THE ROYAL
SOCIETY

The temporal resolution of neural codes: does response latency have a unique role?

M. W. Oram[*], **D. Xiao, B. Dritschel and K. R. Payne**

School of Psychology, University of St Andrews, St Andrews, Fife KY16 9JU, UK

This article reviews the nature of the neural code in non-human primate cortex and assesses the potential for neurons to carry two or more signals simultaneously. Neurophysiological recordings from visual and motor systems indicate that the evidence for a role for precisely timed spikes relative to other spike times (*ca.* 1–10 ms resolution) is inconclusive. This indicates that the visual system does not carry a signal that identifies whether the responses were elicited when the stimulus was attended or not. Simulations show that the absence of such a signal reduces, but does not eliminate, the increased discrimination between stimuli that are attended compared with when the stimuli are unattended. The increased accuracy asymptotes with increased gain control, indicating limited benefit from increasing attention. The absence of a signal identifying the attentional state under which stimuli were viewed can produce the greatest discrimination between attended and unattended stimuli. Furthermore, the greatest reduction in discrimination errors occurs for a limited range of gain control, again indicating that attention effects are limited. By contrast to precisely timed patterns of spikes where the timing is relative to other spikes, response latency provides a fine temporal resolution signal (*ca.* 10 ms resolution) that carries information that is unavailable from coarse temporal response measures. Changes in response latency and changes in response magnitude can give rise to different predictions for the patterns of reaction times. The predictions are verified, and it is shown that the standard method for distinguishing executive and slave processes is only valid if the representations of interest, as evidenced by the neural code, are known. Overall, the data indicate that the signalling evident in neural signals is restricted to the spike count and the precise times of spikes relative to stimulus onset (response latency). These coding issues have implications for our understanding of cognitive models of attention and the roles of executive and slave systems.

Keywords: neural codes; synchrony; response latency; visual system; attention; central executive

1. INTRODUCTION

Understanding brain function requires a precise knowledge of the information present at each stage of processing, how this information is encoded (the neural code) and how different signals associated with different processing stages are combined (integration of neural codes). While the general processing of many cortical and subcortical brain areas in both human and non-human primates is known, the way in which neurons encode inputs and process information to generate output signals is still the subject of intense debate. The neural code underlies the way in which the brain operates and therefore helps constrain the ways in which we think the brain can function by defining the mechanisms by which the brain transmits information. The processing power of the brain is restricted by the capacity of the neural responses to transmit information. If the information capacity of a single channel (neural code) is limited, then additional neural codes (signals) would increase the total information-processing potential of the brain without requiring an increase in the number of neurons.

If neural responses carry different types of information (e.g. stimulus colour and stimulus shape) using different forms of encoding (e.g. spike count and response latency), then these types of information can be processed together ('is a red square present?') or processed separately ('is there a square present?' and 'is a red object present?'). Of current interest is the hypothesis that dynamic links between different neural populations are used to signal particular associations (Singer & Gray 1995). For example, imagine four neural populations, one of which signals the presence of the colour red, another the presence of blue, while the third and fourth populations signal squares and triangles, respectively. Simultaneous presentation of a red square and a blue triangle would activate all four neural populations. However, the decoding is not simple: does the 'red object' belong with the 'square' or the 'triangle'? If an additional code were available that linked these responses in an appropriate way, then the system would be capable of distinguishing 'red square' and 'blue triangle' from 'red triangle' and 'blue square'. Such linking or binding of features (see Singer & Gray 1995; von der Malsburg 1995; for a review) requires that there are separate codes for the attributes (red, blue, square, triangle) and the links (red square, blue triangle etc.).

The traditional method for investigating a signal that is carried by a neural response has been to count the number of spikes elicited within a relatively long time window (the

---

[*] Author for correspondence (mwo@st-andrews.ac.uk).

One contribution of 14 to a Discussion Meeting Issue 'The physiology of cognitive processes'.

DOI 10.1098/rstb.2002.1113

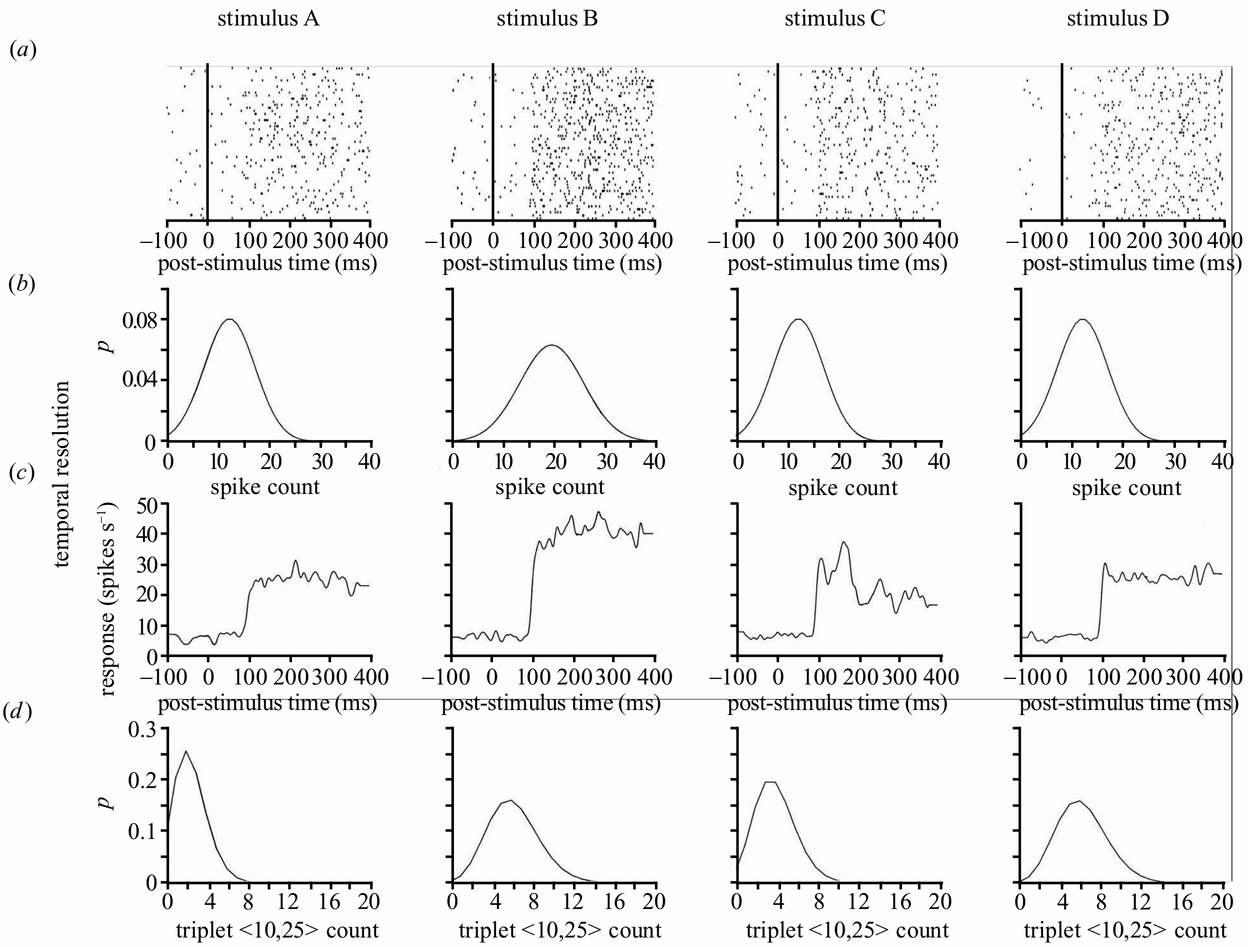

Figure 1. Schematic showing the multiple neural signals. (*a*) Raw neural data. The possible responses to four stimuli (A–D) are shown in rastergram form (each row of dots is the response from a single trial with each dot representing the time when a spike occurred). (*b*) Coarse temporal resolution signals. The probability or relative frequency of each spike count (measured over the period 0–400 ms post-stimulus onset) being elicited by each stimulus (the variance of spike count is twice the mean). The spike-count distributions are identical for stimuli A, C and D (mean of 12 spikes, variance of 24 spikes^2), but the spike-count distribution elicited by presentations of stimulus B is different (mean of 21 spikes, variance of 42 spikes^2). Thus, the spike count discriminates between input stimulus B and the other stimuli (i.e. carries stimulus-related information).
(*c*) Medium temporal resolution signals. The spike-density function (firing rate as a function of time) for each of the stimuli (temporal resolution of 5 ms). The shape of the spike-density function of the responses to stimulus C is different from those of the responses to stimuli A, B and D (after adjusting for the changing spike count in the case of stimulus B). Therefore, the intermediate temporal resolution code can carry information unavailable from the spike count. (*d*) Fine temporal resolution signals (distributions of precisely timed triplets). The fine temporal measure of the probability or relative frequency of observing different numbers of triplet ⟨10,25⟩ in the response to each presentation of stimuli A–D. Triplet ⟨10,25⟩ is a triplet of spikes with intervals of 10 and 25 ms. The differences in the distributions of triplet ⟨10,25⟩ in the responses to stimuli A, B and C can be attributed to changes in the spike-count distributions (A versus B, B versus D) or the spike-density function (A versus C, C versus D). The distributions of triplet ⟨10,25⟩ differ for the responses to stimuli A and D and this difference is not a reflection of differences in either the spike count or spike-density function. Therefore, the fine (1–20 ms) temporal resolution code can carry information unavailable from either the coarse or intermediate resolution code (the spike count and spike-density function shape, respectively). Substantial evidence indicates that the mid-range temporal measures of neural responses carry information that is unavailable for coarse temporal measures. Recently it has been speculated that the fine temporal measures of responses (lower right) may carry yet more information.

spike count), typically measured in hundreds of milliseconds. As the spikes from a neuron are indistinguishable from each other, the spike-count measure captures all of the available information in the neural activity using this coarse temporal resolution measure. For the neural activity to carry more information, therefore, the encoding must utilize a different temporal resolution.

Figure 1 shows schematically how multiple neural codes of different temporal resolutions could be encoded in the responses of a single neuron. Responses of an imaginary neuron to four stimuli (A–D) are shown as rastergrams (*a*). Figure 1*b*–*d* depict three neural codes of different temporal resolution. The spike-count distributions (coarse temporal resolution) are identical for stimuli A, C and D, and different from that of stimulus B. Thus, the spike count conveys stimulus related information because it distinguishes stimulus B from the other stimuli. Figure 1*c*,*d* depicts neural codes of intermediate and fine temporal resolution. The spike-density functions of stimuli A and D (firing rate as a function of time) are statistically indistinguishable. Although the spike-density functions of the responses elicited by stimuli A and B are different, the difference reflects the difference in the coarse temporal measure of spike count (i.e. one is a scaled version of the other). The spike-density function elicited by stimulus C, however, has a different shape from that of the responses to stimuli A, B and D. Thus, the intermediate temporal resolution neural signals can carry stimulus-related information (distinguishing stimulus C from the others) that is unavailable from the coarse temporal measure of spike count. Differences in the probability or frequency distributions of the number of triplets of spikes defined by intervals of 10 and 25 ms exist in the responses elicited by each stimulus. However, only the differences between the distributions of the triplets in the responses elicited to stimulus D and the responses to the other stimuli cannot be attributed to differences in spike-count distributions or spike-density function shape. Thus, fine temporal resolution response measures have the potential to carry information unavailable from either the coarse or intermediate temporal resolution signals.

