From single cells to social perception

Nick E. Barraclough and David I. Perrett

References	**This article cites 117 articles, 42 of which can be accessed free** http://rstb.royalsocietypublishing.org/content/366/1571/1739.full.html#ref-list-1
Rapid response	Respond to this article http://rstb.royalsocietypublishing.org/letters/submit/royptb;366/1571/1739
Subject collections	Articles on similar topics can be found in the following collections neuroscience (492 articles) cognition (508 articles)
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

To subscribe to *Phil. Trans. R. Soc. B* go to: **http://rstb.royalsocietypublishing.org/subscriptions**

Review

From single cells to social perception

Nick E. Barraclough[1],* and David I. Perrett[2]

[1]*Department of Psychology, University of Hull, Hull, East Yorkshire, HU6 7RX, UK*
[2]*School of Psychology, University of St Andrews, Scotland KY16 9JP, UK*

Research describing the cellular coding of faces in non-human primates often provides the underlying physiological framework for our understanding of face processing in humans. Models of face perception, explanations of perceptual after-effects from viewing particular types of faces, and interpretation of human neuroimaging data rely on monkey neurophysiological data and the assumption that neurophysiological responses of humans are comparable to those recorded in the non-human primate. Here, we review studies that describe cells that preferentially respond to faces, and assess the link between the physiological characteristics of single cells and social perception. Principally, we describe cells recorded from the non-human primate, although a limited number of cells have been recorded in humans, and are included in order to appraise the validity of non-human physiological data for our understanding of human face and social perception.

Keywords: face; neuron; neurophysiology; adaptation

1. INTRODUCTION

The behaviour of single neurons is often incorporated into models of social perception and used to explain neuroimaging and psychological data. Many models of perceptual processing of social information in humans feature components that mimic the physiological properties of cells recorded in monkeys (e.g. [1–7]). Interpretation of functional magnetic resonance imaging (fMRI) data needs an understanding of how the measured blood oxygen level-dependent (BOLD) response relates to activity at the single-neuron level [8–10]. Psychophysical adaptation studies investigating social perception in humans make the explicit assumption that there are individual neurons in the brain that are selectively suppressed by prolonged exposure to perceptual stimuli [11–14]. The physiological characteristics of single cells are used to constrain the interpretation of these psychological data. Therefore, precisely defining the characteristics of sensory neurons is important for both the physiology and psychology of social cognition.

Perhaps the most important source of social information is the face. Observing another individual's face can provide information about the individual's identity, gender, age, health, emotions and intentions. This information can be used to form strategies of how to behave, and interact with that individual. In the early 1970s, single neurons were described that selectively responded to faces rather than other tested stimuli [15]. While the characterization of such cells was limited to the brief statement 'three units, complex coloured patterns (e.g. photographs of faces, trees) were more effective than standard stimuli, but the crucial features of these stimuli

were never determined' [15, p. 103], nonetheless debate started as to whether neurons, such as these, could form the basis of a brain system underlying face processing. Correspondingly, the behaviour of these neurons received considerable interest and studies followed detailing the properties of single cells coding faces in several distinct cortical regions. It was not until 2006 that a causal link between cell responses to faces and face perception was established. Monkeys were first trained to categorize different ambiguous images into face or non-face categories [16]. Cells were recorded from the inferior temporal (IT) cortex during passive fixation, to establish selectivity for faces compared with other non-face stimuli. Face-selective cells were found to be grouped into clusters (as has been noted in a variety of reports: [17–21]), and direct microstimulation of these clusters during a subsequent face categorization task biased the monkeys' decisions towards the face category.

Here, we consider the link between single cells coding faces and social perception by assessing early and more recent findings documenting the physiological properties of cells responding selectively to faces. Principally, this information comes from single unit recordings in the non-human primate brain (and typically macaque monkeys: genus *Macaca*). There are, however, some studies investigating cell coding of faces in humans. Where possible we compare between non-human and human recordings in order to appraise the validity of non-human physiological data for our understanding of human face perception.

2. FACE-SENSITIVE CELL GENERAL RESPONSE PROPERTIES
(a) *Face selectivity*
During typical physiological experiments, it is impossible to be entirely sure that one has found a
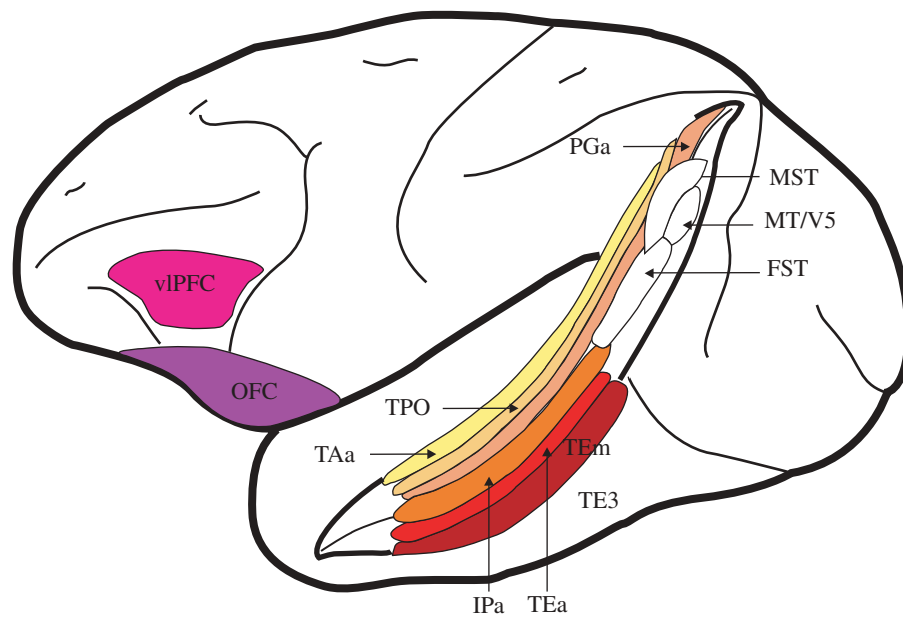

Figure 1. Illustration of some cortical regions in the monkey shown to contain face-sensitive cells (adapted from Seltzer & Pandya [25]). The STS has been opened out in order to illustrate regions within the sulcus. Inferotemporal cortex (IT) consists of regions on the surface of the temporal lobe TE3 and TEm that also extends into the lower bank of the STS. TEa is entirely situated on the lower bank. The floor of the superior temporal sulcus (FST) is usually considered part of the motion-processing system (including also MT/V5 and the medial superior temporal sulcus, MST), and to our knowledge contains no face-sensitive cells. Still within the floor of the STS, but more anterior to FST is IPa. PGa extends anteriorly within the floor of the STS. The upper bank of the STS consists of two elongated regions, TPO and TAa. Orbitofrontal cortex (OFC) and ventro-lateral pre-frontal cortex (vlPFC) also contain face-sensitive cells, as well as the anygdala, located within the pole of the temporal lobe, but not seen in this illustration.

'face-selective cell'. To establish this, it would be necessary to measure and compare the response of the cell to all visual stimuli, a clearly unfeasible approach. Despite these obvious practical difficulties, it has become well understood that there are individual neurons both in the non-human primate and human that respond to faces significantly more than other stimuli (figure 1).

The degree of response selectivity for faces (or 'tuning') has been investigated using a range of different sets of stimuli. The trade-off is between using a small stimulus set to isolate cells sensitive to faces quickly in order to investigate other properties of the cells, while knowing very little about the tuning of the cell, or indeed if the cell responds preferentially to faces when compared with other likely visual stimuli. Indeed, some cells will respond to faces, but also to other complex stimuli [22,23]. Alternatively, the time available to record from the cell can be spent determining the degree of selectivity for a face by comparing responses to faces against responses to a large range of non-face stimuli, at the cost of establishing other cell characteristics. By using a rapid serial visual presentation (RSVP) of large numbers (greater than 1000) of stimuli, Földiák et al. [24] attempted to establish an objective description of the functional characteristics of cells sensitive to faces in a viable timescale. From the sample of cells recorded in superior temporal sulcus (STS), a distribution was found between cells that responded to the images containing faces and cells that responded to the images without faces, confirming the existence of a population of cells in the cortex that respond 'selectively' for faces (e.g. figure 2; see also [26]).

The response of a face-selective cell represents one point within a multi-dimensional stimulus space (although see [27]). Rather than measuring the narrowness/broadness of cell tuning within that entire space, typically researchers assess cell tuning along a restricted dimension, for example, view of the face (e.g. [28]), by comparing the significance of responses to the preferred stimulus against other non-preferred stimuli and spontaneous activity using parametric statistics (analysis of variance (ANOVA)). Alternative approaches might use stimulus optimization procedures similar to those used successfully in identifying cell sensitivity in early visual cortex (e.g. [29]), unfortunately the visual elements contributing to responses to faces are more complex than pixels [30,31]. To date, such approaches for identifying tuning functions of face-sensitive cells in high-dimensional stimulus space have been limited by computational challenges [32]. Nonetheless, design of efficient algorithms in order to optimally adapt experimental design during the collection of neurophysiological data will help maximize information about the overall stimulus tuning function of face-selective cells [24].

Single-cell recording indicates that there are six patches of cells along the monkey STS (three in the right hemisphere, three in the left hemisphere) that respond to faces. These patches have distinct connections with other brain areas (e.g. the parietal cortex, [17]). Functional brain imaging shows the patches to be active when faces are viewed and 99 per cent of the cells within the middle patches respond more to images of faces than to images of other objects [20].

