
A sideways look at configural encoding: Two different effects of face rotation

Michael B Lewis, Thomas E Glenister

School of Psychology, Cardiff University, PO Box 901, Cardiff CF10 3YG, UK;

e-mail: LewisMB@Cardiff.ac.uk

Received 3 July 2002, in revised form 29 October 2002; published online 12 December 2002

Abstract. Inversion has a disproportionate disruptive effect on the recognition of faces. This may be due to the disruption of holistic or configural encoding employed to recognise upright faces. The paradigm developed by Tanaka and Farah (1993 *Quarterly Journal of Experimental Psychology, Section A* **46** 225–246) was used to investigate the effect of 90° (or orthogonal) rotation on configural encoding. Faces learnt in the orthogonal condition were not recognised as well as upright faces, but a whole-face advantage was found in both cases. This whole-face advantage did not occur for inverted faces. It appears that 90° rotation affects recognition but not specifically configural encoding. It is concluded that rotating a face can have at least two different effects on face processing depending on the range of rotation. Implications for the nature of facial dimensions and the expertise account of the inversion effect are considered.

1 Introduction

A great deal of research has focused on the perception of inverted faces compared with upright faces. Yin (1969) first showed that inversion disproportionately disrupts the recognition of faces compared to houses, stick men, or aeroplanes. Further research into the effects of inversion revealed a wealth of findings (see Valentine 1988, for a review). The observed inversion effects can be used as evidence that the perception of faces is, in some way, a special perceptual process and is different from the perception of other objects. Alternatively, it has been suggested that the inversion effect is a result of specific expertise with a set of homogeneous stimuli. Evidence for an expertise explanation comes from the fact that a similar inversion effect has been observed for dog experts when recognising species of dogs (Diamond and Carey 1986). Further, Gauthier and Tarr (2002) have shown similar inversion effects and a whole-object advantage using novel objects (Greebles) learnt over several hours of training. Whether or not face recognition is special remains a debated issue (see Maurer et al 2002, for a review).

Tanaka and Farah (1993) offered one explanation of this disproportionate effect of inversion on face recognition.⁽¹⁾ They suggested that faces are perceived by two processes: holistic (or configural) encoding and featural (or rational) encoding. Tanaka and Farah demonstrated that component features of a face are recognised better when participants are presented with whole faces rather than isolated features. These effects occurred even though the two faces offered in a forced-choice task were identical except for the one feature. If, however, the face is either scrambled or inverted during learning, then this whole-face advantage disappears. In their experiment, inversion only affected recognition of the whole face, and so it was postulated that it was holistic information that was disrupted by this transformation.

The differences between the two encoding methods (configural/holistic and featural) are not clear-cut, and opinions differ as to what is meant by each. For example, the

⁽¹⁾Rakover and Teucher (1997) have suggested an alternative explanation for the face-inversion effect based on feature saliency, upright orientation, and mental rotation. Leder and Bruce (2000), however, have suggested that this explanation might not be so different from others if configural information is interpreted as relations between features.

featural and holistic encoding used by Tanaka and Farah (1993) may be different from the first-order and second-order features suggested by Rhodes et al (1993) or the component and spatial-relation information proposed by Murray et al (2000). In spite of these nominal differences, there is general agreement in the partitioning of encoding strategies into featural (first-order, component, or rational) and configural (holistic, second-order, or spatial-relationship) types. In general, featural information consists of the presence of a particular feature or type of feature (eg a mole or a big nose), whereas configural information is derived from the arrangement of features relative to each other within a face. This distinction, however, is not simple. The presence or absence of a big nose can be considered to be a featural element, but if one considers the nose to be big only in relation to a small face, then it is a configural feature. This interdependence was demonstrated by Tanaka and Sengco (1997) who showed that the recognisability of a person by a feature was affected by the configuration in which it occurred.

While most authors agree that there are two types of encoding, the range of terms employed means that a definition of the visual properties of a face remains elusive. It appears that the best description of the two types of encoding is based on a working definition (see Lewis 2001). One can describe configural encoding as that processing of a face which is disrupted by inversion. Configural encoding, therefore, is not necessarily restricted to holistic encoding because, as we will see later, it is possible to have configural encoding of an isolated feature. Featural encoding can be described, therefore, as that processing which can occur for inverted faces. This definition does not describe what these encoding strategies look like, but it does offer an objective method for determining whether a change is featural or configural.