A stimulus-elicited response could convey information using the three different temporal resolutions as depicted in figure 1. Information about the overall stimulus class could be carried by spike count (measured over hundreds of milliseconds; stimulus A versus stimulus B in figure 1). Details of the particular stimulus could be determined from intermediate temporal resolution signals (the 'shape' of the spike-density function, stimulus A versus stimulus C in figure 1, with a temporal resolution of 20–50 Hz bandwidth; Optican & Richmond 1987; Richmond *et al.* 1987). Finally, a third type of information related to, for example, whether the stimulus was attended or not, could be encoded by the fine temporal resolution signals (stimulus A versus stimulus D in figure 1; Lestienne & Strehler 1987; Abeles *et al.* 1993; Lestienne & Tuckwell 1998; Prut *et al.* 1998). We review in this article the evidence for such coding schemes and discuss the implications of the findings both for our understanding of neurophysiological encoding of information and for higher-level cognitive processes.

In visual-system neural responses, both spike count and intermediate resolution temporal variation in firing rate (20–50 ms precision) carry information that allow stimuli to be differentiated (Optican & Richmond 1987; Richmond *et al.* 1987; Tovee *et al.* 1993). Importantly, the temporal variation in firing rate carries information that is not available from spike count (Richmond *et al.* 1987; Richmond & Optican 1990; Richmond *et al.* 1990; McClurkin *et al.* 1991*a–c*; Eskandar *et al.* 1992*a*,*b*; Heller *et al.* 1995). Thus, different neural codes exist at coarse and intermediate temporal resolutions. It is also natural to wonder whether fine temporal characteristics (*ca.* 1–10 ms precision) carry information unavailable from either the coarse or intermediate temporal resolution signals (for reviews, see von der Malsburg & Schneider 1986; Abeles 1991; Engel *et al.* 1992; Softky & Koch 1993; Singer & Gray 1995; von der Malsburg 1995; Shadlen & Newsome 1998).

The potential for multiple neural coding to increase the processing capacity of a brain with a fixed number of neurons compared with a unitary coding scheme gives rise to the following questions.

(i) What criteria must be met to establish whether or not neural encoding utilizes more than one code?
(ii) At what temporal resolution does the brain operate and to what extent do neurons show multiple signals?
(iii) What are the implications of the neural code for higher cognitive processes (i.e. is there evidence that the brain might use neural codes of different temporal resolution)?

In the present article we review current experimental results that shed light on these issues.

2. PRECISELY TIMED SPIKE PATTERNS RELATIVE TO OTHER SPIKES

Precisely timed spike patterns relative to other spikes within and between neural responses have been postulated to play an important role in functions such as selection and coordination of motor output (Abeles 1991; Riehle *et al.* 1997; Villa & Bajo 1997; Prut *et al.* 1998) and linking or binding of different attributes of visual stimuli (von der Malsburg & Schneider 1986; Singer & Gray 1995; von der Malsburg 1995). Central to the proposal for a special or unique role of precisely timed spike patterns (*ca.* 1 ms precision) in brain functioning is the requirement that such patterns carry information that is unavailable from the coarse and intermediate temporal response measures. For example, if the number of synchronous spikes between the neural responses can be controlled independently of the spike count, then the number of synchronous spikes could provide information that is unavailable from the spike count. Other precisely timed spike patterns that have been investigated include oscillatory activity between pairs of neurons (for reviews, see Engel *et al.* 1992; Singer & Gray 1995) and more complex patterns of precisely timed spikes such as triplets of spikes with precisely timed intervals (the 'synfire chains' of Abeles (1991), see figure 1; Lestienne & Strehler 1987; Abeles *et al.* 1993; Prut *et al.* 1998; Lestienne & Tuckwell 1998; Oram *et al.* 1999*a*,*b*).

Precisely timed spike patterns can only carry information unavailable from spike count and slow variations in firing rate (coarse and intermediate temporal codes, respectively) if the precise spike patterns are controlled rather than occurring by chance. Complex relationships exist between the different temporal measures of neural responses, requiring the use of statistical models to determine the potential significance of precisely timed spikes (Dayhoff & Gerstein 1983*a*,*b*; Lestienne & Strehler 1987; Abeles & Gerstein 1988; Aertsen *et al.* 1989; Abeles 1991; Abeles *et al.* 1993; Vaadia *et al.* 1995; Lestienne & Tuckwell 1998). Several studies have indicated that the observed numbers of precisely timed spike patterns in a variety of neural systems occur more frequently than predicted from stochastic models (Lestienne & Strehler 1987; Palm *et al.* 1988; Abeles 1991; Singer & Gray 1995; Prut *et al.* 1998; Oram *et al.* 2001). However, the assumptions underlying the analysis have pronounced effects on the number and type of precisely timed spike patterns expected by chance (figure 2, and see Brody 1999*a–c*; Oram *et al.* 1999*a*,*b*, 2001; Prut *et al.* 1998; Richmond *et al.* 1999; Treves *et al.* 1999; Baker & Lemon 2000).

The major determinants of the number of precisely timed spike patterns expected by chance are the coarse

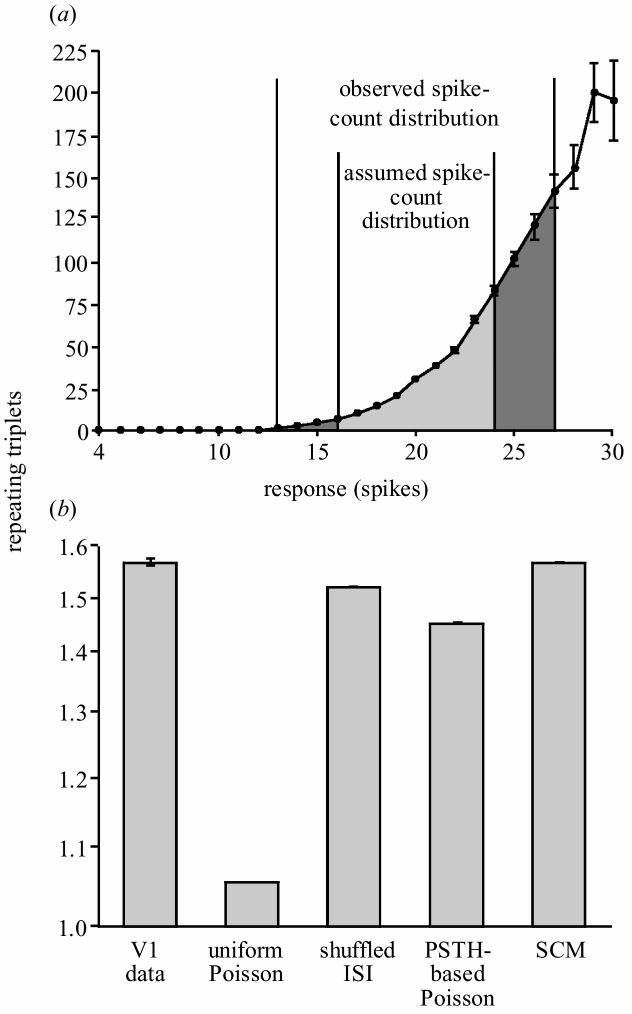

Figure 2. The effect of different assumptions on the expected number of precisely timed spike patterns. (*a*) The number of precisely timed repeating triplets in a trial (mean ± s.e.m.) is a nonlinear function of the spike count. The light shaded area indicates the number of precisely timed spike patterns expected in the responses to a stimulus with the indicated assumed spike-count distribution. If the observed spike-count distribution is larger, the total shaded area (light and dark) represents the number of precisely timed spike patterns expected by chance. Thus, if the assumed distribution of spikes is smaller than the observed distribution, the chosen model will underestimate the number of precisely timed spike patterns expected by chance. (*b*) Assuming a Poisson spike-count distribution will underestimate the number of precisely timed spike patterns expected by chance (V1 data versus Poisson) as spike-count distributions of cortical neurons are super-Poisson. Correcting the coarse temporal response measure (the spike count) but ignoring the mid-range temporal measures (PSTH envelope) also underestimates the number of precisely timed spike patterns expected by chance (shuffled ISI versus V1 data). Correcting for the mid-range temporal measures but not the coarse measures also underestimates the number of precisely timed spike patterns (PSTH-based Poisson versus V1 data). When the coarse temporal response measures (spike-count distribution) and the mid-range temporal measures (PSTH envelope) are matched, the number of precisely timed spike patterns is expected by chance (SCM versus V1 data) (adapted from Oram *et al.* 1999*b*).

and intermediate temporal resolution aspects of the response (spike-count frequency distribution and the PSTH shape, respectively). Precisely timed patterns of spikes reflect temporal relationships or correlation within and between responses. At a sufficiently fine temporal resolution, neural responses can be described as binary events (spike or not) with a low probability of a spike occurring. Very short time bins will have a spike-count Poisson distribution. Assuming independent bins over an extended period, the distribution of spike counts of responses will also be Poisson. Whenever the observed spike-count frequency distribution over repeated trials deviates from a Poisson distribution there must be covariation between different periods of the response. The misestimation of the correlation between periods within a response (as measured by precisely timed spike patterns) due to assuming any response property is increased with measures obtained from techniques that measure the summed activity across many neurons (e.g. local field potentials and more global techniques such as electroencephalogram, functional magnetic resonance imaging and magnetoencephalogram). The accurate assessment of the statistical significance of precisely timed spike patterns is therefore restricted to data collected from multiple single neurons.

Stochastic models that assume that the spike counts follow a Poisson distribution predict fewer precisely timed spike patterns than seen in neural data (Lestienne & Strehler 1987; Abeles & Gerstein 1988; Aertsen *et al.* 1989; Abeles *et al.* 1993; Lestienne & Tuckwell 1998). As non-Poisson distributions of spike count have been reported in responses of neurons in the retina, lateral geniculate nucleus, V1, TE, parietal and frontal lobes (Tolhurst *et al.* 1983; Levine & Troy 1986; Bradley *et al.* 1987; Vogels *et al.* 1989; Snowden *et al.* 1992; Britten *et al.* 1993; Victor & Purpura 1996; Berry *et al.* 1997; Reich *et al.* 1997; Baddeley *et al.* 1997; Berry & Meister 1998; Gershon *et al.* 1998; Buracas *et al.* 1998; Lee *et al.* 1998; Oram *et al.* 1999*a*,*b*, 2001), we introduced a new stochastic model (the SCM model; Oram *et al.* 1999*a*,*b*, 2001). The SCM model extended earlier models based on nonhomogenous Poisson processes by replacing the assumed Poisson distribution of spike counts with the observed frequency distribution of spike counts. Other models match the frequency distribution of spike counts but disrupt other measures that are known to influence the numbers of precisely timed spike patterns that are expected by chance. For example, randomly reordering the interspike intervals within a train maintains the interval distribution but not the PSTH; exchanging spikes between trains maintains the PSTH but not the interval distribution. Jittering the times of spikes in trains retains both the PSTH response and the interval distribution approximately but not exactly. All these model types underestimate the number of precisely timed spike patterns randomly expected (Richmond *et al.* 1999). When models incorporate the observed spike-count distribution, PSTH shape and interspike intervals, the numbers and types of precisely timed patterns seen in neural data from visual and motor system responses are almost exactly predicted (Oram *et al.* 1999*b*, 2001; Richmond *et al.* 1999; Baker & Lemon 2000).

