Taylor & Francis
Taylor & Francis Group

Integrating neuronal coding into cognitive models: Predicting reaction time distributions

MIKE W. ORAM

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

Q1

Abstract

Neurophysiological studies have examined many aspects of neuronal activity in terms of neuronal codes and postulated roles for these codes in brain processing. There has been relatively little work, however, examining the relationship between different neuronal codes and the behavioural phenomena associated with cognitive processes. Here, predictions about reaction time distributions derived from an 10 accumulator model incorporating known neurophysiological data in temporal lobe visual areas of the macaque are examined. Results from human experimental studies examining the effects of changing stimulus orientation, size and contrast are consistent with the model, including qualitatively different changes in reaction time distributions with different stimulus manipulations. The different changes in reaction time distributions depend on whether the image manipulation changes neuronal response 15 latency or magnitude and can be related to parallel or serial cognitive processes respectively. The results indicate that neuronal coding can be productively incorporated into computational models to provide mechanistic accounts of behavioural results related to cognitive phenomena.

Keywords:

Q2

Introduction

20   Understanding visual brain function requires precise knowledge of the information present at each stage of processing, how this information is encoded (the neuronal code), how different signals associated with different stages of processing are combined (integration of neuronal codes) and how neuronal signals are decoded to produce behaviour. While the general properties of many cortical and sub-cortical areas of both human and non-human primate brains 25 are known, the way in which neurones encode inputs and process information to generate output signals is still under debate.

Neuronal codes

As the processing ability of the brain is limited by the number of neurones and their con- 30 nections, the presence of additional neuronal codes (signals) within the response of a single neurone (channel) would increase the total information processing power of the brain without requiring an increase in the number of neurones. Both spike count assessed over hundreds of milliseconds and fine temporal resolution codes ($\sim$1 ms precision) in visual system neuronal responses have been postulated to carry stimulus related information. Response strength is

Q3

Q4

influenced by changes in stimulus orientation, perspective view, object part visibility, size   35
or motion direction (Gross et al. 1972; Oram & Perrett 1992; Oram et al. 1993; Oram &
Perrett 1994, 1996; Perrett et al. 1998; Wiener et al. 2001). Here, the impact of variation of
response magnitude in neurones of the superior temporal sulcus with stimulus orientation
and stimulus size on behaviour is examined.

The precise times of spikes relative to stimulus onset (neuronal response latency) form   40
another neuronal code. Response latency of individual neurons in temporal cortex show
little dependency on response magnitude (Oram & Perrett 1992; Oram et al. 1993; Oram
& Perrett 1994; 1996) whereas changes in stimulus contrast cause changes of up to 400 ms
in response latency, with an average increase of 150 ms when stimulus contrast is dropped
from 100 to 6%, some 3 to 4 times the change observed in primary visual cortex (Oram   45
et al. 2002). Given the dissociation of response magnitude and response latency, it is not
surprising that response latencies, with a temporal resolution of ~10 ms (Reich et al. 2001b),
convey stimulus related information that is unavailable from the spike count (Gawne et al.
1996; Reich et al. 2001a, 2001b; Oram et al. 2002). Here, the impact of neuronal response
latency on behaviour is examined by changing stimulus contrast.   50

In summary, the available neurophysiological data offers support for both coarse (spike
count) and fine (neuronal response latency) temporal resolution neuronal codes as poten-
tial codes involved in brain processing. While precise times of individual spikes within and
between neuronal responses (e.g., synchrony, oscillations and synfire chains) have been pos-
tulated to play a role in many psychological phenomena, including binding or linking of dif-   55
ferent stimulus attributes with behavioural relevance (von der Malsburg & Schneider 1986;
Abeles 1991; Engel et al. 1992; Singer & Gray 1995; von der Malsburg 1995; Prut et al.
1998) the evidence is less clear as to how and under which conditions these codes carry
information unavailable from coarse temporal response measures (Softky & Koch 1993;
Shadlen & Newsome 1994; Shadlen & Newsome 1995; Shadlen & Newsome 1998; Shadlen   60
& Movshon 1999; Oram et al. 1999; Baker & Lemon 2000; Oram et al. 2001; Baker et al.
2001; Oram et al. 2002; Rolls et al. 2003, 2004). This study therefore focuses on the impli-
cations of spike count and neuronal response latency for behavioural phenomena associated
with visual recognition.

