic categories, this type of viewpoint invariant recognition can only be achieved after initial viewpoint-dependent neural processing has been accomplished. Depending on current perceptual goals, changes in viewpoint may impose certain recognition costs, observable in terms of increased response latencies or reduced accuracy. These costs are likely to reflect increased cognitive demands associated with recognition of misoriented shapes such as detailed analysis of object features or mental rotation of the shape to its canonical upright. In this sense, recognition of objects will always be affected by changes in viewpoint early on in the visual processing stream, but these effects will taper off with time. At later visual processing stages, some types of perceptual goals such as object identification or parity discrimination, will require additional processing operations which will give rise to viewpoint dependent behavioral performance.

ACKNOWLEDGMENTS

REFERENCES

Banko, E. M., Gal, V., Kortvelyes, J., Kovacs, G., and Vidnyanszky, Z. (2011). Dissociating the effect of noise on sensory processing and overall decision difficulty. *J. Neurosci.* 31, 2663–2674.

Brandman, T., and Yovel, G. (2010). The body-inversion effect is mediated by face-selective, not body-selective, mechanims. *J. Neurosci.* 30, 10534–10540.

Cooper, L. A., and Shepard, R. N. (1973). "Chronometric studies of the rotation of mental images," in *Visual Information Processing*, ed. W. G. Chase (New York: Academic Press), 75–176.

Corballis, M. C., and McLaren, R. (1984). Winding one's ps and qs: mental rotation and mirror-image discrimination. *J. Exp. Psychol. Hum. Percept. Perform.* 10, 318–327.

Corballis, M. C., Zbrodoff, N. J., Shetzer, L. I., and Butler, P. B. (1978). Decisions about identity and orientation of rotated letters and digits. *Mem. Cognit.* 6, 98–107.

Doniger, G. M., Foxe, J. J., Murray, M. M., Higgins, B. A., Snodgrass, J. G., Schroeder, C. E., and Javitt, D. C. (2000). Activation timecourse of ventral visual stream object-recognition areas: high density electrical mapping of perceptual closure processes. *J. Cogn. Neurosci.* 12, 615–621.

Epstein, R. A., Higgins, J. S., Parker, W., Aguirre, G. K., and Cooperman, S. (2006). Cortical correlates of face and scene inversion: a comparison. *Neuropsychologia* 44, 1145–1158.

Hamm, J. P., Johnson, B. W., and Corballis, M. C. (2004). One good turn deserves another: an event-related brain potential study of rotated mirror-normal letter discriminations. *Neuropsychologia* 42, 810–820.

Hamm, J. P., and McMullen, P. A. (1998). Effects of orientation on the identification of rotated objects depend on the level of identity. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 413–426.

Haxby, J. V., Ungerleider, L. G., Clark, V. P., Schouten, J. L., Hoffman, E. A., and Martin, A. (1999). The effect of face inversion on activity in human neural systems for face and object perception. *Neuron* 22, 189–199.

Heekeren, H. R., Marrett, S., and Ungerleider, L. G. (2008). The neural systems that mediate human perceptual decision making. *Nat. Rev. Neurosci.* 9, 467–479.

Heil, M., Bajric, J., Rösler, F., and Hennighausen, E. (1996). Event-related potentials during mental rotation: disentangling the contributions of character classification and image transformation. *J. Psychophysiol.* 10, 326–335.

James, W. (1890). *Principles of Psychology*. London: Macmillan.

Jolicoeur, P. (1990). Identification and disoriented objects: a dual systems theory. *Mind Lang.* 5, 387–410.

Jolicoeur, P., and Landau, M. J. (1984). Effects of orientation on the identification of simple visual patterns. *Can. J. Psychol.* 38, 80–93.

Kung, E., and Hamm, J. P. (2010). A model of rotated mirror/normal letter discriminations. *Mem. Cognit.* 38, 206–220.

Lawson, R., and Jolicoeur, P. (2003). Recognition thresholds for plane-rotated pictures of familiar objects. *Acta Psychol. (Amst.)* 112, 17–41.

Mack, M. L., and Palmeri, T. J. (2011). The timing of visual object categorization. *Front. Psychol.* 2:165. doi: 10.3389/fpsyg.2011.00165

Milivojevic, B., Corballis, M. C., and Hamm, J. P. (2008). Orientation sensitivity of the N1 evoked by letters and digits. *J. Vis.* 8(10), 1–14. doi: 10.1167/1168.1110.1111

Milivojevic, B., Hamm, J. P., and Corballis, M. C. (2009a). Functional neuroanatomy of mental rotation. *J. Cogn. Neurosci.* 21, 945–959.

Milivojevic, B., Hamm, J. P., and Corballis, M. C. (2009b). Hemispheric dominance for mental rotation: it is a metter of time. *Neuroreport* 20, 1507–1512.

Milivojevic, B., Hamm, J. P., and Corballis, M. C. (2011). About turn: how object orientation affects categorisation and mental rotation. *Neuropsychologia* 49, 3758–3767.

Milivojevic, B., Clapp, W. C., Johnson, B. W., and Corballis, M. C. (2003a). Turn that frown upside down: ERP effects of thatcherization of misoriented faces. *Psychophysiology* 40, 967–978.

Milivojevic, B., Johnson, B. W., Hamm, J. P., and Corballis, M. C. (2003b). Non-identical neural mechanisms for two types of mental transformation: event-related potentials during mental rotation and mental paper folding. *Neuropsychologia* 41, 1345–4356.

Muthukumaraswamy, S. D., Johnson, B. W., and Hamm, J. P. (2003). A high-density ERP comparison of mental rotation and mental size transformation. *Brain Cogn.* 52, 271–280.

Perrett, D. I., Oram, M. W., and Ashbridge, E. (1998). Evidence accumulation in cell populations responsive to faces: an account of generalisation of recognition without mental transformations. *Cognition* 67, 111–145.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., and Boyes-Braem, E. (1976). Basic objects in natural categories. *Cogn. Psychol.* 8, 382–439.

Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., and Crommelinck, M. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: an electrophysiological account of face-specific processes in the human brain. *Neuroreport* 11, 69–74.

Schendan, H. E., and Lucia, L. C. (2009). Visual object cognition precedes but also temporally overlaps mental rotation. *Brain Res.* 1294, 91–105.

Schendan, H. E., and Lucia, L. C. (2010). Object-sensitive activity reflects earlier perceptual and later cognitive processing of visual objects between 95 and 500ms. *Brain Res.* 1329, 124–141.

Thompson, P. (1980). Margaret Thatcher: a new illusion. *Perception* 9, 483–484.

Yin, R. K. (1969). Looking at upside down faces. *J. Exp. Psychol.* 81, 141–145.