A further difficulty in assessing the potential for precisely timed spikes to carry information that is unavailable from coarse temporal response measures arises from the correlation structure of the coarse temporal response measures. The information content of populations of neural responses is dependent on the correlation between the responses of the constituent neurons (Oram *et al.* 1998; Zhang *et al.* 1998; Abbott & Dayan 1999; Panzeri *et al.* 1999; Treves *et al.* 1999). Additionally, the correlation of spike counts assessed over seconds between neurons has a substantial influence on the fine temporal structures (*ca.* 1–5 ms precision) between responses of different neurons (Brody 1999*a–c*; Oram *et al.* 2001). For example, synchronous spikes between responses of different motor cortical neurons carry information about the timing and direction of limb movement that is unavailable from spike count if one assumes statistical independence between the engaged neurons (Abeles 1991; Vaadia *et al.* 1995; Reich *et al.* 1997; Hatsopoulos *et al.* 1998; Maynard *et al.* 1999). When the observed statistical relationships of spike counts are incorporated into the analysis, the information carried by synchronous spikes is found to be only a small fraction of that available from the coarse temporal response measures and does not therefore form a neural code that carries information that is unavailable from the spike count (Baker & Lemon 2000; Baker *et al.* 2001; Oram *et al.* 2001). Similarly, the incorporation of all the observed coarse and intermediate temporal resolution response properties reveals that the stimulus-related information carried by the precise times of spikes relative to the times of other spikes in the neural responses of the lateral geniculate nucleus and primary and inferotemporal cortices to static visual stimuli is available from the spike count assessed over hundreds of milliseconds (Oram *et al.* 1999*a*,*b*; Richmond *et al.* 1999).

Observing that a particular precisely timed spike pattern occurs more frequently than randomly expected does not imply that the precisely timed code conveys information that is unavailable from any other temporal measures of the neural response. For example, the mean number of synchronous spikes predicted from the SCM model underestimates the observed mean number of synchronous spikes between the responses of pairs of motor cortical neurons. Thus, there is an excess of synchrony despite the SCM model incorporating all the observed coarse and intermediate temporal characteristics both within and between the responses of the constituent neurons. Furthermore, the amount by which the observed mean number of synchronous spikes exceeded the predicted mean varied with the arm-movement direction (figure 3 and Oram *et al.* 2001; see also Baker *et al.* 2001). However, when the predicted and observed mean numbers of synchronous spikes are plotted against each other, the observed mean numbers of synchronous spikes are linearly related to the numbers randomly expected (Oram *et al.* 2001). For example, if a mean of two synchronous spikes per trial was randomly expected, a mean of three was observed; whereas if a mean of six synchronous spikes was expected, a mean of nine synchronous spikes was observed. In other words, knowing the number of synchronous spikes found in the responses associated with one arm-movement direction enables the accurate prediction of the number of synchronous spikes from the spike count and the PSTH for all other conditions. Given this predictability, it is not surprising that the number of synchronous spikes carries information only available from the spike counts and the correlation of the spike counts (Oram *et al.* 2001).

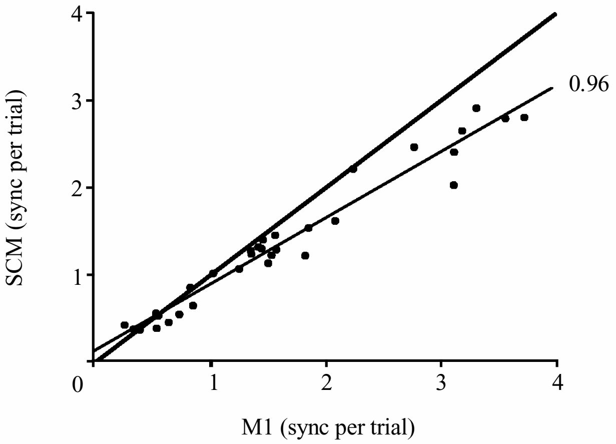

Figure 3. Excess of synchrony does not imply a separable neural code. Sample plot of the number of synchronous spikes expected by chance (SCM model) and the number of synchronous spikes observed in data from simultaneously recorded pairs of neurons in motor cortex for different arm-movement directions. The SCM model underestimates the number of synchronous spikes expected by chance (dots lie to the right of the thick equality line). Regression of the model prediction on the observed number indicates that 96% ($R^2 = 0.96$) of the variability in number of synchronous spikes is explained by the stochastic SCM model. Given that virtually all the variability in numbers of synchronous spikes is explained, it is not surprising that synchronous spikes only carried information available from the coarse and mid-range temporal response measures (adapted from Oram *et al.* 2001).

(a) *Interim summary: spike times relative to the times of other spikes*

Precisely timed spikes have been postulated to form a neural code at fine temporal resolution that carries information that is unavailable from the coarse and intermediate temporal resolution codes. All the coarse and intermediate temporal response characteristics of the individual neurons, including the observed spike-count distribution, and the correlation between the spike count of different neurons influence the predicted relationship between the coarse and fine temporal codes. Even if the number of precisely timed spike patterns exceeds the number randomly expected, it is possible that the observed number of precisely timed spike patterns can be related to the predicted number by scale factor. When all of these criteria have been taken into account, precisely timed spike patterns can be related to and therefore predicted from the coarse and intermediate temporal resolution signals. Both this predictability and direct information measures in the data to date indicate that precisely timed spike patterns carry information that is also available from coarse temporal neural codes (Oram *et al.* 1999*a*,*b*, 2001).

3. EFFECTS OF ATTENTION ON THE NEURAL CODE AND THEIR IMPLICATIONS

Human psychophysical studies have shown that visual images such as shape or colour combinations are not always perceived correctly. When subjects are instructed to attend to specific feature combinations, a reduction in the number of errors (false conjunctions) is observed (Treisman 1996). As solutions to the binding problem have been explicitly linked with precisely timed spikes (von der Malsburg & Schneider 1986; Singer & Gray 1995; von der Malsburg 1995) and because behavioural modulation of stimulus significance has been shown to influence perceptual binding (Treisman & Kanwisher 1998), it is important to assess how the behavioural significance of stimuli influences the information that is carried by the neural responses.

The effects of behavioural significance on the coarse temporal response measures (spike count over hundreds of milliseconds) in the non-human primate indicate that attention enhances responses to effective stimuli (for review, see Desimone & Duncan 1995). The DMS task manipulates the attention given to a stimulus depending on whether the stimulus is presented as a sample (attended), a non-match (ignored) or a match (attended and target). It is found that the mean number of spikes elicited in response to stimuli in an attended condition is slightly higher than when ignored (e.g. Moran & Desimone 1985; Richmond & Sato 1987; Sato 1988; Fuster 1990; Desimone & Duncan 1995; Treue & Maunsell 1996; Reynolds *et al.* 2000). The information carried by precisely timed spikes within and between the responses of simultaneously recorded TE neurons of the macaque monkey during a DMS task has also been examined (Oram *et al.* 1999*a*). Concurrently with the changes in the coarse temporal measures, the numbers and types of precisely timed triplets of spikes (both within the responses of individual neurons and between responses of different neurons) changed with DMS phase. However, both the number and type of precisely timed spike patterns were predictable from the changes in the coarse and intermediate temporal response measures (spike count, spike-count variability and temporal variations in the PSTH). Furthermore, the information available from the precisely timed spike patterns was redundant with the information available from the spike count (Oram *et al.* 1999*a*,*b*). Thus, the influence of behavioural significance on stimulus-elicited neural responses in the non-human primate indicates that attention acts as a 'gain control', scaling the responses to each stimulus by a constant (Motter 1994; Desimone & Duncan 1995; Treue & Maunsell 1996) but otherwise leaving the neural code unchanged (Oram *et al.* 1999*a*,*b*; Wiener *et al.* (2001).

The information that can be carried by a given neural code is determined by the range of possible responses and the variability of those responses. There is a strong relationship between the mean spike count elicited when viewing a single stimulus and the variance of the spike counts on individual trials elicited by that stimulus (e.g. Tolhurst *et al.* 1983; Vogels *et al.* 1989; Wiener *et al.* 2001). The probability of stimulus s being present given the response r, $p(s|r)$, is given by the frequency distribution of responses to stimulus s. Due to the fact that the variance is determined by the mean–variance relationship, the spike-count distribution to each stimulus is fully described by the mean response. Thus, $p(s|r)$ can be determined on a trial by trial basis for each stimulus (or more strictly the probability of each possible mean response given the response r) and hence a single rule, governed by the mean–variance relationship, can be used to decode the neural responses (Oram *et al.* 1998; Gershon *et al.* 1998; Wiener & Richmond 1999; Wiener *et al.* 2001). As the mean–variance relationship remains constant across the different conditions in a DMS task, a single rule can be used to decode the neural responses regardless of the particular DMS condition (Wiener *et al.* 2001).

As behavioural modulation of neural responses affects the mean spike count but not the mean–variance relationship or the fine temporal measures, there is an ambiguity in the signal from inferotemporal neurons during a DMS task. Imagine the responses to two stimuli, A and B. Under passive viewing conditions the mean spike count elicited by stimulus A could be higher than that elicited by stimulus B. With no signal to indicate the behavioural condition, a high spike count could be due to stimulus A being viewed under passive conditions or stimulus B being viewed when attended. This ambiguity could be addressed by the presence of a signal from a second source indicating whether the responses should be decoded using the mean responses associated with the attended or unattended state. The available data indicate that is unlikely to be the case. If changes in decoding are given by a concurrent signal, then the effects of attention on neural responses should be evident from the response onset. For example, responses elicited by different stimuli in the inferotemporal cortex deviate within the first few milliseconds of the response onset (Oram & Perrett 1992, 1994, 1996; Tovee *et al.* 1993). The responses to attended and unattended stimuli, however, begin to deviate from each other after a period of some 150–200 ms (e.g. Chelazzi *et al.* 1993; Lueschow *et al.* 1994). The delay from response onset to the differentiation of the neural responses due to attention modulation is consistent with top–down competitive models of attention rather than a change in the decoding of input signals (e.g. Usher & Niebur 1996; Reynolds *et al.* 1999, 2000).

The ambiguity inherent in signals modulated by behavioural context with no 'context-specific' signal can be overcome by examining the outputs from several neurons at a time—a population code. Figure 4*a* shows the mean spike counts of simulated neurons to different exemplars or instances of a stimulus class (e.g. 'Is this an apple or a pear?') when attended (thick lines) and when ignored (thin lines). We use the maximum-likelihood method to estimate which stimulus instance was presented. The error of the decoded estimate is taken as the difference between the estimated and actual instance averaged over 10 000 simulated trials. Under an ideal observer scheme (Foldiak 1993; Oram *et al.* 1998), the decoding of the population response utilizes the mean and variability to each stimulus–behavioural-condition combination. With the ideal observer, the accuracy of the population response to unattended stimuli remains constant across all gain-control levels (white circles in figure 4*b*) while the accuracy to attended stimuli increases as the gain control increases (black circles).