Psychological models of visual processing   65

In the 1970s, a linear relationship was found between increasing angular differences and
reaction time to match two images of novel stimuli (Shepard & Metzler 1971). This linear
relationship, combined with the subjects' self-report of feeling as though they were 'rotating
the image in their minds eye', led Shepard & Metzler (1971) to propose subjects use a men-
tal process to rotate and match the input images. This process is hypothesised to perform   70
a smooth continuous transformation until the two images are aligned. When accessing rep-
resentations held in memory the recognition process is assumed to involve transformation
of the input image to match the stored representation or previously seen image rather than
transformation of the stored representation to match the input image (Koriat & Norman
1984; Robertson et al. 1987; Koriat & Norman 1988; Jordan & Huntsman 1995).   75

Linear relationships between reaction time and angular mismatch of stimuli are
widespread, having been observed for pigeons (Hamm et al. 1997), sea-lions (Mauck &
Dehnhardt 1997), baboons (Hopkins et al. 1993) and macaque monkeys (Perrett et al.
1988). Similar relationships occur for recognition of recently learnt objects (Tarr & Pinker
1989; Edelman & Bulthoff 1992; Humphrey & Khan 1992; Bulthoff & Edelman 1992;   80

Logothetis et al. 1994; Vetter et al. 1994; Gauthier & Tarr 1997b; Gauthier et al. 1998) as well as novel stimuli. Highly familiar objects (e.g., faces, car, horse, house, clock) are recognised when viewed in 'canonical' views better than other views (Newell & Findlay 1997). The observed dependency of recognition reaction time and presentation of whole or part object (Biederman 1987; Baker Cave & Kosslyn 1993) can also be interpreted as the mental process of 'object completion'. Image transformations other than rotation also produce changes in reaction time for object recognition. A mental 'size normalisation' process has been proposed to account for the observed relation between image size and recognition reaction time (Bundesen & Larsen 1975; Besner 1983; Jolicoeur & Besner 1987; Ellis et al. 1989; Tombu & Jolicoeur 2002b) and reaction times increase with decreasing stimulus contrast (Williams & Lit 1983; Johannes et al. 1995; Van Selst et al. 1999; Nasanen et al. 2001).

Mental transformation processes have been captured in computational models of object recognition in which either single or multiple orientation, size and viewpoint specific instances of objects are stored (Koenderink & van Doorn 1979; Poggio & Edelman 1990; Edelman & Weinshall 1991; Edelman & Bulthoff 1992; Bulthoff & Edelman 1992). The implementation of the putative mental transformation processes to match input images and multiple viewpoint specific stored representations allows recognition of objects from unusual or novel viewpoints (Tarr & Pinker 1989; Bulthoff et al. 1995; Tarr & Bulthoff 1995). Thus, the predominant view within cognitive frameworks is of a process that implements mental transformations (e.g., rotation) to explain experimental results. In this paper, the effect of changes in orientation, size and contrast on reaction times are modelled using observed changes in neuronal activity to these three image manipulations without invoking mental transformation processes to explain the transformation-reaction time relationships (see also Perrett et al. 1998; Oram et al. 2002).

Serial or parallel processing and the dual task paradigm

A fundamental issue concerning any process is whether or not it operates in parallel with other processes. Within current cognitive frameworks of sensory processing, the initial processing of stimuli is assumed to be parallel (i.e., can occur simultaneously with other processes). Subsequent serial processing is thought to prevent the decision based on one stimulus interfering, possibly catastrophically, with the decision based on a second stimulus. The dual task paradigm is used in psychological experiments to dissociate parallel from serial processes (also referred to pre-bottleneck versus bottleneck, peripheral versus central or slave versus executive processing).

Figure 1 illustrates the rationale of the dual task paradigm and how it can discriminate between parallel and serial processing. Within the dual task paradigm, two stimuli are presented sequentially, each stimulus requiring a behavioural response. To avoid interference, the modality of the two stimuli are presented in different modalities (here, the first stimulus is auditory, the second visual). The auditory stimulus is assumed to undergo some parallel auditory processing and then serial processes (also known as central or bottle-neck processes) leading to a decision which is then processed by the motor system to initiate the behavioural response. The processing of the visual stimulus proceeds in an analogous manner (parallel visual processing, serial decision processing, leading to a behavioural response). Reaction time can be measured to different visual stimuli within this paradigm.