In humans, functional brain imaging also indicates six face processing patches in posterior cortical areas
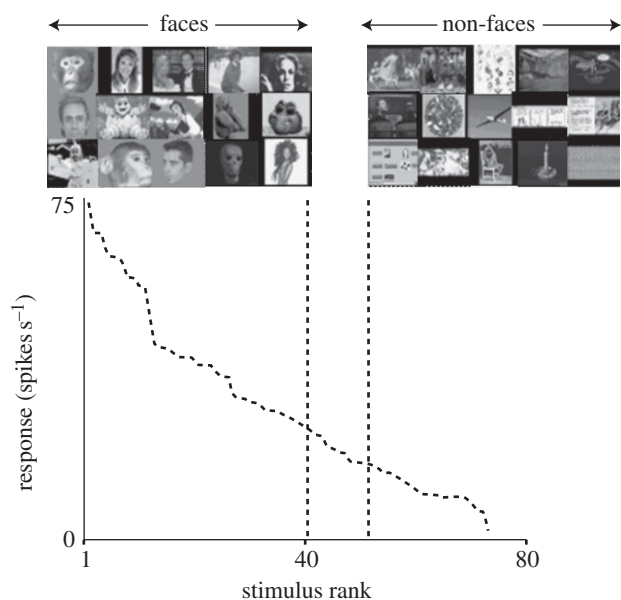
Figure 2. Responses of one cell selective for faces. Responses to more than a thousand images were compared with a progressive narrowing of the collection of images keeping the most effective 75% and the least effective 25% stimuli [24]. After a series of trials, the images were ranked in effectiveness and the 40 most effective stimuli all contained faces. Given the proportion of test images containing faces, the probability of obtaining this pro-face selectivity by chance would be one in 1 000 000 000 000.

(three per hemisphere); these occur in the fusiform gyrus which lies on the ventral surface of the temporal lobe, an occipital region [33], and a region within the STS (e.g. [34]). Of these the right fusiform area is the most consistently located and reliably activated by faces, and activity in the STS is most closely related to gaze signals. High-resolution mapping in humans indicates that the size of the fusiform patch in humans [33] is roughly equivalent to those described in the monkey STS [17,18]. The homology between the human and monkey face-responsive patches is not yet clear, but the similarities are striking.

(b) Latency and response profile

The time from the onset of a stimulus to the moment that a cell's firing rate raises significantly above its background rate is called the cell response latency. The magnitude of a significant response, however, can vary from a few spikes per second to over 50 Hz. Cell response latencies can be calculated separately for those stimuli to which they respond, although the latency of the response to the most effective stimulus is usually reported. For each stimulus, a spike density function (SDF) is first calculated by summing across trials and smoothing using a Gaussian function. Response latencies can be measured as the first 1 ms time bin, where the SDF exceeds 3 s.d. above the background firing rate (calculated within a time-window prior to stimulus onset, e.g. 100 ms) for a significant period (e.g. 15 ms) after stimulus onset [35–37]. There is some considerable variation in the latency of face-sensitive cell responses. First, different cells recorded in one cortical area can have different
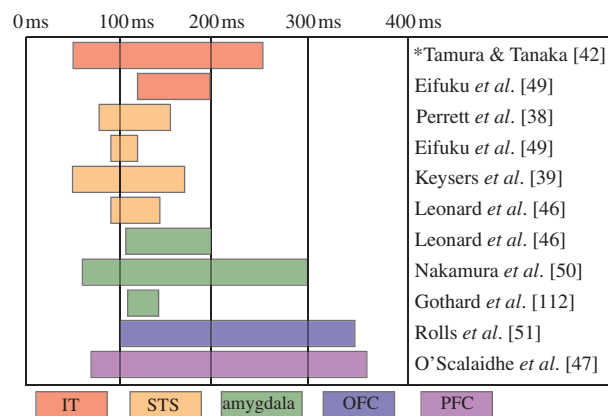

Figure 3. Latency ranges for face-sensitive cells recorded in different cortical areas in the monkey. Asterisk denotes that the estimate includes cells recorded with selectivity for other non-face objects.

response latencies to the same stimulus. For example, large latency ranges have been reported for STS cell responses to the same face stimuli: 80–160 ms [38] and 56–171 ms [39].

Second, a single cell's response latency can vary depending upon the nature of the stimulus. Kiani *et al.* [40] measured the response latencies of face-sensitive cells in IT to a range of different human, monkey and animal faces. Responses to human faces were significantly earlier (mean latency 103 ms) than responses to animal faces (mean latency 118 ms); response latencies to monkey faces were not significantly different from those to human faces, but significantly earlier than responses to non-primate animal faces.

Image contrast has additional and independent effects on latency with cells responding 220 ms later to faint low-contrast images than to high-contrast versions of the same image [35]. Remarkably, how much a given cell responds and how sluggishly it responds are largely independent: responses to low-contrast stimuli are invariably delayed but not necessarily decreased in magnitude. Many image transformations, such as change in the perspective view of the face, produce no change in latency [35,41] although they diminish cell responses.

Finally, response latencies depend upon where the cell is recorded (see figure 3). Cells sensitive to faces can be recorded from several different cortical areas. It is generally agreed that the initial selectivity for faces occurs in the temporal lobe via a hierarchical process where each stage feeds forward to the next. The first evidence of face selectivity occurs in IT cells. IT (or anterior-ventral TE; TEav) cells, in general, respond with a latency range 50–250 ms, mean ± s.d. 135.2 ± 92.1, median 106 [42]. This estimate includes the subset of cells in IT cortex that respond to faces. In other areas, the latency range for cells that respond selectively to faces does not differ from the latency range of cells selective for other stimulus classes [39].

The upper bank, lower bank and fundus of the STS receive input from IT and response latencies of cells responsive to faces range from 56 to 171 ms [39]. STS connections with IT cortex are bidirectional and it is difficult to be certain of the sequence of

processing between the two areas, particularly as both the STS and IT cortex are composed of a variety of functionally and anatomically distinct areas along the anterior–posterior extent [43–45].

Both IT and STS provide feed-forward input to amygdala. In a comparison between cell response latencies to the same stimuli, STS cells showed earlier responses (90–140 ms) than amygdala cells (110–200 ms, [46]). Pre-frontal cortex also contains visually responsive cells that are selective for face stimuli; cells here can respond as early as 70 ms, mean latency is 138 ms, although cells have been found with latencies as late as 360 ms [47].

A few studies have measured the response latencies of face-sensitive cells recorded in humans. Face-selective cells were recorded from the medial temporal lobe (MTL: comprising the hippocampus, amygdala, entorhinal cortex and parahippocampal gyrus) and found to respond approximately 250–350 ms after stimulus onset [27]. Much larger variances in response latencies have been reported for visually responsive cells in the amygdala (95–385 ms), entorhinal cortex (90–328 ms) and hippocampus (107–371 ms); indeed some MTL cells showed latencies as early as 52 ms [48].

3. CELL-RECEPTIVE FIELD SIZE

For a long time, it has been assumed that face-sensitive neurons in temporal cortex have very large receptive fields. The earliest findings reported large (greater than $50°$) receptive fields that always covered the fovea, which could be contralateral, or ipsilateral to the recorded hemisphere [15]. Furthermore, many of these cells in the STS had receptive fields that extended both ipsilaterally and contralaterally [22], and were thought to be the first cells in the visual-processing hierarchy that could represent objects in either visual field. In temporal cortex cells selective for faces, the large receptive fields and broad tuning profiles (across low-level image statistics, e.g. size), were considered commensurate with their position at the latter end of the ventral visual-processing stream.

Along with large receptive fields, face-sensitive cells can show a relative tolerance to the position of the face within the receptive field and the size of the face itself, although typically, there is a foveal 'hot spot' where faces elicit maximal cell responses [52]. Cell responses decline slowly as faces are presented progressively further away from this central region into the periphery. Tovee *et al.* [53] studied the sensitivity of a range of IT and STS cells, which were selective for different face identities, to changes in the size of the face and the position within the receptive field. Receptive fields extended at least $5°$ into the ipsilateral field, and the greatest cell response was observed when the face was presented at the fovea. These cells could tolerate quite large shifts in the position of the face without a significant decrease in the cell response. Responses to images of 'large' faces subtending an angle of $17°$ showed no significant diminution in response even when fixation was beyond the edge of the face itself. The cell firing rate, which provided information about the stimulus, predominantly coded facial identity rather than face position. Relative

position invariance, as well as slow decline in cell responses towards the periphery, seems to underlie and support findings of position invariance for face adaptation in human studies [54,55].

Since the early descriptions of the receptive field profiles of face-sensitive cells, there have been an increasing number of reports of much smaller receptive fields (discussed in Afraz & Cavanagh [54]). Furthermore, the concept of the face-sensitive cell receptive field as a static filtering device for faces might be questioned. Rolls & Tovee [56] describe STS cell receptive field position sensitivity changing depending upon the presence of other non-face stimuli in the visual scene. STS cells responses to images of faces away from the fixation point were markedly reduced when a non-effective stimulus was presented at the fovea. Thus, the typical translation invariance observed in these cells [57] reduces with the presence of other stimuli. Shrinking of the effective receptive field size and weighting of the response to the stimulus present at the fovea allow these cells to effectively represent the face that is being fixated, rather than responding when the face occurs anywhere in the receptive field.

We have found, testing at multiple locations in the visual field, that STS cells selective for faces (and those cells selective for other social stimuli such as hand actions) can have restricted and eccentrically located fields (D.-K. Xiao, N. E. Barraclough & D. I. Perrett 2004, unpublished data). A response field would include the fovea, but the maximally sensitive receptive field position could lie away from the fovea by $3–5°$. Considering cells collectively, the fovea would be the most effective single location for responses, but individual cells would have receptive fields centred away from the fovea. This finding is particularly relevant for understanding how face-sensitive cells operate in naturalistic environments. Faces are very rarely experienced in isolation, being only one part of our rich and cluttered social scene. If cells responded to faces almost anywhere within a large scene (showing complete translational invariance across central vision), then a large number of face-sensitive cells could be simultaneously active. This large population could not be used to determine where a face lies exactly in relation to the fixation point. With a population of cells that are selective for both pattern (be it a face or a hand) and position, then it is possible to define from the population the presence of different objects and their locations. Indeed, conjoint tuning for object and location (within moderately large receptive fields $5–10°$ across) makes it possible to derive the relation of objects to one another, for example, how a hand or face is interacting with another object [58].