The configural nature of facial encoding implies that faces are recognised by default as whole objects and so not as a sum of the recognition of discrete featural parts. This explanation of the face-inversion effect has subsequently been supported by a range of evidence. Leder and Bruce (2000), for example, investigated the ease with which featural or configural changes could be detected in either upright or inverted faces. In these experiments, inversion effects were larger for configural changes (eg relative feature positions) than for featural changes (eg hair colour).

While much research has explored inversion (or 180° rotation), little attention has been given to intermediate levels of rotation. What discussion of rotation there has been has often focused on one specific aspect of the face-inversion effect, that being the Thatcher illusion (Thompson 1980). To generate a Thatcher illusion, take a face and invert its mouth and eyes. The image looks grotesque because configural changes have been made. If the whole image is then inverted, this configural information is disrupted and the image does not look abnormal.

The Thatcher illusion can be used as a benchmark for whether configural processing is taking place (see Lewis and Johnston 1997). The longer it takes to detect a Thatcherisation, the less configural encoding is being employed. Sjöberg and Windes (1992) measured the time to detect a Thatcherisation at steps of 60° rotation from upright to inverted. They found a steady increase in reaction times as the angle increased, although the greatest increase occurred between 60° and 120°. Stürzel and Spillmann (2000) took these results to suggest that there is a sudden change from configural to non-configural processing of faces. Their further research, with Thatcherised faces on a rotating card, localised this transition to between 90° and 120° from upright (a region in which subjects reported faces going from normal looking to grotesque or the reverse).

Murray et al (2000) used 'bizarreness' ratings of Thatcherised and teeth-blackened faces to investigate face rotation. They report a performance discontinuity in the 90°–120° rotation region that is present for the former but not the latter stimuli. Their data, however, suggest that the change in bizarreness is continuous and nonlinear

rather than being linear with a discontinuity [the latter being predicted by Stürzel and Spillmann (2000)]. Lewis (2001) also found such a continuous nonlinear change during rotation. In that experiment, subjects had to detect Thatcherisation of faces rotated at 10° intervals. Lewis suggested that there might be no sharp change in the configural processing as a face is rotated, but instead a steady decrease in the amount or utility of configural information for increasingly rotated faces. Further, Lewis suggested that facial processing may use a range or continuum of features ranging from very configural to very featural, rather than the dichotomy that has been described previously. Some features are disrupted by a small degree of rotation whereas other features are disrupted only by a large degree of rotation.

If this idea of a continuum of features were true, then it would be possible for different parts of the rotation, from upright to inverted, to have different effects on face processing. A similar prediction can be made from the feature-dichotomy account, with rotation before the switch in processing style being different from rotation including the switch. This issue was addressed in the current study by exploring the effect on configural encoding of 90° rotation and 180° rotation. The method employed was similar to that used by Tanaka and Farah (1993), which demonstrated clearly the effects of inversion on configural encoding.

2 Experiment

The purpose of this experiment was to replicate and extend Tanaka and Farah's (1993) original study. In their study, subjects learnt to recognise a set of six faces that differed only in their eyes, nose, or mouth. Subjects were then asked to identify which of two faces belonged to a particular individual or which of two features were part of a particular individual's face. An advantage was found for performance in the first of these conditions, but this whole-face advantage disappeared when the faces were inverted during learning and testing. For upright faces, therefore, the holistic configuration aids recognition, but inversion disrupts configural encoding and so this advantage is not apparent for inverted faces.

The stimuli used in Tanaka and Farah's (1993) experiment were Mac-a-Mug faces. These are black-and-white grainy images. Although this produces images that are recognisable as faces, they contain little textural information and the final quality of the images is not comparable to that of photographic images commonly used in experiments. Tanaka and Farah's experiment, however, has been replicated with better-quality facial images (see Palermo and Rhodes 2002, or Tanaka and Gauthier 1997). In the experiment reported here, we attempted to replicate their result using greyscale faces generated from a modern computer-based face-reconstruction package. This package was used to produce realistic faces in which individual component features could be varied whilst keeping the rest of the face constant.

The second difference from the original study was the inclusion of a 90° rotation (or orthogonal) condition. Comparisons between performance for upright and inverted conditions with this orthogonal condition would reveal how configural encoding changes during rotation.

2.1 Method

2.1.1 *Participants.* Twenty-one undergraduates from the Cardiff University received course credit for their participation in this experiment.