With a long delay between the presentation of the auditory and visual stimuli (Long Inter-Stimulus-Interval in Figure 1), all of the serial decision processing required for the first (auditory) task is completed before the visual stimulus appears, thus allowing uninterrupted processing of the visual stimulus. In Figure 1, visual stimulus 1 leads to shorter reaction time

Long Inter-Stimulus-Interval

Auditory task

| Auditory Processing | Response Selection | Response Initiation |

Visual stimulus 1: Short RT

| Visual Processing | Response Selection | Response Initiation |

Visual stimulus 2: Long RT

Option A

| Visual Processing | Response Selection | Response Initiation |

More parallel processing

Option B

| Visual Processing | Response Selection | Response Initiation |

More serial processing

Short Inter-Stimulus-Interval

| Auditory | Response Selection | Response Initiation |

| Vis Stim 1 Processing | Response Selection | Response Initiation |

| Vis Stim 2 Processing | Response Selection | Response Initiation | A: RT1 = RT2

| Vis Stim 2 Processing | Response Selection | Response Initiation | B: RT1 < RT2
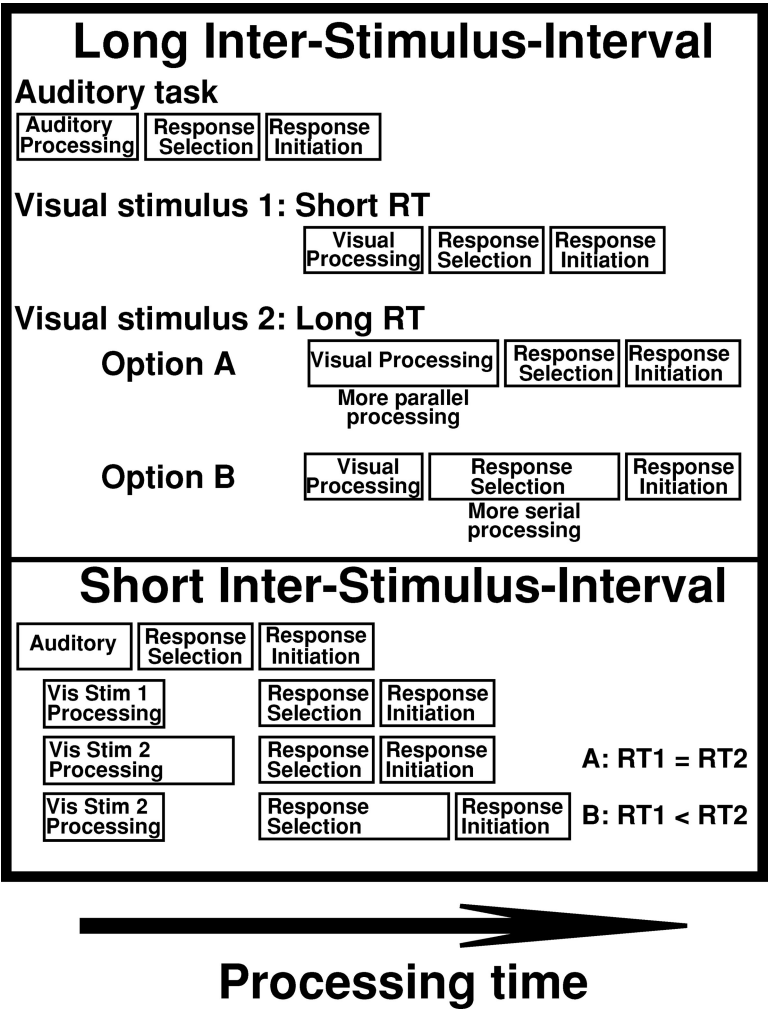
Processing time

Figure 1. Rationale of the dual task paradigm. Within the dual task paradigm, two stimuli are presented sequentially, each stimulus requiring a behavioural response. In this schematic, the first stimulus is auditory, the second visual. The auditory stimulus is assumed to undergo some 'auditory' processing (which can occur in parallel with other tasks), then serial central or bottle-neck processes make a 'decision' which is then processed by the motor system to make the behavioural 'response'. See text for details.

(RT1) than visual stimulus 2 (RT2). There are two options within the cognitive framework as to why visual stimulus 2 gives rise to longer reaction times than visual stimulus 1. One possibility (Option A in Figure 1) is increased parallel processing (i.e., the visual processing of stimulus 2 takes longer than the visual processing of stimulus 1). The second possibility (Option B) is that the serial bottleneck decision processing of stimulus 2 takes longer than for stimulus 1. These options can be distinguished by decreasing the interval between the presentation of the auditory and visual stimuli.