If no signal is present for the behavioural condition, as with inferior temporal neural signals, decoding could occur in three ways. First, decoding may occur using a strategy that assumes the spike count on a given trial

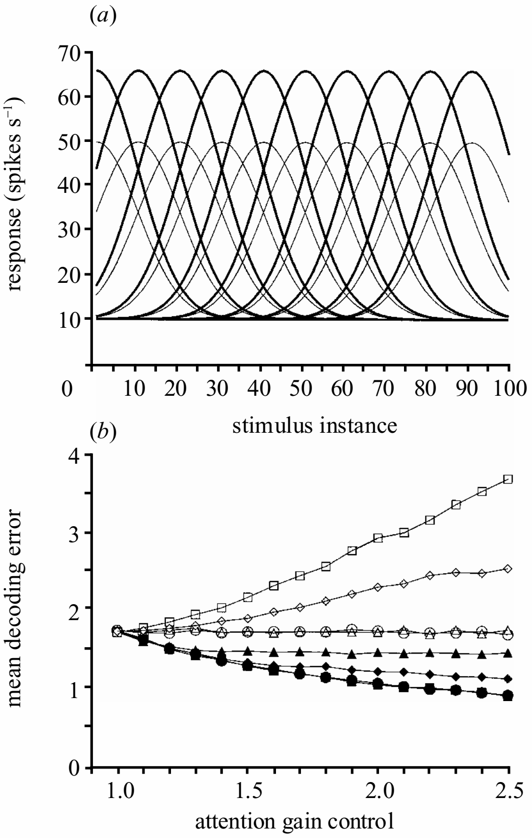

Figure 4. Limitations of attention: discriminating among attended stimuli. Modelling attention as a spike-count gain control reveals the possible limitations of attention. (*a*) The mean spike counts of 10 simulated neurons are plotted for 100 different instances of stimuli (e.g. apple, pear and orange are all instances of fruit). The variance of the responses to each instance was set as twice the mean spike count. The spike counts were either simulated as being in an unattended state (gain control = 1.0, thin lines) or in an attended state (gain control = 1.4, thick lines). (*b*) Plot of the error as a function of the multiplicative gain control due to attention. Black symbols show the error when the stimuli are presented whilst attended. White symbols show the error when the stimuli are presented but unattended. The error in estimating the most likely input stimulus from the responses of a simulated population of neurons is plotted as a function of the strength of attention gain. Mean errors from 10 000 simulated trials using four decoding strategies are plotted as follows. (1) The ideal observer who 'knows' the response distributions to each stimulus–attention condition (circles). (2) Constant decoding where the observer always assumes spike-count distributions associated with unattended stimuli (triangles). (3) Constant decoding where the observer assumes the spike-count distributions associated with attended stimuli (triangles). (4) Decoding assuming that the response distribution for each stimulus could have come from either the attended or unattended state (diamonds) (see § 3 for details). The method assumes an equal probability of each stimulus instance, $p(s_1) = p(s_2)\ldots p(s_n)$; the probability of each of the possible 100 instances of the stimulus was assessed for each simulated trial for each neuron using $p(s|r) = p(r|s)/\Sigma_{i=1}^{100} p(r|s_i)$ where $p(s|r)$ is the probability of stimulus s being present given the trial response of r spikes, $p(r|s)$ is the probability of observing r spikes given that stimulus s was present and $\Sigma_{i=1}^{100} p(r|s_i)$ is the sum of the $p(r|s)$) over all n stimulus instances (for review, see Oram *et al.* 1998). The $p(r|s)$ were obtained from the truncated Gaussian distributions given by the mean and

comes from the means and variances of each stimulus when the stimuli are ignored (thin lines, figure 4*a*). This gives optimal decoding of unattended stimuli (white triangles, figure 4*b*) but a decreased accuracy of decoding attended stimuli compared with ideal observer decoding (black triangles). With this decoding system the benefit of increased attention becomes negligible at a gain control above *ca.* 1.3–1.4. In addition, decoding could occur assuming that the incoming signal is always derived from attended stimuli (thick lines, figure 4*a*). This gives rise to the optimal decoding of stimuli that are attended (black squares, figure 4*b*) but decreased accuracy in determining which stimulus was present when the stimuli are unattended compared with ideal observer decoding (white squares). Finally, decoding could occur using a frequency distribution of spike counts given by a combination of the distribution associated with the unattended responses and the distribution associated with the responses from the attended state. This can be thought of as representing a decoding strategy that is learnt from repeated exposures to the stimuli in both the attended and unattended conditions. The effects of this decoding system lie between the schemes described above (unattended, white diamonds; attended, black diamonds; total distribution, 50% attended, 50% unattended; figure 4*b*). Thus, attention-modulated gain control enhances the ability of a population of neurons to transmit information about which particular exemplar of the attended stimulus class is present, even when no context-specific signal is available. However, in the absence of a context-specific signal, the decrease in the size of the error with increasing attention gain control is smaller than in the presence of a context-specific signal. This indicates that there is a limited benefit in the attention-driven gain control's increasing (more than 1.3–1.4). Restriction of the attention gain-control magnitude also reduces the potential decrement in discrimination between members of a stimulus group when the group is not being attended.

Attention not only increases the discrimination between attended stimuli but can also act as a selective filter (e.g. 'Tell me when you see a fruit' rather than 'Is the fruit an apple or pear?'). Figure 5 shows the effects of attention-modulated gain control on the discrimination between attended (fruit) and unattended stimuli (non-fruits). The thin lines in figure 5*a* show the simulated mean spike counts of different neurons when stimuli of different classes are presented in the unattended state. The thick line shows the changes in mean responses during selective attention (e.g. attending stimulus class 50—fruit—selectively enhances the responses of neurons tuned to class 50 but not neurons maximally responsive to other classes). The errors of decoding such a population under different schemes (details as for figure 4) are plotted in figure 5*b* as a function of different levels of attention gain control.

Decoding of the population response using the mean and variability to each stimulus class–behavioural-condition combination indicates that there is only a small

---

variance (Foldiak 1993; Oram *et al.* 1998; Gershon *et al.* 1998). Assuming independence, the probability of stimulus instance s being present is given by the product of the $p(s|r)$ from all neurons (for details, see Oram *et al.* 1998).

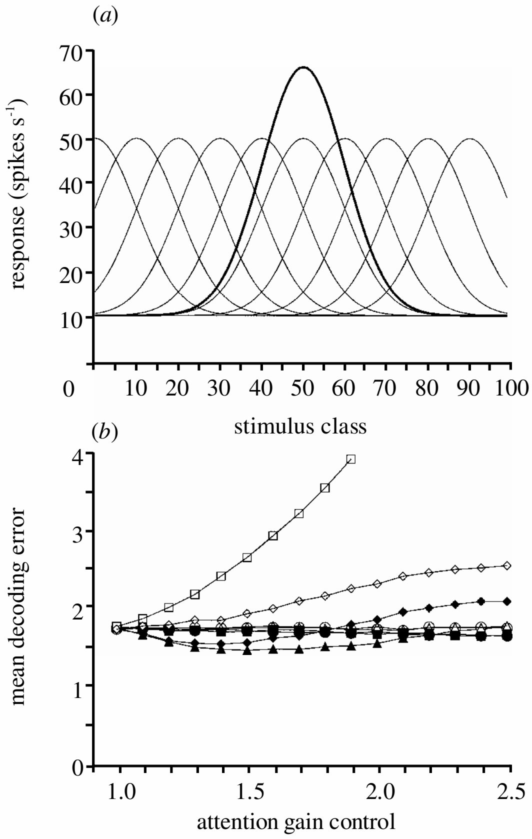

Figure 5. Limitations of attention: discriminating attended from unattended stimuli. (*a*) The mean spike counts of 10 simulated neurons are plotted for 100 different classes of stimuli (e.g. fruit, vegetable, furniture). The variance of the responses to each instance was set as twice the mean spike count. The spike counts were either simulated as being in an unattended state (gain control = 1.0, thin lines) or in an attended state (gain control = 14, thick lines) when attention was selective towards only one stimulus class (class 50). (*b*) Plot of the error as a function of the multiplicative gain control due to attention. The black symbols show the error when the stimuli are presented whilst attended. The white symbols show the error when the stimuli are presented but unattended. The error in estimating the most likely input stimulus from the responses of a simulated population of neurons is plotted as a function of the strength of attention gain. Mean errors from 10 000 simulated trials using the following four decoding strategies are plotted. (1) The ideal observer who 'knows' the response distributions to each stimulus–attention condition (circles). (2) Constant decoding where the observer always assumes spike-count distributions associated with unattended stimuli (triangles). (3) Constant decoding where the observer assumes the spike-count distributions associated with attended stimuli (triangles). (4) Decoding assuming that the response distribution for each stimulus could have come from either the attended or unattended state (diamonds) (see § 3 for details). For a description of the method see the legend to figure 4 and Oram *et al.* (1998) for details.

decrease in the error when the stimulus class is attended compared with when the stimulus class is ignored (black versus white circles). If the responses are decoded assuming all responses are derived from attended stimuli, the error for unattended stimuli rises rapidly (white squares). Decoding assuming that all responses are elicited by unattended stimuli produces a surprising result. For gain-control values from *ca.* 1.2 to 2.0, this decoding strategy shows mean errors in decoding attended stimuli that are below the error rates for all other decoding strategies. This is because the 'amplified' responses to an attended stimulus (in relation to the unattended mean and variance) are unequivocally decoded as being the attended stimulus. When the gain control becomes too large, however, stimuli that share features with the attended stimulus (i.e. are close in response space) also become so large that they are (incorrectly) decoded as being the attended stimulus.

(a) *Interim summary: implications of neural coding for attention mechanisms*

The effects of changing the amount of attention given to visual stimuli on neural responses are well described as a multiplicative gain control (Desimone & Duncan 1995). Changes in the fine temporal structure of the responses of individual neurons and between the responses of different neurons are predictable from the coarse and intermediate temporal resolution signals and carry only a small proportion of the information available from the spike count (Oram *et al.* 1999*a*,*b*). While both the mean and the variance of the spike count change with changes in attention, the mean–variance relationship remains constant, allowing for a decoding mechanism that is constant across different attentional states (Wiener *et al.* 2001). The effects of attention on neural responses are not seen for some 150–200 ms; this is consistent with a competitive mechanism of attention operating between different neurons within a given brain area (Usher & Neiber 1996; Reynolds *et al.* 1999, 2000) and inconsistent with a change in the decoding of the input signals. The absence of evidence for changes in the decoding of inputs with changes in attention and a multiplicative gain control of spike count gives rise to possible ambiguities in neural signals. Simulations of population decoding show that even without a signal of the attention state, decoding the neural spike counts elicited by attended stimuli can still give rise to improved discrimination between those stimuli compared with when the stimuli are unattended (figure 4). The absence of a signal about the attentional state can provide better discrimination between attended and unattended stimuli than when the attentional state during stimuli presentation is known (figure 5).