2.1.2 *Materials.* Three Caucasian greyscale prototype faces were generated with the computer-based face-reconstruction system Faces 3.0 (published by IQ Biometrix, Inc.). These were constructed such that each specified feature (head shape, hairstyle, etc) was different. This package produces good-quality greyscale faces (see figure 1). From these prototypes, three sets of different faces were generated by changing just the eyes,

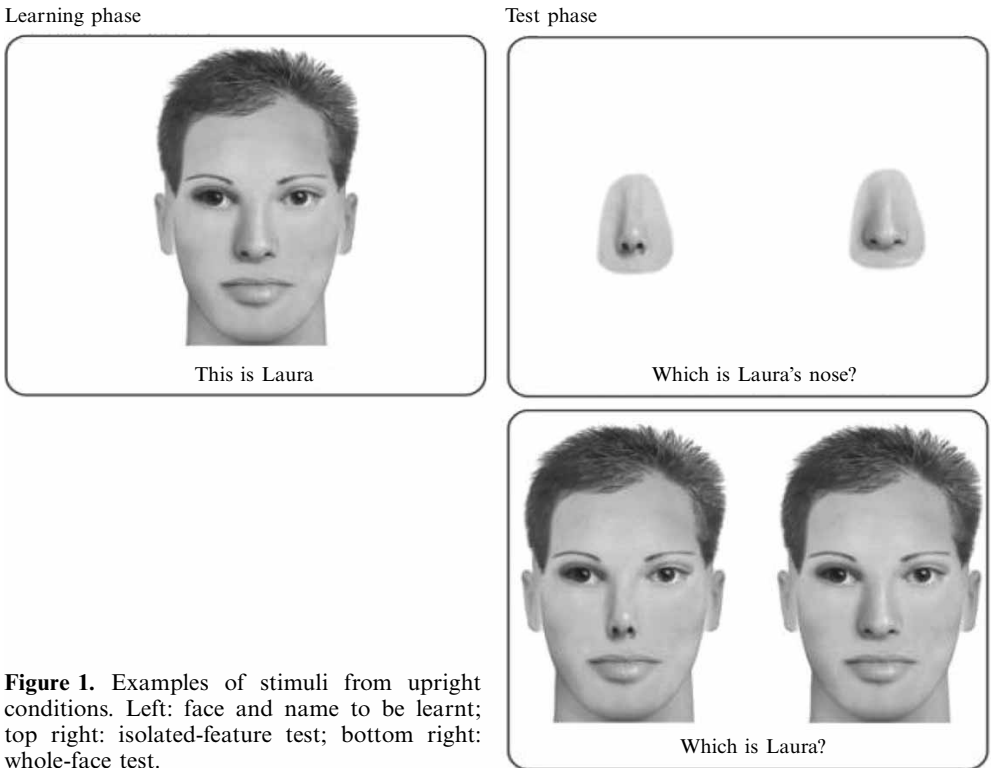


Figure 1. Examples of stimuli from upright conditions. Left: face and name to be learnt; top right: isolated-feature test; bottom right: whole-face test.

nose, and mouth. The same internal features were never used across prototype sets. Each prototype set had three types of eyes, three types of noses, and three types of mouths which could be interchanged to generate any of twenty-seven unique faces.

Six faces from each prototype set were chosen as the learning set. These were chosen so that none of them had more than one feature in common with any of the other learning faces. A set of full-face distractors was generated, such that each differed from the learning faces by only one feature. Images for the isolated-feature tests were generated by reproducing just a part of the face (eyes, nose, or mouth). All images were presented on a Macintosh computer.

Eighteen common names were generated to be associated with the faces.

2.1.3 Procedure. The experiment was conducted in three parts: one for each of the three prototype sets. Each part was made up of a learning phase and a test phase. The parts were arranged so that each participant did one with all faces upright, one with all faces at 90° rotation (ie orthogonal), and one with all faces inverted. The prototype set used in each condition was counterbalanced between participants.

During the learning phase, participants were required to associate six names with the six faces from one of the prototype sets. In each learning trial, a name would appear with its associated face for 5 s. A learning block consisted of one trial with each of the six faces in a random order and there were five blocks in total. All faces were presented either upright, orthogonal, or inverted, depending on the part of the experiment.