With short inter-stimulus-intervals (Short Inter-Stimulus-Interval in Figure 1), the serial bottle-neck processing of the auditory stimulus is ongoing during the parallel processing of the visual stimulus. This does not interfere with the parallel visual processing but leads to a delay, known as the psychological refractory period, before the serial processing of the second

stimulus can start (i.e., the serial response selection processes associated with the auditory stimulus must complete before the serial processes associated with the visual stimulus can start). If the increased reaction time to visual stimulus 2 compared to visual stimulus 1 is due to increased pre-bottleneck processing time (Option A), then this parallel processing can occur in the psychological refractory period, so RT1 = RT2 at short inter-stimulus-intervals. If the increased reaction time of responses to visual stimulus 2 compared to visual stimulus 1 is due to increased serial processing time (Option B), then the difference in reaction times to the two different stimuli will be the same as at long inter-stimulus-intervals (i.e., RT2 > RT2).

Neural coding and the dual task paradigm

An alternative framework explains results from dual task paradigms in terms of the encoding of different visual attributes by neuronal response latency and magnitude (Oram et al. 2002). The activity of single neurones can be viewed as evidence supporting the presence of a particular visual stimulus (Barlow 1972, 1985). Once the evidence or information exceeds a threshold, a decision about the presence or absence of the stimulus can be made (Loftus et al. 1983; Busey & Loftus 1994; Loftus & Ruthruff 1994b; Chun & Wolfe 1996; Palmer & McLean 1996; Busey & Loftus 1998; Verghese 2001).

The visual information acquisition and evidence accumulation hypotheses are captured by diffusion models and accumulator models (Ward & McClelland 1989; Carpenter & Williams 1995; Palmer & McLean 1996; Perrett et al. 1998; Schall & Thompson 1999; Ratcliff & Rouder 2000; Gold & Shadlen 2001; Usher & McClelland 2001; Verghese 2001; Oram et al. 2002; Schall 2002; Roitman & Shadlen 2002; Schall 2003). Accumulator models integrate the information or evidence in favour of each decision separately and operate as a race: whichever set of accumulated evidence crosses the threshold first determines the decision. Diffusion models accumulate the difference between incoming signals and are therefore competitive in nature. The importance of competitive or inhibitory interactions is highlighted by the fact that accumulator models can only successfully model RT distributions when inhibitory interactions occur between the incoming signals (Smith & Ratcliffe 2004). Competition between different neuronal populations, each population representing a different interpretation of the inputs, can be used to model the behavioural decision about which stimulus is present (e.g., Shadlen et al. 1996; Oram et al. 2002).