4. IMPLICATIONS OF PRECISELY TIMED SPIKE PATTERNS RELATIVE TO EXTERNAL EVENTS

The precise times of spikes *relative to stimulus onset* (response latency) are another mechanism by which neurons could signal different stimulus attributes within a single response. Recordings from the macaque visual system show that response latency, but not response strength, is heavily influenced by stimulus contrast (Gawne *et al.* 1996; Opara and Worgotter 1996; Oram *et al.* 1997; Reich *et al.* 1997; Oram *et al.* 2002; Xiao *et al.* 2001). The temporal precision of response latency, estimated as the 'noise' required to reduce the information transmitted by latency by 50%, is *ca.* 10 ms (Reich *et al.* 1997). Conversely response strength, but not response latency, is influenced by stimulus orientation, change in perspective view, object part visibility, size or direction of motion

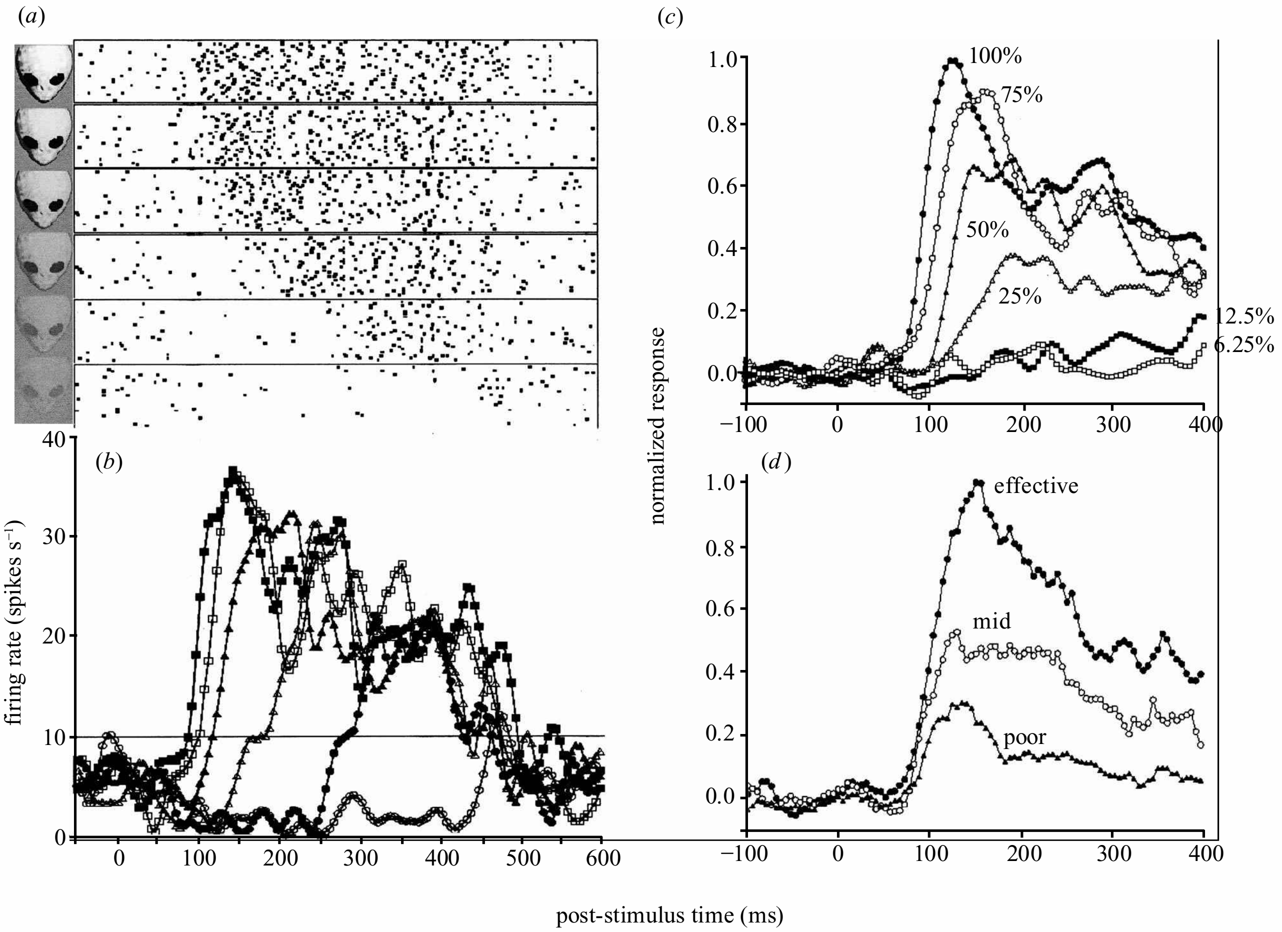

Figure 6. The stimulus contrast influences the response latency. Recordings from neurons in the superior temporal cortex were made using standard techniques. (*a*,*b*) The effect of changing stimulus contrast on responses of a single neuron. (*a*) The rastergrams of responses to a stimulus of 100%, 75%, 50%, 25%, 12.5% and 6.25% internal contrast (contrast $= 100(L_{\max} - L_{\min})/(L_{\max} + L_{\min})$, where $L_{\max}$ is the maximum luminance and $L_{\min}$ is the minimum luminance within the image). (*b*) The spike-density functions (the sum of the individual rastergrams smoothed with a Gaussian of standard deviation of 5 ms). Black squares, 100%; white squares, 75%; black triangles, 50%; white triangles, 25%; black circles, 12.5%; white circles, 6.25%. (*c*,*d*) Normalized population responses across neurons. The response latency and response magnitude change with changes in the stimulus contrast (*c*); however, the response latency changes very little when the effectiveness of the stimulus is varied despite large changes in the response magnitude (*d*). Black circles, 100%; white circles, 75%; black triangles, 50%; white triangles, 25%; black squares, 12.5%; white squares, 6.25%. (Adapted from Oram & Perrett (1992) and Xiao *et al.* (2001).)

(Oram & Perrett 1992; Oram *et al.* 1993, 2002; Oram & Perrett 1994, 1996; Gawne *et al.* 1996; Opara & Worgotter 1996; Oram *et al.* 1997; Perrett & Oram 1998; Xiao *et al.* 2001).

We have recently examined the responses of single neurons in the inferior temporal cortex and the superior temporal sulcus of the macaque monkey to stimuli of constant average luminance that varied in internal contrast (Xiao *et al.* 2001). The increase in response latency with decreasing stimulus contrast was seen for all neurons and could not be attributed to the observed changes in response magnitude (figure 6). Decreasing the stimulus contrast from 100% to 6% produced an average latency increase of over 150 ms (Xiao *et al.* 2001), some three to four times the magnitude of the effect reported in the striate cortex (Gawne *et al.* 1996). Thus, changes in stimulus contrast, but not other visual transformations, cause changes in response latency, and this relationship between stimulus contrast and response latency seems to increase as one moves through the visual system.

Changes in stimulus contrast elicit changes in response latency with little or no dependency on response magnitude, allowing for the simultaneous transmission of two neural signals. This has led to the suggestion that response latency could act as a potential signal that facilitates the binding or linking of the responses of different neurons (Gawne *et al.* 1996). For example, the stripes of a zebra close to the observer will be seen as highly contrasting. The stripes of the zebra viewed across the plain become blurred because, from the observer's point of view, the spatial frequency of a zebra's stripes increases with viewing distance. The blurring reduces the effective contrast and so will give rise to visual responses with longer response latency than the visual responses elicited by a zebra seen at close quarters. The changes in response latency, therefore, provide a cue of 'belonging-ness'. Neurons whose response onsets are close together are likely to be signalling information in their spike counts about the same object (Gawne *et al.* 1996). Computational models of object recognition that use a winner-take-all form of competition based on precise arrival times of spikes have been proposed that could utilize contrast-related signals in exactly this way (e.g. Thorpe 1990; Van Rullen *et al.* 1998).

Changes in response latency and changes in response magnitude have implications for cognitive models of brain function. We assume that behavioural decisions are only made when the relevant information has exceeded some threshold (Loftus & Ruthruff 1994; Bussey & Loftus 1994). As the neural code about stimulus identity can be summarized by spike count and response latency (see previous paragraph), we use the time for the cumulative spike count to reach a given threshold to give the predicted pattern of recognition RTs (Hanes & Schall 1996; Perrett *et al.* 1998). Figure 7*a* shows the responses from a population of inferotemporal neurons to stimuli of different contrast (data from Xiao *et al.* 2001). (The responses have been normalized so that the average background activity is 0 and the peak response is 1.) Figure 7*b* shows the cumulative response from stimulus onset (time = 0) and the cumulative response if the accumulation of the neural signal starts 200 ms after stimulus onset. Figure 7*c* shows the time at which the cumulative response reaches 10 spikes. The time to threshold decreases with increasing contrast when the cumulative spike count starts at stimulus onset (no delay). There is an increase in time to threshold when the cumulative count starts after 200 ms and there is no longer a dependency of time to threshold on the stimulus contrast.

Figure 7*d–f* shows the predicted activity of temporal cortical neurons to stimuli of different orientations and the times to threshold. Figure 7*d* shows the spike-density functions of temporal cortical neurons to stimuli that elicited three different response strengths (data from Oram & Perrett 1992, and see figure 6). As the responses of temporal cortical neurons are reduced when stimuli are rotated about 60° from upright (Ashbridge *et al.* 2000), these responses correspond to the responses to upright images, images rotated 60° from upright and inverted images (see also Perrett *et al.* 1988). The middle right panel shows the cumulative response from stimulus onset (time = 0) and the cumulative response if the accumulation of the neural signal starts 200 ms after stimulus onset. Figure 7*f* shows the time for the cumulative responses to reach a threshold of 10 spikes above background activity. The time to threshold depends on the stimulus orientation and, as with changes in stimulus contrast, the time to reach the threshold changes when there is a delay before the cumulative spike count starts. Critically, the response elicited by an ineffectual (inverted) stimulus takes longer to reach the threshold than the response to an effective (upright) stimulus regardless of the delay in monitoring the neural responses.

As a preliminary investigation into the validity of these predictions, we used human subjects in a dual-task experiment. The principle underlying dual tasks is that the decision-making processes (occurring in the 'central executive' of the working memory; Baddeley 1986) do not start on a second task until the ongoing processes associated with the first task have been completed. If one task—a counting task, for example—is being performed then no executive processing related to a second task—visual discrimination and recognition—can occur. Under this scheme, a counting task followed immediately by a visual discrimination task would involve a delay (while the first task was performed) before executive processes could monitor and process incoming visual system signals. Conversely, a long interval between the first and second tasks should result in no delay in monitoring incoming visual signals associated with the second task. We therefore predicted that recognition RTs would depend on stimulus contrast when there was no delay in monitoring but would be independent of stimulus contrast if subjects were still making decisions about another task (figure 7*c*). Furthermore, we expected that recognition RTs would depend on stimulus orientation at both long and short intervals between a counting task and the presentation of the visual stimulus.

The preliminary results indicate that the human observers follow the pattern of results predicted in figure 7. Subjects were presented with a sequence of one to four auditory 'pips' followed by the presentation of a letter (R or G) in either its normal or mirror form (reflected about the vertical). The letter (normal or mirror) was presented in different orientations (0°, ± 45°, ± 90°, ± 135° or 180° from upright) and at two contrasts (100% and 6%). The interval between the last pip and the onset of the visual stimulus was either 0 or 1000 ms. The subjects' task was first to indicate whether the letter was normal (press the Y key) or mirror (press the N key), then to indicate whether the number of pips was odd (Y) or even (N). Subjects were instructed to make their responses as fast and as accurately as they could. Recognition RTs were taken as the lag between the onset of the visual stimulus and the key press.