The test phase followed after the learning phase with only a short pause to remind the participants of the task. The test phase consisted of a series of two-choice recognition tests. In each test trial, the participant was presented with a name and two faces or two isolated features (eyes, noses, or mouths). Participants had been informed that

identification of the person by a single feature would be necessary in some cases. One of the test stimuli was the correct face, or the correct isolated feature, for the given name. The distractor item in the isolated-feature test condition was a similar feature from one of the other faces in the prototype set. For the whole-face tests, the distractor face was a face that differed by only one feature from the target face. This one feature was substituted by a different feature also used in the prototype set. Figure 1 illustrates the images used for the learning and test phases. Participants were required to decide which face or feature was associated with the given name. The test phase was conducted in the same orientation as the learning phase, so, if faces had been learnt upside-down, they were tested upside-down. All features from all faces were tested twice: once isolated and once in a whole face, making 36 tests for each of the three parts.

2.1.4 Design. There were two within-participant independent variables: orientation (upright, orthogonal, or inverted); and type of test stimulus (whole face or isolated feature). The dependent measure was the response (correct or incorrect) to the forced-choice decision which face or feature belonged to a name. The order of presentation of upright, orthogonal, and inverted sets was counterbalanced across participants. The learning trials and test trials were randomised within each block.

2.2 Results

The proportion of correct identifications of a face or feature was calculated for each participant in each of the six conditions (three orientations by either whole face or isolated feature). The means of these proportions are shown in figure 2. Each condition was significantly greater than chance performance ($\mu_i = 50\%$; $ps < 0.05$) and significantly less than ceiling ($\mu_i = 100\%$; $ps < 0.05$). Performance over the three orientations (collapsed across stimulus type) showed a gradual decrease as the faces were rotated: upright faces = 71%; orthogonal faces = 66%; and inverted faces = 65%. Overall, whole faces (average = 75%) at test gave better accuracies than isolated features (average = 64%).

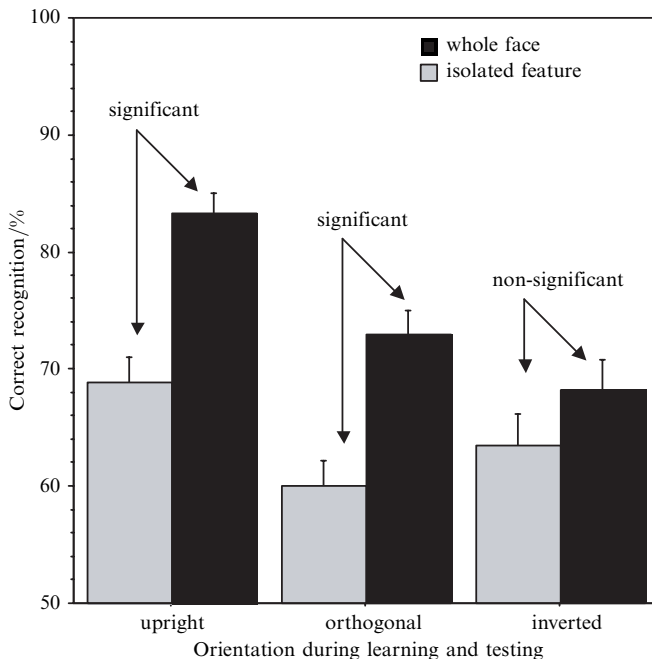


Figure 2. Percentage of correct recognition for the six experimental conditions. Error bars show standard errors.

A two-way ANOVA was conducted on the proportions of correct choices. The two factors were orientation of the faces during learning and test (upright, orthogonal, or inverted) and the type of test stimuli (whole faces or isolated features). This analysis found significant effects of orientation ($F_{2,40} = 12.519$, $p < 0.001$) and test stimuli ($F_{1,20} = 48.152$, $p < 0.0001$) as well as a significant interaction ($F_{2,40} = 4.574$, $p < 0.05$). Newman–Keuls a posteriori tests on the orientation factor revealed that upright-learned faces produced significantly better performance than either orthogonally learned or inverted-learned faces. The simple main effects were analysed because the interaction was significant. This analysis showed a significant effect of orientation for both whole-face tests ($F_{2,40} = 14.292$, $p < 0.001$) and for isolated-feature tests ($F_{2,40} = 4.925$, $p < 0.05$). Further, there was a significant whole-face advantage when the faces were learned and tested either upright ($F_{1,20} = 57.996$, $p < 0.001$) or orthogonally ($F_{1,20} = 23.690$, $p < 0.001$). This effect was not significant for inverted conditions ($F_{1,20} = 2.749$, $p > 0.05$). A Newman–Keuls a posteriori test was conducted on the interaction to investigate how the size of the whole-face advantage changed over different orientations. A significant difference was found in the size of whole-face advantage between the inverted-face condition and the other two conditions (ie orthogonal and upright).