Figure 2 shows schematically the accumulation of evidence during a dual task experiment when an image manipulation changes the neuronal response latency (as induced by changes in stimulus contrast (Figure 2 left)) or the neuronal response magnitude (as induced by changes in stimulus orientation (Figure 2 right)). The top panels show schematic plots of neuronal activity over time. The middle row plots the accumulation of evidence when accumulation begins from response onset. These schematic figures illustrate that both increases in neuronal response latency and decreases in neuronal response magnitude predict increased time to threshold (reaction times). The lower panels plot the accumulation of evidence when accumulation is delayed. These lower panels illustrate a condition in which these two alternatives predict differential patterns of reaction time: only changes in response magnitude predict changes in reaction times when accumulation of evidence is delayed. This type of delay is assumed to be induced in the dual task experimental paradigm with short inter-stimulus-intervals (Loftus & Ruthruff 1994a; Ruthruff & Miller 1995; Ruthruff et al. 1995; Jolicoeur 1999; Pashler et al. 2001; Tombu & Jolicoeur 2002a; Tombu & Jolicoeur 2002b; Tombu & Jolicoeur 2003; Ruthruff et al. 2003).
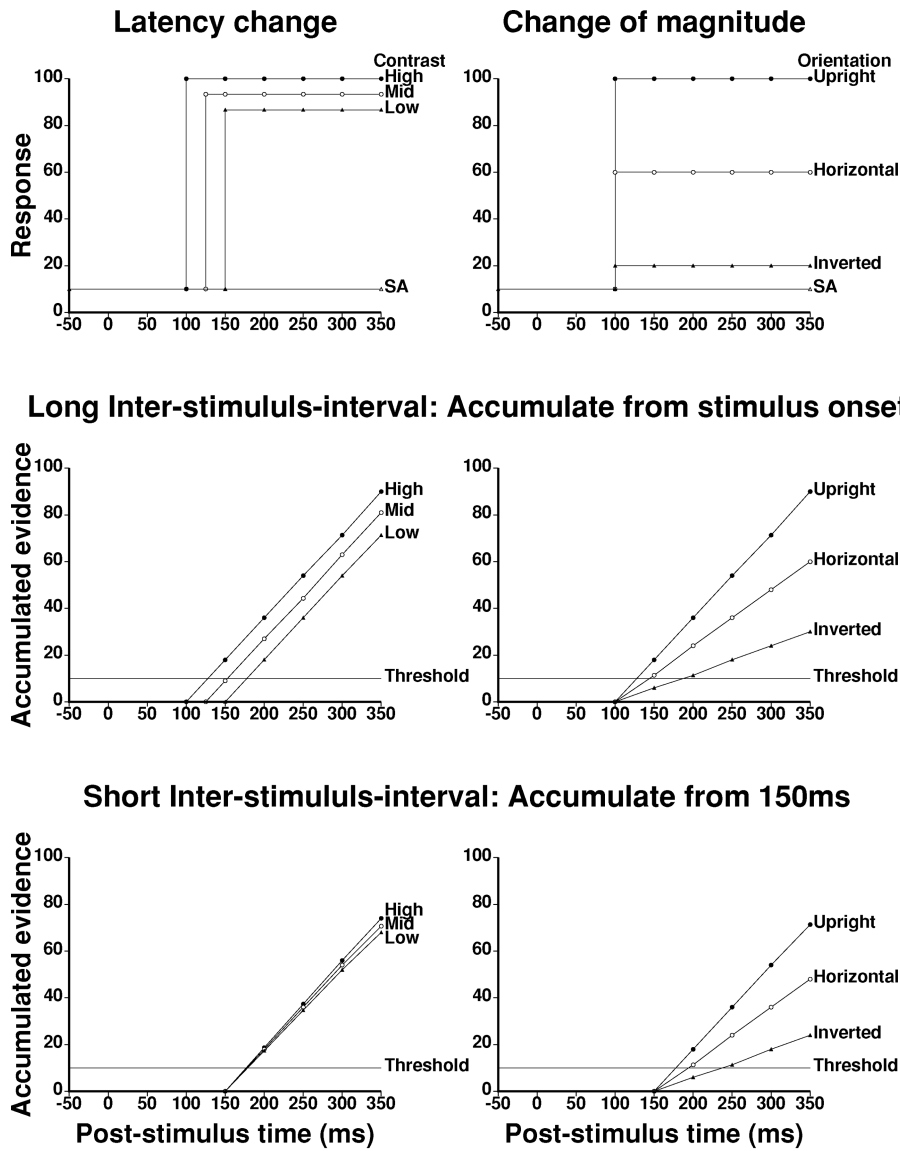
Figure 2. Neuronal coding, accumulator models and dual task paradigms. The schematic figure shows how changes in the neuronal coding of visual representations predict reaction time performance in a dual task paradigm. Left column: predicted effect on reaction times following a change in neuronal response latency induced by changing stimulus contrast. Right column: predicted effect on reaction times following a change in neuronal response magnitude induced by changing stimulus orientation. Top row: schematic of the time course of neuronal activity with a change in response latency (left) or a change in response magnitude (right). Stimulus onset = 0 ms. Middle row: accumulation of evidence when the accumulator starts at stimulus onset. This situation is assumed to occur with long inter-stimulus-intervals in dual task paradigm. The time to threshold (predicted reaction time) increases when either the response latency increases (left) or response magnitude decreases (right). Bottom row: accumulation of evidence when the accumulator starts some period after stimulus onset (here after a delay of 150 ms). This situation is assumed to occur with short inter-stimulus-intervals in dual task paradigm. The time to threshold (predicted reaction time) does not increase as the neuronal response latency increases with, for example, decreasing contrast (left) but does increase when response magnitude decreases (right).