Figure 8 plots the mean recognition RTs from 11 subjects in such a task as a function of stimulus orientation. The mean RTs were longer when the interval between the last pip and the onset of the visual stimulus was 0 compared with 1000 ms. The mean RTs taken to indicate whether the letter was normal or its mirror image were longer when there was no interval between the last pip and stimulus presentation than when there was a long interval. This indicates that performing the counting task delayed the processing of the visual task (for a review, see Pashler & Johnston 1998). High-contrast images were recognized faster than low-contrast images when the interval between the two tasks was long (compare the black and white triangles in figure 8). The effect of contrast on the RTs was attenuated when the visual stimulus was presented immediately after the final pip ($p < 0.005$, compare the black and white circles). The RT increased with increasing rotation from vertical across the conditions in a statistically indistinguishable way ($p > 0.3$), importantly being independent of the interval between the two tasks (compare the triangles and circles in figure 8).

(a) *Interim summary: latency as a precisely timed neural code*

Response latency in visual-system responses forms a neural code of relatively high temporal precision (*ca.* 10 ms; Reich *et al.* 2001) where the timing is defined relative to an external event (stimulus onset). Response latency conveys information that is unavailable from the spike counts and their correlation (Gawne *et al.* 1996; Wiener *et al.* 1999; Reich *et al.* 2001) and therefore seems unlike

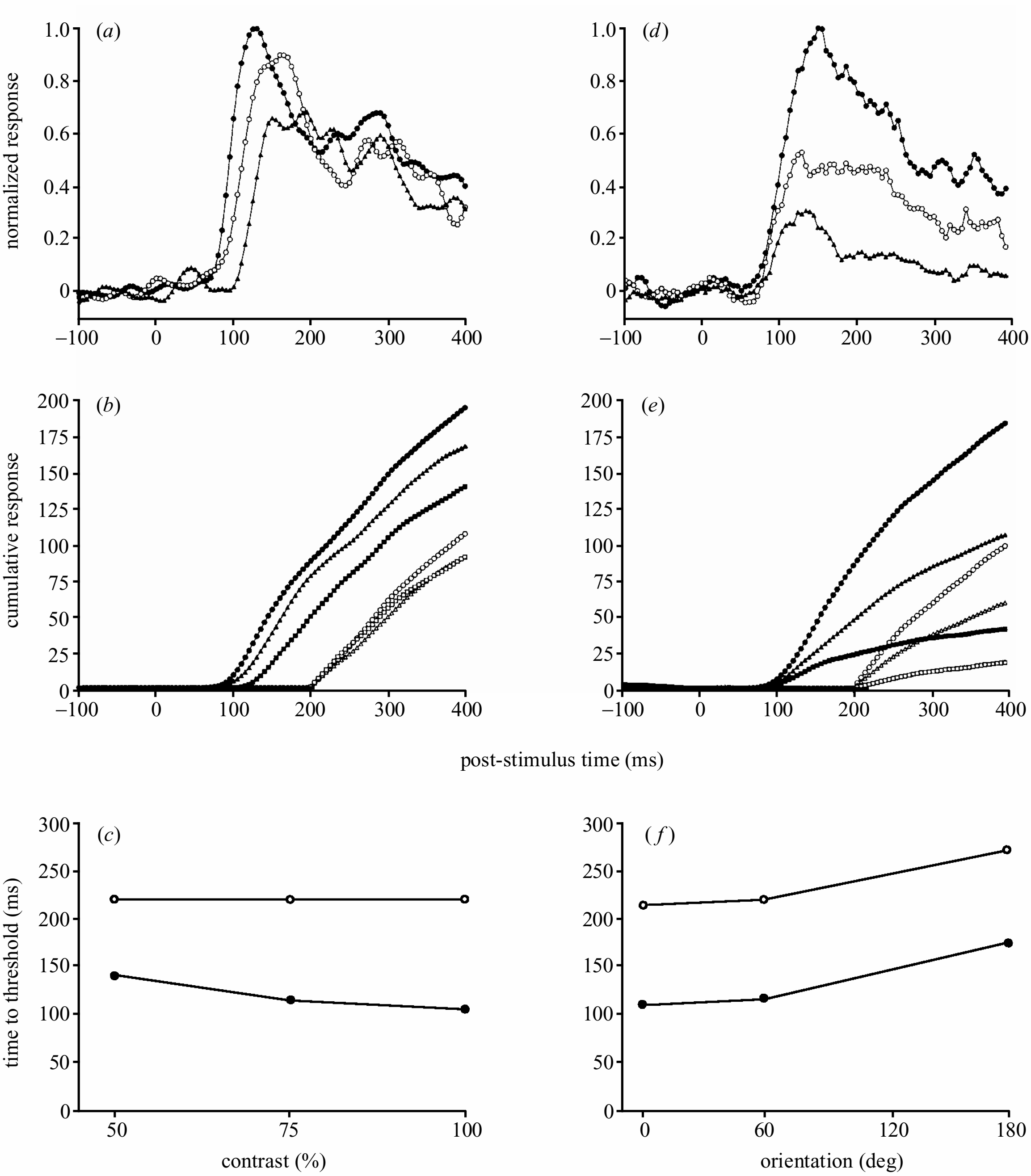

Figure 7. Predicting recognition RTs from neural responses. The population responses to stimuli of different contrasts (*a*–*c*) and different rotations (*d*–*f*) are shown. The black symbols show the cumulative response when there is no delay in monitoring the population response. (*a*) The population responses to stimuli of 100% (black circles), 75% (white circles) and 50% (black triangles) contrast are shown. The response latency varies with the stimulus contrast. (*b*) The cumulative response is shown to stimuli of 100% (circles), 75% (triangles) and 50% (squares) contrast. (*c*) Time to threshold (ms) versus contrast (%): delay (white circles); no delay (black circles). (*d*) Population responses to three stimuli of different effectiveness: upright (black circles); 60° (white circles); inverted (black triangles). The responses have approximately the same latency but different magnitudes. (*e*) Cumulative response to stimuli of three different levels of effectiveness: upright (circles); 60° (triangles); inverted (squares). (*f*) Time to threshold (ms) versus orientation (°): delay (white circles); no delay (black circles).

other precisely timed neural codes where the spike times are defined relative to the times of other spikes. The latency of visual-system responses is influenced by changes in stimulus contrast but not other image changes such as changes in stimulus size, shape, lighting or direction of motion. Preliminary studies with human subjects indicate that changes in response latency with changes in stimulus contrast are relevant to behaviour; in particular how cognitive processes interact (figures 7 and 8).

5. DISCUSSION

The potential for neural responses to convey information using multiple codes is appealing in terms of increasing the information processing capacity without needing to increase the number of processing elements. However, we are only beginning to understand the highly complex relationship between the fine and coarse temporal codes inherent in neural activity. Even slight errors

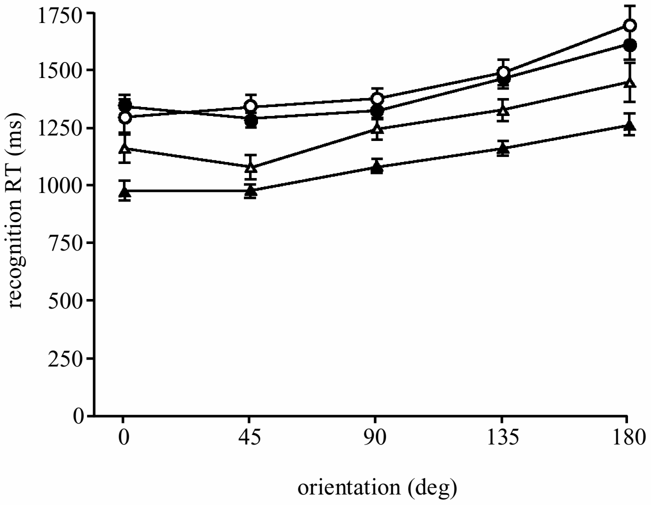

Figure 8. Mean recognition RTs (± s.e.m.) for discriminating a mirror image of a letter from a normal letter. Only data from trials where the responses to both the auditory and the visual task were correct have been included. The slope relating RT to stimulus rotation was highly significant and statistically equivalent across the four experimental conditions. The effect of decreasing the stimulus contrast was to increase the recognition RT. When there was a 1000 ms interval between the end of the counting task and the presentation of the visual stimulus an effect of stimulus contrast was seen. As predicted, the effect of stimulus contrast on the recognition RT was reduced when the visual stimulus was presented immediately after the counting task. See § 4 for details. Overall analysis: effect of stimulus rotation $F_{4,40} = 34.2$, $p < 0.0005$; rotation by delay $F_{4,40} = 0.27$, $p > 0.8$; rotation by contrast $F_{4,40} = 0.31$, $p > 0.8$; rotation by contrast by delay $F_{4,40} = 1.06$, $p > 0.3$. Effect of stimulus contrast $F_{1,10} = 58.1$, $p < 0.0005$; effect of delay between pips and image presentation $F_{1,10} = 36.7$, $p < 0.0005$; delay by contrast $F_{1,10} = 15.9$, $p = 0.003$. Circles, 0 ms interval; triangles, 1000 ms interval; white symbols, low contrast; black symbols, high contrast.

in the assumptions underlying analysis can lead to very different conclusions about the statistical significance of observed numbers of precisely timed spike patterns where the timing is defined relative to other spikes (Brody 1999*a–c*; Oram *et al.* 1999*a*,*b*; Oram *et al.* (2001). Furthermore, the presence of a precisely timed spike pattern, such as synchronous spikes, above chance levels does not imply that the observed spike pattern carries information unavailable from coarse temporal response measures and their correlation. This is true even if the observed number of precisely timed spike patterns exceeds the number expected by chance in responses to some experimental condition and not others (Oram *et al.* 2001). Overall, the evidence that precisely timed neural codes (where the timing is relative to other spikes) convey information unavailable from coarse temporal measures of neural activity is still inconclusive (see also Shadlen & Newsome 1994, 1995, 1998; Treves *et al.* 1999). Additionally, precisely timed spike patterns only reflect a small proportion of the information available from the spike counts (e.g. Oram *et al.* 1999*a*,*b*, 2001). These observations do not rule out that the *decoding* of the signals involves monitoring and processing based on precisely timed spike patterns. The data to date indicate, however, that precisely timed spike patterns do not form a code that is separable from coarser temporal neural codes in the sense of carrying information that is different from the information carried by coarse temporal neural codes (see Richmond *et al.* 1999).