An alternative way to explore the interaction is to break the analysis down further. Two separate 2×2 ANOVAs were conducted. The first explored the interaction limited to the upright and orthogonal conditions. This was found to be non-significant ($F_{1,20} = 0.178$, $p > 0.05$). The second explored the interaction limited to the orthogonal and inverted conditions. This interaction was found to be significant ($F_{1,20} = 5.536$, $p < 0.05$). The interaction, therefore, was only present in the orthogonal-to-inverted transition and was such that there was a significantly greater whole-face advantage for orthogonal faces than for inverted faces. This result is less conservative than the Newman–Keuls analysis described above but it yields the same results in this instance.

3 Discussion

The results of the experiment demonstrate that upright faces are learned (or retrieved) better than inverted faces. This result has already been demonstrated many times before (eg Yin 1969). The pattern of results obtained here is similar to those obtained in Tanaka and Farah's (1993) experiment from which this experiment was derived. Recognition was most accurate when faces had been learned in the upright orientation and were tested as whole faces rather than isolated parts. The use of just isolated features in the test, in the upright condition, has a significant detrimental effect on recognition. This suggests that configural information is being used to judge between the two test images.

Tanaka and Farah (1993) found a significant interaction between the effects of orientation and test type. A similar interaction is found here, with the effect of isolated-feature versus whole-face conditions being smaller for inverted faces than upright faces. Although the pattern of results obtained here for upright and inverted faces is similar to the one found by Tanaka and Farah, there are a couple of differences. Tanaka and Farah found that isolated-feature tests for upright faces produced a level of performance equivalent to that of inverted faces. In our results, performance on isolated-feature tests with upright faces was significantly better than that with inverted faces. This result is consistent with the findings of Palermo and Rhodes (2002) who also employed greyscale faces. They also found better performance for upright isolated features than for inverted isolated features (although in their case they were not compared statistically). This demonstrates that inversion disrupts the coding of isolated features as well as whole faces. To understand this isolated-feature-inversion effect, not found by Tanaka and Farah, it is useful to consider the stimuli used. Tanaka and Farah used features

generated by Mac-a-Mug which, while producing reasonable representations, does not provide much fine detail at the feature level. The images used here were almost photographic-quality greyscale faces. Isolated features from these faces are highly detailed and it is quite possible that a configural encoding of the individual features takes place as well as configural (or holistic) encoding of the entire face. Inverting the isolated features disrupts this configural encoding (within the isolated feature), thus producing poorer performance in recognition than when upright isolated features are used. The inversion effect, therefore, is not restricted to the disruption of holistic processing.

One problem with the interpretation of configural encoding of isolated features is that there are some studies that have failed to show an advantage for upright isolated features over inverted isolated features. Tanaka and Farah (1993) is one such study, but this finding can be attributed to the poor quality of the facial representations employed. Rhodes et al (1993) also used grainy (two-tone) faces and they also failed to show an advantage for upright isolated features. It appears, therefore, that configural encoding of smaller features only takes place when higher-quality images are employed.

The current study extends previous studies in that 90° rotation was also employed. If rotation has a gradual and constant effect on recognition, then one might expect that performance from the orthogonal condition would lie between that of the upright and inverted conditions. For whole-face targets, this was the case, but recognition of isolated features was poorer for orthogonal faces than for either upright or inverted faces (although orthogonal and inverted conditions were not significantly different from each other). The change in accuracy scores between upright and orthogonal conditions is similar for both whole faces and isolated features. This suggests that while the first 90° of rotation affects recognition, it does so equally to whole faces and isolated features. This rotation does not differentially affect configural and featural processing.

The transition from the orthogonal condition to the inverted condition does, however, show a differential effect on the whole faces compared with isolated features. This transition makes recognition of whole faces harder, whereas it does not significantly affect the isolated features. It would appear, therefore, that this transition, unlike the one described above, disrupts highly configural (or holistic) encoding of faces.

The conclusion from the results presented here is that the rotation of a face from upright to inverted has at least two separable effects on recognition. The first occurs over acute rotations and equally affects isolated features and whole faces. The second occurs between orthogonal faces and inverted faces and only affects the recognition of whole faces, in the same way that Tanaka and Farah's (1993) original inversion experiments only affected whole faces.