4. EFFECTS OF ADAPTATION ON FACE-SENSITIVE CELLS

In the past decade, there has been considerable use of adaptation, employed both during psychophysical experiments and neuroimaging experiments, as a technique to investigate the brain mechanisms underlying face processing [55,59–63]. In its most basic form, adaptation results from prolonged exposure to a stimulus

that causes a selective suppression of the neurons that code that particular stimulus, sparing neurons that code different stimuli. This short period of selective suppression can result in a period of imbalance in activity across the perceptual system causing 'after-effects' in which perception is biased.

During the 1970s and 1980s, there was a proliferation of experiments demonstrating adaptation of both neural responses in monkey neurons, and monkey and human perception, after exposure to a range of simple stimuli (e.g. colour, oriented lines and moving dots; for a recent review see [64]). These experiments suggested that it was possible to investigate the responses of individual neurons coding particular stimuli in humans using psychophysical adaptation techniques. Since that time, it has become clear that there is not a simple link between exposure to a stimulus and the resultant response decrease in a select population of neurons. Adaptation can have differing effects on neurons recorded in different cortical areas [65,66], and more than one mechanism of adaptation might be acting after exposure to a stimulus (e.g. [67]), resulting in different post-adaptation perceptual effects.

Despite increasing evidence that adaptation has a complex effect on neural mechanisms underlying perception, and the reliance of studies investigating face processing on adaptive techniques, there are few reports of how adaptation changes neuronal coding of faces. Much of the early evidence of adaptation to repeated presentations of complex stimuli comes from studies recording from the IT cortex. Cells in IT respond to a range of different complex stimuli (shapes, pictures or faces) and after an initial presentation of the novel stimulus, a subsequent presentation of that same stimulus can result in a smaller cell response [26,68–71]. In a study of adaptation in IT cells (referred to in the paper as a habituation-like response), responsiveness of cells immediately declined following the initial presentation of the stimulus [70]. Cell sensitivity slowly increased with time following this initial presentation; the effect of adaptation was evident 12 s after stimulus exposure, but not after 20 s. These results indicated that the effect of adaptation on IT cells lasted up to about 12 s, but normal sensitivity resumed afterwards.

The above studies did not explicitly describe faces as part of their stimulus sets. Rolls *et al.* [72], however, did test the sensitivity of cells in IT to repeated presentations of face stimuli. Many cells selective for faces showed larger responses to the initial presentation of a novel face stimulus than for a subsequent presentation; a few other face-sensitive cells were observed that showed greater responses to the subsequent presentation of the face [72]. In an investigation of IT cell responses to familiarity, Li *et al.* [73] included faces in their stimulus set, although did not distinguish face-sensitive cells from cells preferentially sensitive to other complex stimuli. Again, in this study, responses to repeated presentation of the stimulus declined to a level about 40 per cent of the response to the initial presentation of that stimulus. This decline in sensitivity was observed even when up to 150 intervening stimuli were presented between the first and second presentation of the stimulus [73]. In contrast to Rolls

et al. [72], no cells showed a larger response on subsequent presentations of the same stimulus, although this may be owing to an influence of the delayed match to sample task in which the monkeys were engaged during Li *et al.*'s [73] experiment. Li *et al.*'s study indicates that cellular adaptation resulting from the presentation of a face can last a significant period of time, considerably longer than seen during psychophysical adaptation experiments [74].

Psychophysical adaptation experiments and functional magnetic resonance-adaptation (fMR-A; [75,76]) studies of face processing make assumptions about the nature of face-sensitive neuronal adaptation. Often the goal of these studies in the human is to determine the stimulus sensitivity of the mechanisms underlying face processing. The explicit assumption here is that the magnitude of a cell's response to different stimuli is directly and proportionally related to the degree of adaptation of the cell's response with increased exposure to that stimulus. By studying the degree of adaptation (in terms of perceptual after-effects or decreases in the BOLD responses), it would therefore be possible to investigate cellular sensitivity indirectly.

Concern for the relative lack of evidence on the relationship between cell tuning to complex stimuli and the stimulus selectivity of adaptation has led Sawamura *et al.* to investigate this directly [77]. As for many of the earlier studies of responses to repeated stimulation, IT cells were presented with a range of complex images (objects and animals) without specifically studying face-selective cells. Neurons were observed with distinct tuning profiles, where a discrete selection of images would generate responses. The range of stimuli that generated response suppression, however, was more restricted, illustrating tighter tuning profiles. For these IT cells, the response tuning profile and the adaptation tuning profile did not overlap. So, if one was to measure adaptation tuning and then use this information to infer the response selectivity of the neurons, one would underestimate the broadness of the response tuning profile. Psychophysical adaptation and fMR-A [78] experiments typically provide data that show cross adaptation between some stimuli, but not others. For example, adaptation to stimulus A may not only cause an after-effect or suppression of response to repeated instances of stimulus A but also stimulus B. Adaptation to stimulus A, however, may cause very little or no after-effect in, or response to stimulus C. A conclusion might be drawn that the adapted neurons respond equally to stimuli A and B, and not to stimulus C. Sawamura *et al.*'s data indicate caution should be taken in such a conclusion as the adaptive technique may be underestimating the true sensitivity of the adapted neurons to stimulus C.

Is this a problem for studies of face processing using adaptation? First, despite the above evidence, it is still not clear how face-sensitive cells adapt to repeated presentation of face stimuli. Different cortical areas contain neurons that adapt differently [66]; areas that contain significant proportions of neurons that respond selectively to faces [79] may show different effects of adaptation. Second, Sawamura *et al.*'s results [77] pertain to the precise relationship between

response and adaptation tuning to stimuli that lie close to each other in stimulus space. Investigators should bear in mind that psychological and fMR-A experiments involving faces and other complex objects may be underestimating the breadth of tuning of neurons responsible.

(a) Short-term adaptation

The effects of stimulus repetition may well involve several different mechanisms operating over different timescales. Repetition of a similar image after a short time interval (less than 1 s) produces a marked reduction of response (to the second presentation). This response short-term suppression is found for all STS cells responsive to faces and other objects [26]. It is likely to be the counterpart to the forward masking that occurs in human perception when the presentation of one image disturbs the perception or detection of similar images presented after a brief interval [80].

The disruptive effect that one stimulus has on following stimuli is somewhat paradoxical given that experience of visual sequences allows anticipation and faster reactions to upcoming stimuli [26]. To understand this paradox, a comparison was made of cell responses to body images presented individually, in pairs and in action sequences. Responses to one image did suppress responses to similar images for about 500 ms. This suppression led to responses peaking 100 ms earlier to image sequences than to isolated images (e.g. during head rotation, face-selective activity peaks before the face confronts the observer). Thus, forward masking has an unrecognized benefit for perception because it can transform neuronal activity to make it predictive during natural changes in view.

The anticipatory responses that occur in natural sequences of images parallel the speeding up of reaction times to images that are about to occur. Observers detect a face view faster in a film of a head turning to confront a camera than when it is presented in isolation or in a sequence in which the images comprising the film have been randomly reordered [26]. The advantage in reaction times to views anticipated in natural sequences also has a hidden cost: when the natural sequence is violated and the anticipated view does not occur, there is a bias to report (erroneously) its presence. Hence, interactions between successive stimuli allow anticipation which has benefits in speeding up reactions to stimuli that are predictable given the continuation of natural and familiar sequences, but the brain's anticipation can mean observers 'jump the gun' and respond to events and stimuli before they actually occur.

(b) Long-term adaptation

As noted above there may be several different mechanisms, operating over different time courses, that contribute to adaptation at the neural and perceptual levels. Most perceptual experiments on adaptation and after-effects document rather short duration effects lasting over seconds or minutes [74]. One type of face (e.g. expanded) is shown repeatedly and subsequently a negative after-effect is induced, whereby normal-shaped faces now appear contracted (e.g. [62]). Such effects can be contingent on facial characteristics (e.g. age, sex, race, expression, species, view, orientation) and it is possible to adapt different classes of faces in opposite ways such that simultaneous after-effects are induced in opposite directions, for example, after looking at male and female faces transformed in opposite directions (e.g. [81]). These effects are often interpreted as evidence for separate populations of cells encoding different categories of faces (e.g. [61,81,82]). It is possible that category contingent perceptual after-effects with faces could have long duration, much like contingent after-effects reported for simple visual parameters [83]. Long duration after-effects may be missed because participants return to normal face experience following experiments.

There are reports that perceptual after-effects from viewing faces can be detected after much longer periods of more than 24 h [63,84,85]. These after-effects must depend on neuronal effects of repetition-lasting days, and such effects have been described. For example, in reticular nucleus of the thalamus and in the mamillary bodies [86], cell responses change when a stimulus becomes familiar after being viewed for a brief period of 1–5 s. The next time the same stimulus is seen the responses are augmented or depressed depending on the particular cell [87]. Several cells studied showed stimulus effects of stimulus familiarity lasting days, and estimates of the duration of the 'memory' of individual cells indicate that it can exceed more than 700 intervening stimuli between the first and subsequent presentations [86]. The thalamic cells exhibit familiarity effects with all types of visual stimuli and show limited generalization over image transformations (such as rotation in the image plane). The thalamic cells could be pooling long-term perceptual learning and familiarity effects generated in the visually responsive cells within temporal cortex (longer lasting than those demonstrated so far, [73]); or they may be an independent memory effect generated in the entorhinal–hippocampal system and passed to the mamillary bodies and thalamus via the fornix. Exactly how these neural mechanisms for long-term 'familiarity' interact with sensory encoding of faces and other objects is unclear, but they could well be involved in the phenomenology of many perceptual after-effects where unusual faces (or other objects) come to look 'normal' or 'familiar' after prolonged viewing.