Given that behavioural modulation of stimulus significance influences perceptual binding (Treisman 1996, 1998) and there is little evidence that behavioural modulation of neural responses influences the fine temporal structure, it seems unlikely that precisely timed spike patterns, such as synchrony (Singer & Gray 1995) or synfire chains (Abeles 1991), are necessarily involved in overcoming the binding problem. An alternative mechanism to precisely timed spikes to overcome the binding problem has been in the literature for many years (Wicklegren 1969; Rumelhart *et al.* 1986). This scheme differs from the precisely timed spike hypothesis in that one response parameter (e.g. spike count) codes multiple stimulus attributes (e.g. colour and form, or form and motion), with each neuron being sensitive to multiple features (Oram & Perrett 1994; Perrett & Oram 1998; Mel & Fiser 2000). Neural networks using a simple learning rule can generate output nodes that show sensitivity to combinations of stimulus attributes (Oram & Foldiak 1996; Wallis & Rolls 1997). Importantly, the feared combinatorial explosion in the number of feature conjunctions requiring coding (von der Malsburg 1995) need not arise, even when approximately two billion feature conjunctions are possible (Ullman & Soloviev 1999; Mel & Fiser 2000).

The absence of a separate neural code indicating attention modulation has implications for the processing of attention. Maximizing the discrimination between attended stimuli involves the presence of a context-specific signal (figure 4). Alternatively, if the mechanisms underlying attention (the gain control) are constrained to give the maximum benefit (error reduction) in discriminating attended from unattended stimuli, then there should be no signal indicating whether or not a stimulus is attended. A context-specific signal indicating the state of attention is not evident in the neural codes of inferotemporal cortical neurons. Furthermore, the time delay from response onset to the effects of attention indicates that it is unlikely that such a signal is input to inferotemporal cortical neurons. Thus, the neural codes present in the ventral stream of the primate visual cortex indicate that the mechanisms of attention maximize the discrimination of attended from unattended stimuli.

Changes in response latency and response magnitude are associated with changes in stimulus contrast and changes in stimulus shape, orientation or size, respectively. These two types of coding give rise to different predictions if the evidence or information accumulation is delayed. Assuming the executive processes are simply monitoring the 'slave' visual system, the dual task paradigm achieves just this delay. The effect of changes in stimulus contrast on RTs are predicted, and shown, to decrease if subjects are actively engaged in another task when the stimuli are presented. The effects on RTs of image manipulations that change in the strength of activation of a neural representation (e.g. changes in stimulus orientation) are predicted and shown to be independent of ongoing tasks. The traditional interpretation of differences in dependency on stimulus-onset asynchrony (the delay in the onset of monitoring) has been to assume fundamentally different cognitive processes (executive and slave). The observation that the interpretation of the results from psychological experiments requires an understanding of the nature of the representations and how the underlying neural code changes with changes in the access of those representations presents a new challenge to both psychologists and cognitive neuroscientists.

REFERENCES

Abbott, L. F. & Dayan, P. 1999 The effect of correlated variability on the accuracy of a population code. *Neural Comput.* **11,** 91–101.

Abeles, M. 1991 *Corticonics*. Cambridge University Press.

Abeles, M. & Gerstein, G. L. 1988 Detecting spatiotemporal firing patterns among simultaneously recorded single neurons. *J. Neurophysiol.* **60,** 909–924.

Abeles, M., Bergman, H., Margalit, E. & Vaadia, E. 1993 Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. *J. Neurophysiol.* **70,** 1629–1638.

Aertsen, A. M., Gerstein, G. L., Habib, M. K. & Palm, G. 1989 Dynamics of neuronal firing correlation: modulation of 'effective connectivity'. *J. Neurophysiol.* **61,** 900–917.

Ashbridge, E., Perrett, D. I., Oram, M. W. & Jellema, T. 2000 Effect of image orientation and size on object recognition: responses of single units in the macaque monkey temporal cortex. *Cogn. Neuropsych.* **17,** 13–34.

Baddeley, A. D. 1986 *Working memory*. Oxford University Press.

Baddeley, R., Abbott, L. F., Booth, M. C., Sengpiel, F., Freeman, T., Wakeman, E. A. & Rolls, E. T. 1997 Responses of neurons in primary and inferior temporal visual cortices to natural scenes. *Proc. R. Soc. Lond.* B **264,** 1775–1783. (DOI 10.1098/rspb.1997.0246.)

Baker, S. N. & Lemon, R. N. 2000 Precise spatiotemporal repeating patterns in monkey primary and supplementary motor areas occur at chance levels. *J. Neurophysiol.* **84,** 1770–1780.

Baker, S. N., Spinks, R., Jackson, A. & Lemon, R. N. 2001 Synchronization in monkey cortex during a precision grip task: I. Task-dependent modulation in single unit synchrony. *J. Neurophysiol.* **85,** 869–885.

Berry, M. J. & Meister, M. 1998 Refractoriness and neural precision. *J. Neurosci.* **18,** 2200–2211.

Berry, M. J., Warland, D. K. & Meister, M. 1997 The structure and precision of retinal spike trains. *Proc. Natl Acad. Sci. USA* **94,** 5411–5416.

Bradley, A., Skottun, B. C., Ohzawa, I., Sclar, G. & Freeman, R. D. 1987 Visual orientation and spatial frequency discrimination: a comparison of single neurons and behavior. *J. Neurophysiol.* **57,** 755–772.

Britten, K. H., Shadlen, M. N., Newsome, W. T. & Movshon, J. A. 1993 Responses of neurons in macaque MT to stochastic motion signals. *Vis. Neurosci.* **10,** 1157–1169.

Brody, C. D. 1999*a* Disambiguating different covariation types. *Neural Comput.* **11,** 1527–1535.

Brody, C. D. 1999*b* Correlations without synchrony. *Neural Comput.* **11,** 1537–1551.

Brody, C. D. 1999*c* On artefactual spike train cross-correlations. *Neurocomputing* **26–7**, 957–962.

Buracas, G. T., Zador, A. M., DeWeese, M. R. & Albright, T. D. 1998 Efficient discrimination of temporal patterns by motion-sensitive neurons in primate visual cortex. *Neuron* **20,** 959–969.

Bussey, T. A. & Loftus, G. R. 1994 Sensory and cognitive components of visual information acquisition. *Psychol. Rev.* **101**, 446–469.

Chelazzi, L., Miller, E. K., Duncan, J. & Desimone, R. 1993 A neural basis for visual-search in inferior temporal cortex. *Nature* **363,** 345–347.

Dayhoff, J. E. & Gerstein, G. L. 1983*a* Favored patterns in spike trains. I. Detection. *J. Neurophysiol.* **49,** 1334–1348.

Dayhoff, J. E. & Gerstein, G. L. 1983*b* Favored patterns in spike trains. II. Application. *J. Neurophysiol.* **49,** 1349–1363.

Desimone, R. & Duncan, J. 1995 Neural mechanisms of selective visual attention. *A. Rev. Neurosci.* **18,** 193–222.

Engel, A. K., Konig, P., Kreiter, A. K., Schillen, T. B. & Singer, W. 1992 Temporal coding in the visual cortex: new vistas on integration in the nervous system. *Trends Neurosci.* **15,** 218–226.

Eskandar, E. N., Richmond, B. J. & Optican, L. M. 1992*a* Role of inferior temporal neurons in visual memory. I. Temporal encoding of information about visual images, recalled images, and behavioral context. *J. Neurophysiol.* **68,** 1277–1295.

Eskandar, E. N., Optican, L. M. & Richmond, B. J. 1992*b* Role of inferior temporal neurons in visual memory. II. Multiplying temporal waveforms related to vision and memory. *J. Neurophysiol.* **68,** 1296–1306.

Foldiak, P. 1993 The 'ideal homunculus': statistical inference from neural population responses. In *Computation and neural systems* (ed. F. H. Eeckman & J. M. Bower), pp. 53–60. Norwell, MA: Kluwer Academic.

Fuster, J. M. 1990 Inferotemporal units in selective visual attention and short term memory. *J. Neurophysiol.* **64,** 681–697.

Gawne, T. J., Kjaer, T. W. & Richmond, B. J. 1996 Latency: another potential code for feature binding in striate cortex. *J. Neurophysiol.* **76,** 1356–1360.

Gershon, E. D., Wiener, M. C., Latham, P. E. & Richmond, B. J. 1998 Coding strategies in monkey V1 and inferior temporal cortices. *J. Neurophysiol.* **79,** 1135–1144.

Hanes, D. P. & Schall, J. D. 1996 Neural control of voluntary movement initiation. *Science* **274,** 427–430.

Hatsopoulos, N. G., Ojakangas, C. L., Paninski, L. & Donoghue, J. P. 1998 Information about movement direction obtained from synchronous activity of motor cortical neurons. *Proc. Natl Acad. Sci. USA* **95,** 15 706–15 711.

Heller, J., Hertz, J. A., Kjaer, T. W. & Richmond, B. J. 1995 Information flow and temporal coding in primate pattern vision. *J. Comput. Neurosci.* **2,** 175–193.

Lee, D., Port, N. L., Kruse, W. & Georgopoulos, A. P. 1998 Variability and correlated noise in the discharge of neurons in motor and parietal areas of the primate cortex. *J. Neurosci.* **18,** 1161–1170.

Lestienne, R. & Strehler, B. L. 1987 Time structure and stimulus dependence of precisely replicating patterns present in monkey cortical neuronal spike trains. *Brain Res.* **437,** 214–238.

Lestienne, R. & Tuckwell, H. C. 1998 The significance of precisely replicating patterns in mammalian CNS spike trains. *Neuroscience* **82**, 315–336.

Levine, M. W. & Troy, J. B. 1986 The variability of the maintained discharge of cat dorsal lateral geniculate cells. *J. Physiol.* (*Lond.*) **375**, 339–359.

Loftus, G. R. & Ruthruff, E. 1994 A theory of visual information acquisition and visual memory with special application to intensity-duration trade-offs. *J. Exp. Psychol.* **20**, 33–49.

Lueschow, A., Miller, E. K. & Desimone, R. 1994 Inferior temporal mechanisms for invariant object recognition. *Cerebral Cortex* **5**, 523–531.

McClurkin, J. W., Optican, L. M., Richmond, B. J. & Gawne, T. J. 1991*a* Concurrent processing and complexity of temporally encoded neuronal messages in visual perception. *Science* **253,** 675–677.

McClurkin, J. W., Gawne, T. J., Optican, L. M. & Richmond,

B. J. 1991*b* Lateral geniculate neurons in behaving primates. II. Encoding of visual information in the temporal shape of the response. *J. Neurophysiol.* **66**, 794–808.

McClurkin, J. W., Gawne, T. J., Richmond, B. J., Optican, L. M. & Robinson, D. L. 1991*c* Lateral geniculate neurons in behaving primates. I. Responses to two-dimensional stimuli. *J. Neurophysiol.* **66**, 777–793.

Maynard, E. M., Hatsopoulos, N. G., Ojakangas, C. L., Acuna, B. D., Sanes, J. N., Normann, R. A. & Donoghue, J. P. 1999 Neuronal interactions improve cortical population coding of movement direction. *J. Neurosci.* **19**, 8083–8093.

Mel, B. W. & Fiser, J. 2000 Minimizing binding errors using learned conjunctive features (vol. 12, p 247. 1999). *Neural Comput.* **12**, 731–762.

Moran, J. & Desimone, R. 1985 Selective attention gates visual processing in the extrastriate cortex. *Science* **229**, 782–784.

Motter, B. C. 1994 Neural correlates of attentive selection for color or luminance in extrastriate area V4. *J. Neurosci.* **14**, 2178–2189.

Opara, R. & Worgotter, F. 1996 Using visual latencies to improve image segmentation. *Neural Comput.* **8**, 1493–1520.