Outline of experiments

The aim of the present study was to further examine the extent to which neuronal codes of different temporal resolution (firing rate assessed over hundreds of milliseconds and neuronal response latency assessed with millisecond precision) can explain the behavioural phenomena associated with recognition of visual presentation of objects from unusual viewpoints (Oram et al. 2002). Both traditional cognitive and diffusion/accumulator frameworks assume a serial process operates in making a behavioural decision/response. While the cognitive framework provides little or no specification of the underlying mechanism, accumulator and diffusion models assume an integration process occurs over time (the accumulation) and when some threshold is reached a decision (i.e., motor output) will be initiated. Neither class of model specifies the mechanisms by which this serial process 'switches' from one task to another.

Unlike cognitive frameworks, diffusion and accumulator models allow assessment of the distribution of the recognition times as well as their means. While the traditional cognitive framework describes processes as serial or parallel according to the results of dual task experiments, the diffusion and accumulator models allow for these effects to be due to changes in the way in which the stimuli are encoded. Diffusion and accumulator models have a long history within cognitive psychology (e.g., Ratcliffe 1978; McClelland 1979), but have not constrained the input signal, typically fitting Gaussian or Poissonian variability of signal strength to human experimental data. Here the distribution of the signals is constrained by known neurophysiological data and the results of both the mean and the distributions of the predicted reaction times are compared to human experimental results.

The dual task paradigm is used to assess whether the time required for a perceptual process is constant with inter-stimulus-interval or not. Diffusion models predict constant perceptual process time across inter-stimulus-interval when the stimulus difference results in changes in neuronal response magnitude, whereas if image manipulation changes neuronal response latency reaction time differences due to the image manipulation will decrease with decreasing inter-stimulus-interval. Previously recorded neurophysiological responses to complex images (including the face and body) from the macaque superior temporal cortex indicate that changes in stimulus orientation and size influence neuronal response magnitude with no impact on response latency (Oram & Perrett 1992; Wachsmuth et al. 1994; Ashbridge et al. 2000), while changes in stimulus contrast influence neuronal response latency and, to a lesser extent, changes in response magnitude (Oram et al. 2002). A simple accumulator model is developed using parameters derived from these neurophysiological data. Predictions about performance in dual task studies are made using the model for transformations of stimulus orientation, size and contrast. To test the model predictions, human experimental dual task studies were performed using these image transformations.

Methods

An accumulator model, incorporating neurophysiological data was used to generate specific predictions that were then tested using human experimental techniques (see also Oram et al. 2002).

Modelling behavioural reaction time

The model used here incorporated known neurophysiological response properties from the macaque temporal lobe of a monkey performing a simple red/green discrimination task when viewing complex images, including the head and body (Perrett et al. 1991; Oram & Perrett
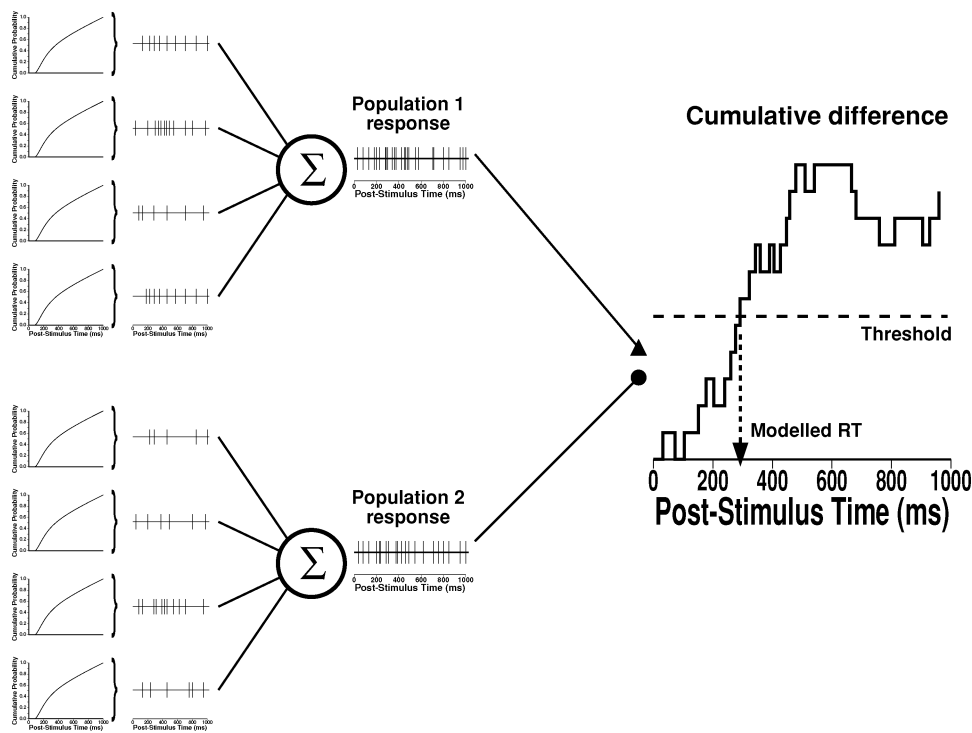
Figure 3. The diffusion model. The mean activity level of each neurone was based known neurophysiological data, with spike times of each neurone were generated using the spike count matched model (see text for details). The cumulative difference of the activity from two populations of simulated neurones was thresholded to obtain a modelled reaction time on a single trial. The mean and distribution of modelled reaction times was obtained from 10000 trials.