5. CODING OF IDENTITY

Accurate recognition of other individuals is critical to successful social functioning. There are clear advantages to being able to dissociate familiar individuals, with whom one may have had a previous positive or negative interaction, from other individuals on whom no information is available. Face structure varies widely, and this information can be used to identify specific individuals. Several reports have described that a significant proportion of cells selective for faces can also distinguish between different facial identities [18,23,28,88–92]. These studies have received considerable interest, as the cells described could be part of a system underlying the recognition of individuals.

In a facial recognition system, the ability to distinguish faces under varying conditions is important. Neurons selective for faces are relatively insensitive to changes in low-level stimulus attributes, like contrast, size and colour [38,93], which may occur under different viewing circumstances. Some neurons in IT and the STS will respond selectively to the faces of specific monkeys, irrespective of other facial characteristics [94] and irrespective of the viewing conditions with changes in lighting, face orientation, size and distance and in some cases change in perspective view [18,23, 88–90]. Monkeys make different facial expressions under different behavioural circumstances, and these expressions change the facial shape considerably. Despite these changes in facial shape, facial identity can be preserved [94]. Thirty-three per cent of recorded neurons in Hasselmo *et al.*'s [94] study showed significant differences recorded from face-selective neurons while systematically varying facial identity and facial expression in their responses to different monkeys, independent of the expression of the face itself. Very few of these neurons in IT showed an interaction between identity and expressions, a very different coding from that observed in amygdala cells (see §7).

The selective coding of individual faces by single neurons provides a basis from which a system of facial recognition can be built. This system is unlikely to consist of a single-neuron coding a specific individual (so-called 'grandmother' cell coding), as this brings some considerable practical and computational problems (for discussion see: [28,89,95–97]). A more likely coding system for facial identity is to involve populations of neurons that are, to a lesser or greater degree, selective for individual faces. Young & Yamane [92] recorded from a large population of anterior IT cells and STS cells while the monkeys performed a facial discrimination task. By applying multi-dimensional scaling (MDS) to the two populations of cells, they quantified the respective population responses, and compared them with measurements of the physical characteristics of the faces and relative familiarity of the faces. The IT population response showed a significant relationship with the physical characteristics of the face, and the STS population response, a significant relationship with the familiarity of the face. Young & Yamane [92] also suggest that a population of cells coding facial characteristics does not have to be particularly large ('sparse coding'); only a few tens of cells were necessary to form a precise code sufficient to identify individual face images from their collection of faces. From the available evidence, it seems that to recognize any familiar face across a variety of viewing situations requires a population of cells, but the population does not need to be vast since individual cells respond reliably.

Investigations of MTL cell coding of faces in humans also indicate that identity may be coded by a sparse population of neurons, despite the apparent precise tuning of a few recorded cells. Some cells in MTL show tuning for a single individual, while generalizing across incidental visual properties of the stimulus, and responding to very different pictures of that specific individual [27]. Quiroga *et al.* [97],

however, argue that the probability of finding neurons that respond *only* to the individuals to which they are apparently tuned is too great, and indeed point out that many cells have specific tuning to more than one individual. In a probabilistic analysis of the data available from several recordings of selective cells in the human MTL, Waydo *et al.* [98] argue that each recorded neuron is likely to respond to 50–150 different images. This analysis indicates a sparse code in which faces and objects are represented by a small subset of neurons rather than either a large population of cells or by individual cells. The selectivity of MTL cells and sparseness of representation are likely to be underestimated because interviews with patients prior to testing allows the experimenters to focus testing on faces, and places that are well known and of particular interest to the patients; hence the probability of finding specific responses is greatly magnified when compared with a random unguided search.

In monkey STS and IT, different parts of the cellular response can carry different information about the face [99]. The early part of the response (peak transmission 117 ms following stimulus onset) appears to distinguish between global categories of stimuli, for example, whether they are monkey faces, human faces or objects. The later part of the cellular response (peak transmission 165 ms following stimulus onset), is sensitive to much finer analysis of the face and can distinguish between different facial expressions and identities within the category [99]. Sugase *et al.* [99] argue that the initial response of these face-sensitive neurons is likely to be owing to a feed-forward process distinguishing between broad categories of faces and may provide an initial orienting of the face-processing network. The later response could be guided by feedback from the amygdala containing face-sensitive neurons that distinguish between monkey face identity and expression.

Recently, psychophysical experiments in humans have indicated that faces with 'average' identities play a particularly important role as a reference to which other face identities are coded [55,100]. The importance of the average identity face is also apparent in monkey IT neurons coding faces [101]. Leopold *et al.* [55,74,101] created a range of different face stimuli by taking four individuals' faces and morphing between them to generate a multi-dimensional 'face-space'. At the centre of this face space is a face with the average of the four identities: the 'average face'. Cells with tuning centred on this average face were particularly well represented; cells with tuning centred on an identity at the periphery of the face-space were less common. For cells most responsive to identity, their responses incremented as the face configuration was moved from the average along the identity axis.

6. OPPONENT AND POPULATION CODING

A recent study of cells responsive to faces in the middle temporal patch confirms that cells are tuned to several facial features and to their configuration (as had been indicated in early recordings, e.g. [38]). The systematic exploration of responses to 16 real and cartoon

faces and to their component features revealed that cell tuning to values of individual face features (e.g. eye separation) was ramp shaped (with maximal response to one end of the feature continuum) [102]. Thus, cells were tuned to faces with an aspect ratio making them slender, or to an aspect ratio making them broad; either cells were tuned to long hair or they were tuned to short hair; likewise they were tuned to thick or to thin hair, to widely separated or to closely separated eyes; to large or small eyes; to eyes fixating centrally or eccentrically; to high- or low-feature positioning within the face outline and to slanting or to non-slanting eyebrows.

Psychophysics of perceptual adaptation to faces is often explained in terms of underlying coding of a face-space centred on average facial values. Within this space, it is argued that perception of particular facial values relies on opponent coding [62,103,104]. This mode of coding is akin to that assumed to underlie perception of motion, where movement to the left is contrasted with motion to the right; and motion up contrasted with motion down. Likewise, opponent processing is assumed to underlie colour perception with red being contrasted with green, and blue contrasted with yellow [104], viewing of downward motion produces a familiar motion after-effect of upward motion (the waterfall illusion) and prolonged viewing of red gives a green after-effect. Adaptation to faces distorted in one way, e.g. features lowered, makes the features of a normal face appear distorted in the opposite way, e.g. raised. The cell tuning for individual facial parameters described by Freiwald *et al.* [102] fits the predictions for opponent processing of facial features with linear coding [104].

Not all dimensions of facial coding are binary or opponent. For example, with respect to head view, there are cells tuned to many different views in the horizontal plane, although there is a preponderance of coding of face, profile and the back of the head views [28,90]. There are further cells tuned to the vertical posture of the head coding head raised towards the sky or head lowered towards the ground. Perceptual adaptation to different gaze directions (to the left, right and directly at the observer) provides evidence that human gaze directions reflect coding by at least three populations rather than relying on a binary opponent-processing mechanism contrasting gaze left and right [105].

At the cellular level, there is evidence of a hierarchy in the encoding of gaze, head and body direction. Many cells are sensitive to two parameters and some are sensitive to all three parameters. Where conjoint sensitivity is found, the cells give most weight to gaze direction, followed by the head direction [23,90]. Given the interactions evident at the cellular level (particularly those between head and gaze), one would expect psychophysics to provide evidence that gaze, head and body direction interact and perhaps cross-adapt.

It is not clear with 'norm' or average-based face coding how many norms there are, or how perceptual judgements are referenced to the appropriate face category average value. Perceptual adaptation effects are indicated for face species, view, race and for age

[63,106–108]. For each of these dimensions, it is possible to think that several categories of cells exist depending on an observer's experience. Take age for example. We distinguish readily between infants, toddlers, children, teens, young adults and old adults. It is insufficient to conceive of a space of faces centred on a single average of all faces experienced. Instead, it is more likely that experience of particular face categories builds up cells tuned to the dimensions of this category (e.g. [107]). It should be possible to define adaptations dependent on such perceptual expertise.

7. CODING EMOTION AND COMMUNICATIVE GESTURES

One's face can convey information about internal states, whether this is involuntary or under conscious control. Involuntary facial movements of an agent can indicate that the agent is in pain, or convey the emotions of the agent, like anger or surprise. Conscious control over the movement of the face allows voluntary communication with other individuals. A sophisticated facial recognition system should be able to distinguish between different facial configurations, and use this information to read another individual's behaviour and intentions.

There is considerable evidence that monkeys, like humans, are able to use the facial gestures of other conspecifics (e.g. [109,110]). Monkey facial expressions and facial communicative gestures, however, are considerably different from human facial gestures. Properties of neurons recorded in the monkey that code facial gestures and expressions need to be interpreted in the light of what is known about monkey behaviour rather than human behaviour. There is not an immediate and obvious translation from the monkey to human model in this case, although in more closely related non-human primates, chimpanzees, it has been suggested that homologous emotional expressions to those observed in humans are found [111].

Monkeys have unique and clear facial expressions that convey positive and negative emotions. They also use oro-facial gestures during social interactions with other monkeys potentially subserving a communicative role (figure 4). Cells responding selectively to static images, movies or real-life monkey facial gestures can be found in many cortical regions, including the upper bank, lower bank and fundus of the STS [18,94], amygdala [46,112] and area F5 of pre-motor cortex [113].