Optican, L. M. & Richmond, B. J. 1987 Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. III. Information theoretic analysis. *J. Neurophysiol.* **57**, 162–178.

Oram, M. W. & Foldiak, P. 1996 Learning generalisation and localisation: competition for stimulus type and receptive field. *Neurocomputing* **11**, 297–321.

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

Oram, M. W. & Perrett, D. I. 1994 Modeling visual recognition from neurobiological contraints. *Neural Networks* **7**, 945–972.

Oram, M. W. & Perrett, D. I. 1996 Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. *J. Neurophysiol.* **76**, 109–129.

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.

Oram, M. W., Gawne, T. J. & Richmond, B. J. 1997 Relationship of response latency and magnitude in the LGN of macaque monkey. *Soc. Neurosci. Abstr.* **23**, 450.

Oram, M. W., Foldiak, P., Perrett, D. I. & Sengpiel, F. 1998 The 'Ideal Homunculus': decoding neural population signals. *Trends Neurosci.* **21**, 259–265.

Oram, M. W., Lui, Z. & Richmond, B. J. 1999*a* Precisely timed spike patterns within the TE neuronal responses of a monkey performing a delayed match to sample task are consistent with chance. *Soc. Neurosci. Abstr.* **25**, 915.

Oram, M. W., Wiener, M. C., Lestienne, R. & Richmond, B. J. 1999*b* Stochastic nature of precisely timed spike patterns in visual system neuronal responses. *J. Neurophysiol.* **81**, 3021–3033.

Oram, M. W., Hatsopoulos, N. G., Richmond, B. J. & Donoghue, J. P. 2001 Excess synchrony in motor cortical neurons provides redundant direction information with that from coarse temporal measures. *J. Neurophysiol.* **86**, 1700–1716.

Oram, M. W., Wiener, M. C. & Richmond, B. 2002 Relationship between stimulus contrast, stimulus pattern, response magnitude and response latency in striate cortex and the lateral geniculate nucleus of the macaque monkey. (In preparation.)

Palm, G., Aertsen, A. M. & Gerstein, G. L. 1988 On the significance of correlations among neuronal spike trains. *Biol. Cybern.* **59**, 1–11.

Panzeri, S., Schultz, S. R., Treves, A. & Rolls, E. T. 1999 Correlations and the encoding of information in the nervous system. *Proc. R. Soc. Lond.* B **266**, 1001–1012. (DOI 10.1098/rspb.1999.0736.)

Pashler, H. & Johnston, J. 1998 Attentional limitations in dual-task performance. In *Attention* (ed. H. Pashler). Hove, UK: Psychology Press.

Perrett, D. I. & Oram, M. W. 1998 Visual recognition based on temporal cortex cells: viewer-centred processing of pattern configuration. *Z. Naturforsch.* C **53**, 518–541.

Perrett, D. I., Mistlin, A. J., Chitty, A. J., Smith, P. A., Potter, D. D., Broennimann, R. & Harries, M. 1988 Specialized face processing and hemispheric asymmetry in man and monkey: evidence from single unit and reaction time studies. *Behav. Brain Res.* **29**, 245–258.

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

Prut, Y., Vaadia, E., Bergman, H., Haalman, I., Slovin, H. & Abeles, M. 1998 Spatiotemporal structure of cortical activity: properties and behavioral relevance. *J. Neurophysiol.* **79**, 2857–2874.

Reich, D. S., Victor, J. D., Knight, B. W., Ozaki, T. & Kaplan, E. 1997 Response variability and timing precision of neuronal spike trains *in vivo*. *J. Neurophysiol.* **77**, 2836–2841.

Reich, D. S., Mechler, F. & Victor, J. D. 2001 Temporal coding of contrast in primary visual cortex: when, what, and why. *J. Neurophysiol.* **85**, 1039–1050.

Reynolds, J. H., Chelazzi, L. & Desimone, R. 1999 Competitive mechanisms subserve attention in macaque areas V2 and V4. *J. Neurosci.* **19**, 1736–1753.

Reynolds, J. H., Pasternak, T. & Desimone, R. 2000 Attention increases sensitivity of V4 neurons. *Neuron* **26**, 703–714.

Richmond, B. J. & Sato, T. 1987 Enhancement of inferior temporal neurons during visual-discrimination. *J. Neurophysiol.* **58**, 1292–1306.

Richmond, B. J. & Optican, L. M. 1990 Temporal encoding of two-dimensional patterns by single units in primate primary visual cortex. II. Information transmission. *J. Neurophysiol.* **64**, 370–380.

Richmond, B. J., Optican, L. M., Podell, M. & Spitzer, H. 1987 Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. I. Response characteristics. *J. Neurophysiol.* **57**, 132–146.

Richmond, B. J., Optican, L. M. & Spitzer, H. 1990 Temporal encoding of two-dimensional patterns by single units in primate primary visual cortex. I. Stimulus–response relations. *J. Neurophysiol.* **64**, 351–369.

Richmond, B. J., Oram, M. W. & Wiener, M. C. 1999 Response features determining spike times. *Neural Plasticity* **6**, 133–145.

Riehle, A., Grun, S., Diesmann, M. & Aertsen, A. 1997 Spike synchronization and rate modulation differentially involved in motor cortical function. *Science* **278**, 1950–1953.

Rumelhart, D. E., McClelland, J. L. & PDP Research Group 1986 *Parallel distributed processing: explorations in the microstructure of cognition*. Cambridge, MA: MIT Press.

Sato, T. 1988 Effects of attention and stimulus interaction on visual responses in inferior temporal neurones in macaque. *J. Neurophysiol.* **60**, 344–364.

Shadlen, M. N. & Newsome, W. T. 1994 Noise, neural codes and cortical organization. *Curr. Opin. Neurobiol.* **4**, 569–579.

Shadlen, M. N. & Newsome, W. T. 1995 Is there a signal in the noise? *Curr. Opin. Neurobiol.* **5**, 248–250.

Shadlen, M. N. & Newsome, W. T. 1998 The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J. Neurosci.* **18**, 3870–3896.

Singer, W. & Gray, C. M. 1995 Visual feature integration and the temporal correlation hypothesis. *A. Rev. Neurosci.* **18**, 555–586.

Snowden, R. J., Treue, S. & Andersen, R. A. 1992 The response of neurons in areas V1 and MT of the alert rhesus monkey to moving random dot patterns. *Exp. Brain Res.* **88**, 389–400.

Softky, W. R. & Koch, C. 1993 The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J. Neurosci.* **13**, 334–350.

Thorpe, S. J. 1990 Spike arrival times: a highly efficient coding scheme for neural networks. In *Parallel processing in neural systems and computers* (ed. R. Eckmiller, G. Hartmann & R. Hauske), pp. 91–94. Amsterdam: Elsevier.

Tolhurst, D. J., Movshon, J. A. & Dean, A. F. 1983 The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vis. Res.* **23**, 775–785.

Tovee, M. J., Rolls, E. T., Treves, A. & Bellis, R. P. 1993 Information encoding and the responses of single neurons in the primate temporal visual cortex. *J. Neurophysiol.* **70**, 640–654.

Treisman, A. 1996 The binding problem. *Curr. Opin. Neurobiol.* **6**, 171–178.

Treisman, A. 1998 Feature binding, attention and object perception. *Phil. Trans. R. Soc. Lond.* B **353**, 1295–1306. (DOI 10.1098/rstb.1998.0284.)

Treisman, A. M. & Kanwisher, N. G. 1998 Perceiving visually presented objects: recognition, awareness, and modularity. *Curr. Opin. Neurobiol.* **8**, 218–226.

Treue, S. & Maunsell, J. H. 1996 Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature* **382**, 539–541.

Treves, A., Panzeri, S., Rolls, E. T., Booth, M. & Wakeman, E. A. 1999 Firing rate distributions and efficiency of information transmission of inferior temporal cortex neurons to natural visual stimuli. *Neural Comput.* **11**, 601–632.

Ullman, S. & Soloviev, S. 1999 Computation of pattern invariance in brain-like structures. *Neural Networks* **12**, 1021–1036.

Usher, M. & Neiber, E. 1996 Modeling the temporal dynamics of IT neurones in visual search: a mechanism for top-down selective attention. *J. Cogn. Neurosci.* **8**, 311–327.

Vaadia, E., Haalman, I., Abeles, M., Bergman, H., Prut, Y., Slovin, H. & Aertsen, A. 1995 Dynamics of neuronal interactions in monkey cortex in relation to behavioural events. *Nature* **373**, 515–518.

Van Rullen, R., Gautrais, J., Delorme, A. & Thorpe, S. 1998 Face processing using one spike per neurone. *Biosystems* **48**, 229–239.

Victor, J. D. & Purpura, K. P. 1996 Nature and precision of temporal coding in visual cortex: a metric-space analysis. *J. Neurophysiol.* **76**, 1310–1326.

Villa, A. E. P. & Lorenzana, V. M. B. 1997 Ketamine modulation of the temporal pattern of discharges and spike train interactions in the rat substantia nigra pars reticulata. *Brain Res. Bull.* **43**, 525–535.

Vogels, R., Spileers, W. & Orban, G. A. 1989 The response variability of striate cortical neurons in the behaving monkey. *Exp. Brain Res.* **77**, 432–436.

von Bonin, G. & Bailey, P. 1947 *The neocortex of* Macaca mulatta. Urbana, IL: University of Illinois Press.

von der Malsburg, C. 1995 Binding in models of perception and brain function. *Curr. Opin. Neurobiol.* **5**, 520–526.

von der Malsburg, C. & Schneider, W. 1986 A neural cocktail-party processor. *Biol. Cybern.* **54**, 29–40.

Wallis, G. & Rolls, E. T. 1997 Invariant face and object recognition in the visual system. *Prog. Neurobiol.* **51**, 167–194.

Wicklegren, W. A. 1969 Context-sensitive coding, associative memory and serial order in (speech) behaviour. *Psychol. Rev.* **76**, 1–15.

Wiener, M. C. & Richmond, B. J. 1999 Using response models to estimate channel capacity for neuronal classification of stationary visual stimuli using temporal coding. *J. Neurophysiol.* **82**, 2861–2875.

Wiener, M. C., Oram, M. W. & Richmond, B. J. 1999 Latency is a better temporal code than principal components in V1. *Soc. Neurosci. Abstr.* **25**, 1549.

Wiener, M. C., Oram, M. W., Liu, Z. & Richmond, B. J. 2001 Consistency of encoding in monkey visual cortex. *J. Neurosci.* **21**, 8210–8221.

Xiao, D. K., Edwards, R. H., Bowman, E. M. & Oram, M. W. 2001 The influence of stimulus contrast on response latency and response strength of neurones in the superior temporal sulcus of the macaque monkey. *Soc. Neurosci. Abstr.* **23**, 450.

Zhang, K., Ginzburg, I., McNaughton, B. L. & Sejnowski, T. J. 1998 Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. *J. Neurophysiol.* **79**, 1017–1044.

GLOSSARY

DMS: delayed-match-to-sample
ISI: inter-spike interval
PSTH: post-stimulus time histogram
RT: reaction time
SCM: spike-count matched
TE: inferior temporal visual cortical area in the parcellation of macaque cortex by von Bonin & Bailey (1947)
V1: visual area 1, or primary visual cortex