1992; Perrett et al. 1998; Ashbridge et al. 2000; Oram et al. 2002). In general, the model has 230 $2..n_p$ populations of input spike trains, one set of inputs representing the 'target' neuronal activity, the other populations of input spike trains representing neuronal activity to n_p-1 'non-target' stimuli (modelled on neuronal responses to 'control stimuli (Perrett et al. 1991). Here, only two neuronal populations are simulated, one representing the target stimulus the second the non-target (Figure 3). The model accumulated the difference in the number of 235 spikes between the 'target' stimulus population and each 'non-target' population. When the absolute value of the accumulated difference reached an arbitrary threshold, a 'decision' was made. If the accumulated difference was positive, a 'target present' decision was made. If the accumulated threshold was negative, a 'target absent' decision was made. The time of the model's 'target' or 'target absent' decision represented the behavioural reaction time.   240

As detailed next, the input spike trains to the model were based on neurophysiological data from neurones in the superior temporal sulcus responsive to the sight of static images of the head and body. Non-linear regression was used to fit an arbitrary function to previously collected neurophysiological data (Oram & Perrett 1992) to obtain appropriate spike density functions (the time course of the activity of the responses of single neurones). Spike density 245 functions obtained from different neurones to stimuli that elicited maximum, mid (50% of maximum) and the smallest response magnitudes were averaged after aligning each spike density function on the neuronal response latency (Oram & Perrett 1992). The averaged

Table I.

	β_1	β_2	τ_2	β_3	τ_3	β_4	τ_4	R^2
Maximum	54.0	5476	48.6	-47.7	286	-5463	49.3	0.91
Mid	38.0	8108	67.8	-7029.0	677803	-8126	68.2	0.77
Minimum	9.6	58	21.1	-10895.0	5200958	-62	49.4	0.56

spike density function was modelled using:

$$Activity = \begin{cases} t >= 0, & \beta_1 + \beta_2 * (1 - e^{-t/\tau_2}) + \beta_3 * (1 - e^{-t/\tau_3}) + \beta_4 * (1 - e^{-t/\tau_4}) \\ t < 0, & 0 \end{cases} \qquad (1)$$

250 where *Activity* is the neuronal response (firing rate in spike per second), t is time relative to response onset (stimulus onset – response latency) and $\beta_1 \ldots \beta_4$, $\tau_2 \ldots \tau_4$ are the coefficients to be estimated. (Inclusion of the third exponential provided a statistically significant improvement ($p < 0.05$) in the R^2 to a sum of two exponentials.) The values for the coefficients are given in Table I.

255   Figure 4 shows neurophysiological data of three levels of response magnitude (Oram & Perrett 1992) with the regression lines superimposed. Linear interpolation between the modelled spike density functions allowed spike density functions to be estimated for any response magnitude. It should be noted that the purpose of the curve fitting was simply to describe the time course of neuronal responses with a high degree of fidelity (see Figure 4 and R^2 values

260 in Table I). While other functional forms would fit the data well using fewer parameters and possibly more stably, the parameters (even though the 95% confidence intervals all include 0) provide a sufficiently good description of the time course that other regressions were not performed (minor changes in the goodness of fit will not substantially alter the results of the simulations).