STS cells will distinguish between images of faces showing neutral expressions, open-mouthed mild and strong threat gestures, and lip-smacking (an affiliative gesture [94,99]). The cells also distinguish between mouth opening made during emotional expressions and chewing or yawning [26,114]. Cells distinguishing between similar categories of expression are also found within several nuclei within the amygdala [112].

Coding of facial expressions appears different within different areas. Within areas IT and STS, neurons separately code either facial expression or facial identity; Hasselmo *et al.* [94] found that only
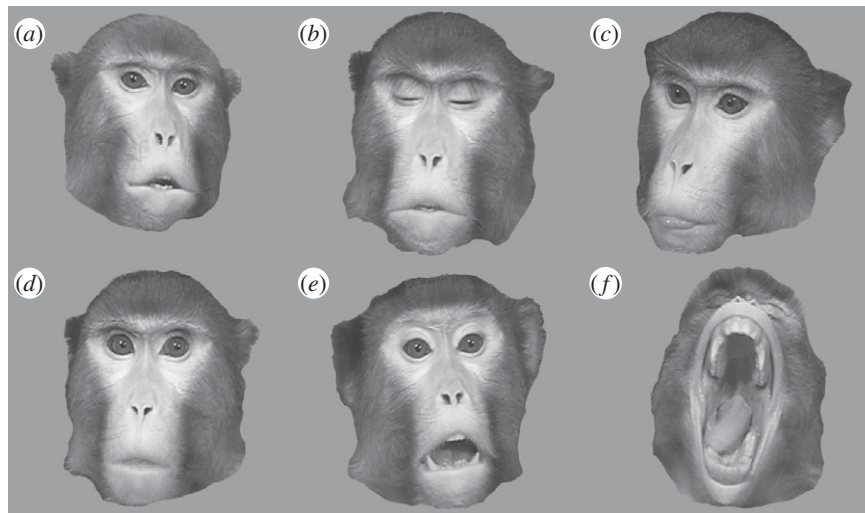
Figure 4. Images of different facial expressions performed by one individual macaque. (*a*) Coo—an affiliative call often used to contact other individuals over short and medium distances, characterized by a pursing and protuberance of the lips. (*b*) Grunt—a soft sounding call associated with positive interactions and food, often accompanied by blinking. (*c*) Licking lips—in the absence of food, associated with grooming. (*d*) Lip-smacking—a repetitive subordinate gesture used to appease more dominant individuals. (*e*) Pant threat—a call used as a threat, characterized by opening of the mouth showing teeth and staring. (*f*) Yawn—opening of the mouth and showing of the teeth, raising of the head and often eyelid closure.

7 per cent of neurons expressed an interaction between expression and identity. Furthermore, selectivity for facial expressions in IT and STS neurons is always expressed as an increase in the firing rate of the neuron to a particular expression [94]. The amygdala shows a different coding pattern. Here, cells are often more selective, coding for both facial expression and facial identity [46,112]. Gothard *et al.* [112] found 64 per cent of amygdala cells showed coding of unique combinations of identity and expression, underlining the role of amygdala cells in the interpretation of complex social interactions. In contrast to neurons recorded in more posterior areas of the temporal lobe, amygdala neuronal selectivity is not always exemplified as an increase in firing rate; many cells will also show a selective decrease in firing rate to a particular expression [112].

Information about the proportion of cells coding a particular expression in one area will of course be subject to the constraints of the search set and methodology used. In the amygdala, there does not appear to be a particular predominance of cells coding one expression over another, but Gothard *et al.* [112] found that there are some differences in the nature of the response. Cells selective for threat expressions typically increased their response compared with responses to other non-threat expressions. Cells selective for lip-smacking typically decreased their response compared with responses to other facial gestures. Expression-selective neurons that responded preferentially to threatening faces also tended to have higher firing rates than neurons selective for other expressions. This difference in population response was found only during the period 120–250 ms following stimulus onset.

In contrast to temporal cortex coding of facial expression, neurons in pre-motor cortex code facial expressions in a different fashion. Ferrari *et al.* [113] analysed the visual responses of motor neurons responding to movements of the face. These neurons express 'mirror' properties, responding to both execution and observation of actions, and show many similar properties to the hand action 'mirror neurons' often cited [115,116]. Ferrari *et al.*'s neurons demonstrated motor properties, responding during the execution of mouth movements alone (e.g. pursing of the lips) or during both mouth and hand actions (e.g. bringing food to the mouth to eat). The neurons often responded selectively to the sight of faces, most (85%) responded to the sight of ingestive mouth actions, but 15 per cent responded to the sight of communicative facial gestures (e.g. lip-smacking). The range of different facial stimuli to which these neurons responded was quite large, including grasping with the mouth, chewing, sucking, lip-smacking, lip protrusion, tongue protrusion and teeth-chattering. All the neurons responded during the execution of ingestive acts. The clear visual selectivity indicates a potentially significant role in the understanding of other agents. The mismatch between the visual selectivity and motor selectivity in those neurons that respond to communicative facial gestures leads to some problems in the interpretation of their role. Ferrari *et al.* [113] argue that this link between communicative gestures and ingestion in single neurons may result from similar motor acts that occur during positive social interactions like grooming and eating.

8. FACE-SENSITIVE CELL RESPONSE MODULATION BY OTHER MODALITIES

Facial movements are often accompanied by sounds. For example, an individual expressing surprise will not only move their face in a distinctive fashion, but this may also be accompanied by an audible intake of air. Furthermore, speech occurs with specific facial movements that, when observed, help disambiguate the auditory signal [117]. The common pairing of specific facial movements with particular sounds provides 'matched' audiovisual information
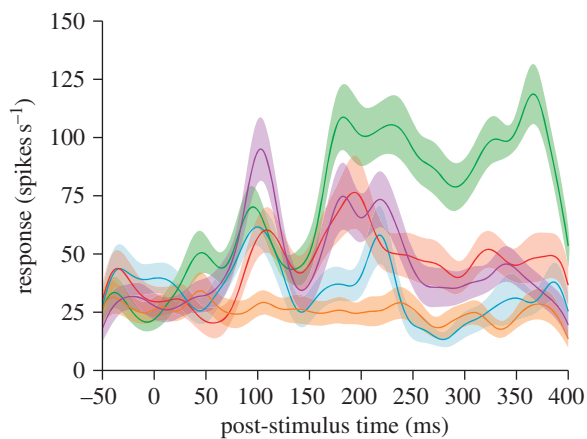
Figure 5. An example, superior temporal sulcus neuron that codes the audiovisual oro-facial communicative gesture 'coo'. Rhesus macaques use coos to communicate over varying distances with allies. There is a small visual response (red), and very little auditory response (orange). When the visual and auditory signals are presented simultaneously, there is a marked augmentation of the response (green). Pant–threat vocalizations (a threatening signal to a macaque) result in a poor response (blue), and when paired with the sight of the coo oro-facial gesture do not result in much of an increase in the cell response (purple). Responses are spike density waveforms with 12 ms Gaussian smoothing; the shaded areas indicate the standard error of the mean. The data are adapted from Barraclough *et al.* [119]. Red line, coo vision; orange line, coo sound; green line, coo vision + coo sound; blue line, pant–threat sound; purple line, coo vision + pant–threat sound.

that is available to the brain in order to decode a particular facial gesture.

There is evidence that neurons in the STS are sensitive to audiovisual oro-facial stimuli and respond in a way that allows them to use both auditory and visual signals to distinguish specific facial communicative gestures from other possibilities [118]. Barraclough *et al.* [118] recorded from single neurons in the upper bank, lower bank and fundus of the STS while searching for neurons that responded to different audiovisual facial gestures performed by monkeys and human (as well as hand and body actions). We found that 23 per cent of cells recorded had their visual response significantly modulated by the concurrent presentation of the 'matching' sound: in half of those cells, the response increased with the addition of the sound; in the other half responses decreased with the sound. In those cells that responded with an increase in the visual responses with the addition of an auditory signal, this was dependent upon sound-matching the vision. Non-matching sounds had little effect on the visual response (figure 5 illustrates the responses of such a cell). Neurons with such specific audiovisual characteristics could form the basis of a multimodal representation of social stimuli [120] and underlie the ability to recognize facial gestures including communication signals (cf. [109]).

In contrast, for cells that showed a decrease in the visual responses with the addition of an auditory signal, this was *not* dependent upon the vision and sound matching; in fact, any auditory signal appeared

to reduce the visual response. What role these neurons have in multimodal perception is not clear but they appear prevalent in other areas of cortex that show multimodal integration. For example, neurons that respond selectively to monkey vocalizations in lateral belt auditory cortex can show response augmentation or attenuation with the addition of visual signals (e.g. [121]).

9. CONCLUDING REMARKS

In this paper, we have reviewed cell properties in relation to face perception. There has been a rapid development of adaptation experiments on face perception and of functional brain imaging studies using adaptation effects. The interplay between these experiments and the physiological studies of cells responsive to faces has been increasingly convergent, although there are limitations between the correspondences. One domain that will need to be clarified concerns the time-courses of adaptation and the duration of after-effects. It is most likely that there are multiple cellular mechanisms at play. It will be important to differentiate separate adaptation effects that are dependent on different levels of visual processing and to differentiate these from adaptation effects that are dependent on brain systems involved in familiarity and memory. Yet, other effects of long-term experience (e.g. [63,107]) may produce shifts in criteria, so that particular nuances in expression, age or mix of ethnicity come to be regarded as typical. It is not unreasonable to speculate that long-term experience of mood disorders (either those affected themselves or close family thereof) may change the level of facial expressions that come to be regarded as normal, which could have consequences for social interactions that are clinically significant.