These conclusions can be seen as being consistent with either the generally accepted featural/configural dichotomy of facial encoding or with Lewis's (2001) continuum approach. In terms of the dichotomy, it has been suggested that configural encoding is not disrupted until a rotation greater than 90° (Stürzel and Spillmann 2000). The upright-to-orthogonal transition does not therefore disrupt configural encoding, whereas the orthogonal-to-inverted transition does. This explanation does not, however, account for the decrease in overall performance observed between the upright and orthogonal conditions.

Lewis's (2001) continuum of dimensions can also account for the pattern of data observed. It is suggested that the upright face is encoded on a spectrum of dimensions. These dimensions are disrupted by differing degrees of rotation. The more holistic dimensions lead to the advantage for the whole faces over the isolated features observed in the upright and orthogonal conditions. Less holistic configural features are disrupted between these two conditions, leading to an overall decrease in performance.

The transition between orthogonal and inverted faces disrupts configural features even further, leading to a reduction of the advantage for the whole face over the isolated features.

While the current experiment does not distinguish between a dichotomy and a continuum of facial dimensions, it does illustrate that there is a great deal to be learnt about how faces are processed in a variety of orientations. It has been shown that rotating a face has two different and dissociable effects on recognition depending on the range of rotation. This finding has consequences for the debate whether faces are special. The whole-face advantage for upright but not inverted faces is consistent with the idea that faces are special just because we are experts (or highly trained) at recognising them (see Gauthier and Tarr 2002). The current study, however, shows that this advantage, in fact, still occurs for orthogonally rotated faces. Unless we are willing to believe that our expertise generalises to orthogonal faces, simple expertise cannot account for why we find a whole-face advantage for orthogonal faces.

References

- Diamond R, Carey S, 1986 "Why faces are and are not special: An effect of expertise" *Journal of Experimental Psychology: General* **115** 107–117
- Gauthier I, Tarr M J, 2002 "Unraveling mechanisms for expert object recognition: Bridging brain activity and behavior" *Journal of Experimental Psychology: Human Perception and Performance* **28** 431–446
- Leder H, Bruce V, 2000 "When inverted faces are recognised: The role of configural information in face recognition" *Quarterly Journal of Experimental Psychology, Section A* **53** 513–536
- Lewis M B, 2001 "The Lady's not for turning. Rotation of the Thatcher illusion" *Perception* **30** 769–774
- Lewis M B, Johnston R A, 1997 "The Thatcher illusion as a test of configural disruption" *Perception* **26** 225–227
- Maurer D, Le Grand R, Mondloch C J, 2002 "The many faces of configural processing" *Trends in Cognitive Sciences* **6** 255–260
- Murray J E, Yong E, Rhodes G, 2000 "Revisiting the perception of upside-down faces" *Psychological Science* **11** 492–496
- Palermo R, Rhodes G, 2002 "The influence of divided attention on holistic face perception" *Cognition* **82** 225–257
- Rakover S S, Teucher B, 1997 "Facial inversion effects: Parts and whole relationship" *Perception & Psychophysics* **59** 752–761
- Rhodes G, Brake S, Atkinson A P, 1993 "What is lost in inverted faces?" *Cognition* **47** 25–57
- Sjoberg W, Windes J D, 1992 "Recognition times for rotated normal and Thatcher faces" *Perceptual and Motor Skills* **75** 1176–1178
- Stürzel F, Spillmann L, 2000 "Thatcher illusion: Dependence on angle of rotation" *Perception* **29** 937–942
- Tanaka J W, Farah M J, 1993 "Parts and wholes in face recognition" *Quarterly Journal of Experimental Psychology, Section A* **46** 225–246
- Tanaka J W, Gauthier I, 1997 "Expertise in object and face recognition", in *Mechanisms of Perceptual Learning* Eds R L Goldstone, D L Medin, P G Schyns (San Diego, CA: Academic Press) pp 83–125
- Tanaka J W, Sengco J A, 1997 "Features and their configuration in face recognition" *Memory & Cognition* **25** 583–592
- Thompson P, 1980 "Margaret Thatcher: a new illusion" *Perception* **9** 483–484
- Valentine T, 1988 "Upside-down faces: A review of the effect of inversion, and race on face recognition" *British Journal of Psychology* **79** 471–491
- Yin R K, 1969 "Looking at upside-down faces" *Journal of Experimental Psychology* **8** 141–145