That cells responsive to faces are sensitive to multiple modalities imply that there should be psychological interactions that depend on the congruency of visual and auditory cues. While some such interactions have been demonstrated (e.g. [122]), future psychophysics, perceptual and fMRI adaptation studies could well tap into these strong cross-modal links evident at the cellular level. High-level cross-adaptations between faces and other conceptually linked visual stimuli should be expected (e.g. between hand and face, between face and body, and between face and gaze).

REFERENCES
1   Casile, A. & Giese, M. A. 2005 Critical features for the recognition of biological motion. *J. Vis.* **5**, 328–360. (doi:10.1167/5.4.6)
2   Giese, M. A. & Poggio, T. 2003 Neural mechanisms for the recognition of biological movements. *Nat. Rev. Neurosci.* **4**, 179–192. (doi:10.1038/nrn1057)
3   Jiang, X., Rosen, E., Zeffiro, T., VanMeter, J., Blanz, V. & Riesenhuber, M. 2006 Evaluation of a shape-based model of human face discrimination using fMRI and behavioural techniques. *Neuron* **50**, 159–172. (doi:10.1016/j.neuron.2006.03.012)
4   Lange, J. & Lappe, M. 2006 A model of biological motion perception from configural form cues. *J. Neurosci.* **26**, 2894–2906. (doi:10.1523/JNEUROSCI.4915-05.2006)

5 Riesenhuber, M., Jarudi, I., Gilad, S. & Sinha, P. 2004 Face processing in humans is compatible with a simple shape-based model of vision. *Proc. R. Soc. Lond. B* **271**, S448–S450. (doi:10.1098/rsbl.2004.0216)

6 Schindler, K., Van Gool, L. & de Gelder, B. 2008 Recognising emotions expressed by body pose: a biologically inspired neural model. *Neural Netw.* **21**, 1238–1246. (doi:10.1016/j.neunet.2008.05.003)

7 Wallis, G. & Rolls, E. T. 1997 Invariant face and object recognition in the visual system. *Prog. Neurobiol.* **51**, 167–194. (doi:10.1016/S0301-0082(96)00054-8)

8 Kreiman, G., Hung, C. P., Kraskov, A., Quiroga, R. Q., Poggio, T. & DiCarlo, J. J. 2006 Object selectivity of local field potentials and spikes in the macaque inferior temporal cortex. *Neuron* **49**, 433–445. (doi:10.1016/j.neuron.2005.12.019)

9 Logothetis, N. K. 2002 The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal. *Phil. Trans. R. Soc. Lond. B* **357**, 1003–1037. (doi:10.1098/rstb.2002.1114)

10 Wilke, M., Logothetis, N. K. & Leopold, D. A. 2006 Local field potential reflects perceptual suppression in monkey visual cortex. *Proc. Natl Acad. Sci. USA* **103**, 17 507–17 512. (doi:10.1073/pnas.0604673103)

11 Barraclough, N. E., Keith, R. H., Xiao, D.-K., Oram M. W. & Perrett, D. I. 2009 Visual adaptation to goal-directed hand actions. *J. Cogn. Neurosci.* **21**, 1806–1820. (doi:10.1162/jocn.2008.21145)

12 Benton, C. P., Etchells, P. J., Porter, G., Clark, A. P., Penton-Voak, I. S. & Nikolov, S. G. 2007 Turning the other cheek: the viewpoint dependence of facial expression after-effects. *Proc. R. Soc. B* **274**, 2131–2137. (doi:10.1098/rspb.2007.0473)

13 Benton, C. P., Jennings, S. J. & Chatting, D. J. 2006 Viewpoint dependence in adaptation to facial identity. *Vision Res.* **46**, 3313–3325. (doi:10.1016/j.visres.2006.06.002)

14 Lawson, R. P., Clifford, C. W. G. & Calder, A. J. 2009 The visual representation of human body orientation revealed by adaptation. *Psychol. Sci.* **20**, 363–371. (doi:10.1111/j.1467-9280.2009.02301.x)

15 Gross, C. G., Rocha-Mirana, C. E. & Bender, D. B. 1972 Visual properties of neurons in the inferotemporal cortex of the macaque. *J. Neurophysiol.* **35**, 96–111.

16 Afraz, S.-R., Kiani, R. & Esteky, H. 2006 Microstimulation of inferotemporal cortex influences face categorization. *Nature* **422**, 692–695. (doi:10.1038/nature04982)

17 Harries, M. H. & Perrett, D. I. 1991 Modular organisation of face processing in temporal cortex: physiological evidence and possible anatomical correlates. *J. Cogn. Neurosci.* **3**, 9–24. (doi:10.1162/jocn.1991.3.1.9)

18 Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D. & Jeeves, M. A. 1984 Neurons responsive to faces in the temporal cortex: studies of functional organisation, sensitivity to identity and relation to perception. *Hum. Neurobiol.* **3**, 197–208.

19 Tanaka, K., Saito, H., Fukada, Y. & Moriya, M. 1991 Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *J. Neurophysiol.* **66**, 170–189.

20 Tsao, D. Y., Freiwald, W. A., Tootell, R. B. H. & Livingstone, M. S. 2006 A cortical region consisting entirely of face-selective cells. *Science* **311**, 670–674. (doi:10.1126/science.1119983)

21 Wang, G., Tanaka, K. & Tanifuji, M. 1996 Optical imaging of functional organization in the monkey inferotemporal cortex. *Science* **272**, 1665–1668. (doi:10.1126/science.272.5268.1665)

22 Bruce, C. J., Desimone, R. & Gross, C. G. 1981 Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *J. Neurophysiol.* **46**, 369–384.

23 Perrett, D. I., Hietanen, J. K., Oram, M. W. & Benson P. J. 1992 Organization and functions of cells responsive to faces in the temporal cortex. *Phil. Trans. R. Soc. Lond. B* **335**, 23–30. (doi:10.1098/rstb.1992.0003)

24 Földiák, P., Xiao, D.-K., Keysers, C., Edwards, R. & Perrett, D. I. 2004 Rapid serial visual presentation for the determination of neural selectivity in area STSa. *Prog. Brain Res.* **144**, 107–116. (doi:10.1016/S0079-6123(03)14407-X)

25 Seltzer, B. & Pandya, D. N. 1978 Afferent cortical connections and architectonics of the superior temporal sulcus and surrounding cortex. *Brain Res.* **149**, 1–24. (doi:10.1016/0006-8993(78)90584-X)

26 Perrett, D. I., Xiao, D.-K., Barraclough, N. E., Keysers C. & Oram, M. W. 2009 Seeing the future: natural image sequences produce 'anticipatory' neuronal activity and bias perceptual report. *Q. J. Exp. Psychol.* **62**, 2081–2104. (doi:10.1080/17470210902959279)

27 Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C. & Fried, I. 2005 Invariant visual representation by single neurons in the human brain. *Nature* **435**, 1102–1107. (doi:10.1038/nature03687)

28 Perrett, D. I., Oram, M. W., Harries, M. H., Bevan, R., Hietanen, J. K., Benson, P. J. & Thomas, S. 1991 Viewer-centred and object-centred coding of heads in the macaque temporal cortex. *Exp. Brain Res.* **86**, 159–173. (doi:10.1007/BF00231050)

29 Földiák, P. 2001 Stimulus optimisation in primary visual cortex. *Neurocomputing* **38**, 1217–1222. (doi:10.1016/S0925-2312(01)00570-7)

30 Kayaert, G., Biederman, I. & Vogles, R. 2005 Shape tuning in macaque inferior temporal cortex. *Cereb. Cortex* **15**, 1308–1321. (doi:10.1093/cercor/bhi014)

31 Mazer, J. A. 2008 So many pixels, so little time. *Nat. Neurosci.* **11**, 1243–1244. (doi:10.1038/nn1108-1243)

32 Lewi, J., Butera, R. & Paninski, L. 2009 Sequential optimal design of neurophysiology experiments. *Neural Comput.* **21**, 619–687. (doi:10.1162/neco.2008.08-07-594)

33 Spiridon, M., Fischl, B. & Kanwisher, N. 2006 Location and spatial profile of category-specific regions in human extrastriate cortex. *Hum. Brain Mapp.* **27**, 77–89. (doi:10.1002/hbm.20169)

34 Haxby, J. V., Hoffman, E. A. & Gobbini, M. I. 2000 The distributed human neural system for face perception. *Trends Cogn. Sci.* **4**, 223–233. (doi:10.1016/S1364-6613(00)01482-0)

35 Oram, M. W. & Perrett, D. I. 1992 Time course of neural responses discriminating views of the face and head. *J. Neurophysiol.* **68**, 70–84.

36 Oram, M. W., Perrett, D. I. & Hietanen, J. K. 1993 Directional tuning of motion-sensitive cells in the anterior superior temporal polysensory area of the macaque. *Exp. Brain Res.* **97**, 274–294. (doi:10.1007/BF00228696)

37 Oram, M. W., Perrett, D. I., Wachsmuth, E. & Emery, N. J. 1996 Motion and visual coding of object parts. In *Cognitive Neuroscience Society Conf.*, San Francisco, CA, April 1996.

38 Perrett, D. I., Rolls, E. T. & Caan, W. 1982 Visual neurons sensitive to faces in the monkey temporal cortex. *Exp. Brain Res.* **47**, 329–342. (doi:10.1007/BF00239352)

39 Keysers, C., Xiao, D. K., Földiák, P. & Perrett, D. I. 2001 The speed of sight. *J. Cogn. Neurosci.* **13**, 90–101. (doi:10.1162/089892901564199)

40 Kiani, R., Esteky, H. & Tanaka, K. 2005 Differences in onset latency of macaque inferotemporal neural responses to primate and non-primate faces. *J. Neurophysiol.* **94**, 1587–1596. (doi:10.1152/jn.00540.2004)

41 Perrett, D. I., Oram, M. W. & Ashbridge, E. 1998 Evidence accumulation in cell populations responsive to faces: an account of generalisation of recognition without mental transformations. *Cognition* **67**, 111–145. (doi:10.1016/S0010-0277(98)00015-8)

42 Tamura, H. & Tanaka, K. 2001 Visual response properties of cells in the ventral and dorsal parts of the macaque inferotemporal cortex. *Cereb. Cortex* **11**, 384–399. (doi:10.1093/cercor/11.5.384)

43 Nelissen, K., Vanduffel, W. & Orban, G. A. 2006 Charting the lower superior temporal region, a new motion-sensitive region in monkey superior temporal sulcus. *J. Neurosci.* **26**, 5929–5947. (doi:10.1523/JNEUROSCI.0824-06.2006)

44 Saleem, K. S., Suzuki, W., Tanaka, K. & Hashikawa, T. 2000 Connections between anterior inferotemporal cortex and superior temporal sulcus regions in the macaque monkey. *J. Neurosci.* **20**, 5083–5101.

45 Seltzer, B. & Pandya, D. N. 1994 Parietal, temporal and occipital projections to cortex of the superior temporal sulcus in the rhesus monkey: a retrograde tracer study. *J. Comp. Neurol.* **15**, 445–463. (doi:10.1002/cne.903430308)

46 Leonard, C. M., Rolls, E. T., Wilson, F. A. W. & Baylis, G. C. 1985 Neurons in the amygdala of the monkey with responses selective for faces. *Behav. Brain Res.* **15**, 159–176. (doi:10.1016/0166-4328(85)90062-2)

47 O'Scalaidhe, S. P., Wilson, F. A. W. & Goldman-Rakic, P. S. 1999 Face-selective neurons during passive viewing and working memory performance of rhesus monkeys: evidence for intrinsic specialization of neural coding. *Cereb. Cortex* **9**, 459–475. (doi:10.1093/cercor/9.5.459)

48 Kreiman, G., Koch, C. & Fried, I. 2000 Category-specific visual responses of single neurons in the human medial temporal lobe. *Nat. Neurosci.* **3**, 946–953. (doi:10.1038/78868)

49 Eifuku, S., De Souza, W. C., Tamura, R., Nishijo, H. & Ono, T. 2004 Neuronal correlates of face identification in the monkey anterior temporal cortical areas. *J. Neurophysiol.* **91**, 358–371. (doi:10.1152/jn.00198.2003)

50 Nakamura, K., Mikami, A. & Kubota, K. 1992 Activity of single neurons in the monkey amygdala during performance of a visual discrimination task. *J. Neurophysiol.* **67**, 1447–1463.

51 Rolls, E. T., Critchley, H. D., Browning, A. S. & Inoue, K. 2006 Face-selective and auditory neurons in the primate orbitofrontal cortex. *Exp. Brain Res.* **170**, 74–87. (doi:10.1007/S00221-005-0191-y)

52 Desimone, R., Albright, T. D., Gross, C. G. & Bruce C. 1984 Stimulus-selective properties of inferior temporal neurons in the macaque. *J. Neurosci.* **8**, 2051–2062.

53 Tovee, M. J., Rolls, E. T. & Azzopardi, P. 1994 Translation invariance in the responses to faces of single neurons in temporal visual cortical areas of the alert macaque. *J. Neurophysiol.* **72**, 1049–1060.

54 Afraz, S.-R. & Cavanagh, P. 2008 Retinotopy of the face aftereffect. *Vision Res.* **48**, 42–54. (doi:10.1016/j.visres.2007.10.028)

55 Leopold, D. A., O'Toole, A. J., Vetter, T. & Blanz, V. 2001 Prototype-referenced shape encoding revealed by high-level after-effects. *Nat. Neurosci.* **4**, 89–94. (doi:10.1038/82947)

56 Rolls, E. T. & Tovee, M. J. 1995 The responses of single neurons in the temporal visual cortical areas of the macaque when more than one stimulus is present in the receptive field. *Exp. Brain Res.* **103**, 409–420. (doi:10.1007/BF00241500)

57 Tovee, M. J., Rolls, E. T. & Azzopardi, P. 1994 Translation invariance in the responses to faces of single neurons in temporal visual cortical areas of the alert macaque. *J. Neurophysiol.* **72**, 1049–1060.

58 Giese, M. A. 2009 Computation mechanisms of the recognition of biological motion and actions: theories and experiments. In *Perceiving bodies in action: from low to high level mechanisms, Leipzig, Germany, 4 December 2009.*

59 Andrews, T. J. & Ewbank, M. P. 2004 Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *Neuroimage* **23**, 905–913. (doi:10.1016/j.neuroimage.2004.07.060)

60 Kovacs, G., Zimmer, M., Banko, E., Harza, I., Antal A. & Vidnyanszky, Z. 2006 Electrophysiological correlates of visual adaptation to faces and body parts in humans. *Cereb. Cortex* **16**, 742–753. (doi:10.1093/cercor/bhj020)

61 Little, A. C., deBruine, L. M. & Jones, B. C. 2005 Sex-contingent face after-effects suggest distinct neural populations code male and female faces. *Proc. R. Soc. B* **272**, 2283–2287. (doi:10.1098/rspb.2005.3220)

62 Rhodes, G., Jeffery, L., Watson, T. L., Clifford, C. W. G. & Nakayama, K. 2003 Fitting the mind to the world: face adaptation and attractiveness after-effects. *Psychol. Sci.* **14**, 558–566. (doi:10.1046/j.0956-7976.2003.psci_1465.x)

63 Webster, M. A., Kaping, D., Mizokami, Y. & Duhamel P. 2004 Adaptation to natural facial categories. *Nature* **428**, 557–561. (doi:10.1038/nature02420)

64 Kohn, A. 2007 Visual adaptation: physiology, mechanisms, and functional benefits. *J. Neurophysiol.* **97**, 3155–3164. (doi:10.1152/jn.00086.2007)

65 Kohn, A. & Movshon, J. A. 2004 Adaptation changes the direction tuning of macaque MT neurons. *Nat. Neurosci.* **7**, 764–772. (doi:10.1038/nn1267)

66 Verhoef, B.-E., Kayaert, G., Franko, E., Vangeneugden, J. & Vogels, R. 2008 Stimulus similarity-contingent neural adaptation can be time and cortical area dependent. *J. Neurosci.* **28**, 10 631–10 640. (doi:10.1523/JNEUROSCI.3333-08.2008)

67 Ghisovan, N., Nemri, A., Shumikhina, S. & Molotchnikoff, S. 2009 Long adaptation reveals mostly attractive shifts of orientation tuning in cat primary visual cortex. *Neuroscience* **164**, 1274–1283. (doi:10.1016/j.neuroscience.2009.09.003)

68 Baylis, G. C. & Rolls, E. T. 1987 Responses of neurons in the inferior temporal cortex in short term and serial recognition memory tasks. *Exp. Brain Res.* **65**, 614–622. (doi:10.1007/BF00235984)

69 Brown, M. W., Wilson, F. A. W. & Riches, I. P. 1987 Neuronal evidence that inferomedial temporal cortex is more important than hippocampus in certain processes underlying recognition memory. *Brain Res.* **409**, 158–162. (doi:10.1016/0006-8993(87)90753-0)

70 Miller, E. K., Gochin, P. M. & Gross, C. G. 1991 Habituation-like decrease in the responses of neurons in inferior temporal cortex of the macaque. *Visual Neurosci.* **7**, 357–362. (doi:10.1017/S0952523800004843)

71 Miller, E. K., Li, L. & Desimone, R. 1991 A neural mechanism for working and recognition memory in inferior temporal cortex. *Science* **254**, 1377–1379. (doi:10.1126/science.1962197)

72 Rolls, E. T., Baylis, G. C., Hasselmo, M. E. & Nalwa E. 1989 The effect of learning on the face selective responses of neurons in the cortex in the superior

temporal sulcus of the monkey. *Exp. Brain Res.* **76**, 153–164. (doi:10.1007/BF00253632)

73 Li, L., Miller, E. K. & Desimone, R. 1993 The representation of stimulus familiarity in anterior inferotemporal cortex. *J. Neurophysiol.* **69**, 1918–1929.

74 Leopold, D. A., Rhodes, G., Muller, K.-M. & Jeffery, L. 2005 The dynamics of visual adaptation to faces. *Proc. R. Soc. B* **272**, 897–904. (doi:10.1098/rspb.2004.3022)

75 Grill-Spector, K., Kushnir, T., Edelman, S., Avidan G., Itzchak, Y. & Malach, R. 1999 Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* **24**, 187–203. (doi:10.1016/S0896-6273(00)80832-6)

76 Grill-Spector, K. & Malach, R. 2001 fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychol.* **107**, 293–321. (doi:10.1016/S0001-6918(01)00019-1)

77 Sawamura, H., Orban, G. & Vogels, R. 2006 Selectivity of neuronal adaptation does not match response selectivity: a single-cell study of the fMRI adaptation paradigm. *Neuron* **49**, 307–318. (doi:10.1016/j.neuron.2005.11.028)

78 Grill-Spector, K., Henson, R. & Martin, A. 2006 Repetition and the brain: neural models of stimulus specific effects. *Trends Cogn. Sci.* **10**, 14–23. (doi:10.1016/j.tics.2005.11.006)

79 Tsao, D. Y., Freiwald, W. A., Knutsen, T. A., Mandeville, J. B. & Tootell, B. H. 2003 Faces and objects in macaque cerebral cortex. *Nat. Neurosci.* **6**, 989–995. (doi:10.1038/nn1111)

80 Macknik, S. L. & Livingstone, M. S. 1998 Neural correlates of visibility and invisibility in the primate visual system. *Nat. Neurosci.* **1**, 144–149. (doi:10.1038/393)

81 Bestelmeyer, P. E. G., Jones, B. C., DeBruine, L. M., Little, A. C., Perrett, D. I., Schneider, A., Welling, L. L. M. & Conway, C. A. 2008 Sex contingent face aftereffects depend upon perceptual category rather than structural encoding. *Cognition* **107**, 353–365. (doi:10.1016/j.cognition.2007.07.018)

82 Little, A. C., DeBruine, L. M., Jones, B. C. & Waitt, C. 2007 Category contingent aftereffects for faces of different races, ages and species. *Cognition* **106**, 1537–1547. (doi:10.1016/j.cognition.2007.06.008)

83 McCullough, C. 1965 Adaptation of edge-detectors in the human visual system. *Science* **149**, 1115–1116. (doi:10.1126/science.149.3688.1115)

84 Carbon, C. C. & Leder, H. 2006 The *Mona Lisa* effect: is 'our' Lisa fame or fake? *Perception* **35**, 411–414. (doi:10.1068/p5452)

85 Carbon, C. C., Strobach, T., Langton, S. R., Harsányi G., Leder, H. & Kovács, G. 2007 Adaptation effect of highly familiar faces: immediate and long lasting. *Mem. Cognit.* **35**, 1966–1976. (doi:10.3758/BF03192929)

86 Perrett, D. I. 1981 A neurophysiological study of aspects of higher visual processing. D.Phil thesis, Department of Psychology, University of Oxford, Oxford, UK.

87 Rolls, E. T., Perrett, D. I., Caan, A. W. & Wilson, F. A. W. 1982 Neuronal responses related to visual recognition. *Brain* **105**, 611–646. (doi:10.1093/brain/105.4.611)

88 Perrett, D. I., Harries, M. H., Bevan, R., Thomas, S., Benson, P. J., Mistlin, A. J., Chitty, A. J., Hietanen, J. K. & Ortega, J. E. 1989 Frameworks of analysis for the neural representation of animate objects and actions. *J. Exp. Biol.* **146**, 87–113.

89 Perrett, D. I., Mistlin, A. J. & Chitty, A. J. 1987 Visual cells responsive to faces. *Trends Neurosci.* **10**, 358–364. (doi:10.1016/0166-2236(87)90071-3)

90 Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin A. J., Head, A. S., Milner, A. D. & Jeeves, M. A. 1985 Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proc. R. Soc. Lond. B* **223**, 293–317. (doi:10.1098/rspb.1985.0003)

91 Rolls, E. T. 1992 Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical areas. *Phil. Trans. R. Soc. Lond. B* **335**, 11–21. (doi:10.1098/rstb.1992.0002)

92 Young, M. P. & Yamane, S. 1992 Sparse population coding of faces in the inferotemporal cortex. *Science* **256**, 1327–1331. (doi:10.1126/science.1598577)

93 Rolls, E. T. & Baylis, G. C. 1986 Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the monkey. *Exp. Brain Res.* **65**, 38–48. (doi:10.1007/BF00243828)

94 Hasselmo, M. E., Rolls, E. T. & Baylis, G. C. 1989 The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behav. Brain Res.* **32**, 203–218. (doi:10.1016/S0166-4328(89)80054-3)

95 Bowers, J. S. 2009 On the biological plausibility of grandmother cells: implications for neural network theories in psychology and neuroscience. *Psychol. Rev.* **116**, 220–251. (doi:10.1037/a0014462)

96 Földiák, P. 2009 Neural coding: non-local but explicit and conceptual. *Curr. Biol.* **19**, R904–R906. (doi:10.1016/j.cub.2009.08.020)

97 Quiroga, R. Q., Kreiman, G., Koch, C. & Fried, I. 2008 Sparse but not grandmother-cell coding in the medial temporal lobe. *Trends Cogn. Sci.* **12**, 2081–2104. (doi:10.1016/j.tics.2007.12.003)

98 Waydo, S., Kraskov, A., Quiroga, R. Q., Fried, I. & Koch, C. 2006 Sparse representation in the human medial temporal lobe. *J. Neurosci.* **26**, 10 232–10 234. (doi:10.1523/JNEUROSCI.2101-06.2006)

99 Sugase, Y., Yamane, S., Ueno, S. & Kawano, K. 1999 Global and fine information coded by single neurons in the temporal visual cortex. *Nature* **400**, 869–873. (doi:10.1038/23703)

100 Rhodes, G., Maloney, L. T., Turner, J. & Ewing, L. 2007 Adaptive face coding and discrimination around the average face. *Vision Res.* **47**, 974–989. (doi:10.1016/j.visres.2006.12.010)

101 Leopold, D. A., Bondar, I. V. & Giese, M. A. 2006 Norm-based face encoding by single neurons in the monkey inferotemporal cortex. *Nature* **442**, 572–575.

102 Freiwald, W. A., Tsao, D. Y. & Livingstone, M. S. 2009 A face feature space in the macaque temporal lobe. *Nat. Neurosci.* **9**, 1187–1196. (doi:10.1038/nn.2363)

103 Rhodes, G. & Jeffery, L. 2006 Fitting the mind to the world: face adaptation and attractiveness aftereffects. *Psychol. Sci.* **14**, 558–566. (doi:10.1046/j.0956-7976.2003.psci_1465.x)

104 Susilo, T., McKone, E. & Edwards, M. 2010 What shape are the neural response functions underlying opponent coding in face space? A psychophysical investigation. *Vision Res.* **50**, 300–314. (doi:10.1016/j.visres.2009.11.016)

105 Calder, A. J., Jenkins, R., Cassel, A. & Clifford, C. W. G. 2008 Visual representation of eye gaze is coded by a non-opponent multichannel system. *J. Exp. Psychol. Gen.* **137**, 244–261. (doi:10.1037/0096-3445.137.2.244)

106 Bestelmeyer, P. E. G., Jones, B. C., DeBruine, L. M., Little, A. C. & Welling, L. L. M. 2009 Face aftereffects

suggest interdependent processing of expression and sex and of expression and race. *Vis. Cogn.* **18**, 255–274. (doi:10.1080/13506280802708024)

107 de Heering, A. & Rossion, B. 2008 Prolonged visual experience in adulthood modulates holistic face perception. *PLoS ONE* **3**, e2317. (doi:10.1371/journal.pone.0002317)

108 Dufour, V., Pascalis, O. & Petit, O. 2006 Face processing limitation to own species in primates: a comparative study in brown capuchins, Tonkean macaques and humans. *Behav. Processes* **73**, 107–113. (doi:10.1016/j.beproc.2006.04.006)

109 Ghazanfar, A. A. & Logothetis, N. K. 2003 Facial expressions linked to monkey calls. *Nature* **423**, 937–938. (doi:10.1038/423937a)

110 Partan, S. R. 2002 Single and multichannel signal composition: facial expressions and vocalizations of rhesus macaques (*Macaca Mulatta*). *Behaviour* **139**, 993–1027. (doi:10.1163/15685390260337877)

111 Parr, L. A., Waller, B. M., Vick, S. J. & Bard, K. A. 2007 Classifying chimpanzee facial expressions using muscle action. *Emotion* **7**, 172–181. (doi:10.1037/1528-3542.7.1.172)

112 Gothard, K. M., Battaglia, F. P., Erickson, C. A., Spitler, K. M. & Amaral, D. G. 2007 Neural responses to facial expression and face identity in the monkey amygdala. *J. Neurophysiol.* **97**, 1671–1683. (doi:10.1152/jn.00714.2006)

113 Ferrari, P. F., Gallese, V., Rizzolatti, G. & Fogassi, L. 2003 Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *Eur. J. Neurosci.* **17**, 1703–1714. (doi:10.1046/j.1460-9568.2003.02601.x)

114 Perrett, D. I. & Mistlin, A. J. 1990 Perception of facial attributes. In *Comparative perception*, vol. II: compex signals (eds W. C. Stebbins & M. A. Berkley), pp. 187–215. New York, NY: John Wiley.

115 di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V. & Rizzolatti, G. 1992 Understanding motor events: a neurophysiological study. *Exp. Brain Res.* **91**, 176–180.

116 Rizzolatti, G., Fadiga, L., Gallese, V. & Fogassi, L. 1996 Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.* **3**, 131–141. (doi:10.1016/0926-6410(95)00038-0)

117 Sumby, W. H. & Pollack, I. 1954 Visual contribution to speech intelligibility in noise. *J. Acoust. Soc. Am.* **26**, 212–215. (doi:10.1121/1.1907309)

118 Barraclough, N. E., Xiao, D.-K., Oram, M. W. & Perrett, D. I. 2005 Integration of visual and auditory information by STS neurons responsive to the sight of actions. *J. Cogn. Neurosci.* **17**, 377–391. (doi:10.1162/0898929053279586)

119 Barraclough, N. E., Xiao, D.-K. & Perrett, D. I. 2004 Understanding audio-visual communication at the level of the single neuron. *Kagaku* (*Science*) **74**, 891–893.

120 Barraclough, N. E., Xiao, D.-K., Oram, M. W. & Perrett, D. I. 2004 Deriving specific meaning through integration of vision and sound by neurons in primate superior temporal sulcus (STS). In 5th *Int. Multisensory Research Forum, Barcelona, Spain, 2–5 June 2004.*

121 Ghazanfar, A. A., Maier, J. X., Hoffmann, K. L. & Logothetis, N. K. 2005 Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *J. Neurosci.* **25**, 5004–5012. (doi:10.1523/JNEUROSCI.0799-05.2005)

122 Dolan, R. J., Morris, J. S. & de Gelder, B. 2001 Cross-modal binding of fear in voice and face. *Proc. Natl Acad. Sci. USA* **98**, 10 006–10 010. (doi:10.1073/pnas.171288598)