265   A population response, consistent with observed neurophysiological responses (Perrett et al. 1991; Oram & Perrett 1992; Perrett et al. 1998), was generated from 25 model neurones, with different neurones having different preferred views (i.e., stimulus orientations that
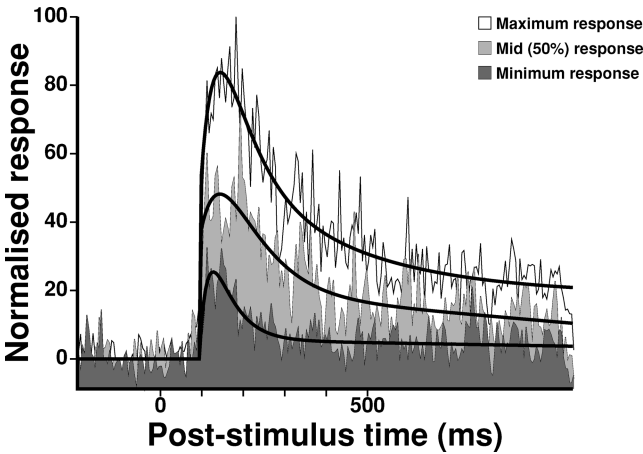

Figure 4. Modelling the time course of neuronal activity. The lines superimposed on previously collected neuronal data indicate the modelled time course of activity (see text for details).

elicited the maximal response (Perrett et al. 1991)). The modelled distribution of response latencies was taken from previously recorded neurophysiological data (Oram & Perrett 1992). The R^2 of the regression of the time courses of the simulated responses on the observed population responses (maximum response magnitude 0.899, mid response magnitude = 0.778, smallest response magnitude = 0.5234) were all highly significant ($p < 0.00001$). Using curve fitting directly on the recorded population data (Perrett et al. 1998) gives only a modest improvement (R^2 maximum = 0.917, mid = 0.809, smallest = 0.592, average improvement of 4% of the variance in the responses 'explained'). The modelled time course of the neuronal responses was therefore in good agreement of observed data.

The mean number of spikes elicited by each stimulus was modelled using the average tuning curves to stimuli of different orientation (Perrett et al. 1991). The trial-by-trial number of spikes elicited by a stimulus was randomly drawn from a truncated Gaussian (Oram et al. 1998; Wiener & Richmond 1999) with a mean-variance relationship described by a fano factor (variance/mean) of 1.5, a value typical of cortical neurones (Mechler et al. 1998; Lee et al. 1998; Oram et al. 1999, 2001). The standard deviation of the predicted reaction times using the observed supra-Poissonian variability, is nearly twice that when Poissonian variability is assumed (data not shown).

The times of individual spikes of each neurone in the population responses was simulated using the spike count matched model (Oram et al. 1999, 2001). Briefly, the spike count matched model uses a standard method to generate random numbers (spike times) with a known probability distribution (spike density function (Press et al. 1992)). The spike times are generated randomly by taking uniform random numbers in the interval (0–1) and applying the inverse of the cumulative probability distribution. Specifically, the spike density function is transformed into a cumulative spike density function (*CSDF*) for each stimulus at each time point t (in general, $CSDF(t) = \sum_{i=1}^{t} SDF(i)$, although here the integral of the modelled spike density functions can be used). Normalization by the value of the *CSDF* at the end of the sample period (T = 1000 ms) gives the cumulative spike probability function $CSPF(t) = \frac{CSDF(t)}{CSDF(T)}$. The *CSPF* gives the probability with which any given spike in a train will have occurred within t ms of stimulus onset. The time bin (width δt, here 1 ms) in which a spike occurs, t_{spike}, in a simulated train is determined from a uniform random distribution $R_{[0..1]}$, such that t_{spike} satisfies $CSPF(k) \leq R_{[0..1]} < CSPF(k+1)$, the time of the kth bin being $k\delta t$ - $(k+1)\delta t$. Only one spike is allowed in each time bin: when a spike was assigned to an already occupied bin, a new random number was drawn and the spike re-assigned. The timing of the spikes in the simulated spike trains can be constrained by the inter-spike-intervals so that both absolute and relative refractory periods of the neurones are incorporated into the simulated data. Here, the constraint was simplified so that no two spikes of the response of an individual model neurone were within 1 ms of each other. The spike count matched model simulates spike trains that have the same fine temporal resolution spike pattern distributions seen in neurophysiological data, including synchronous spikes and synfire chains between responses of different neurones (Oram et al. 1999, 2001).

For each simulated trial, the population response was the summed spike trains across neurones within that population. Estimates of the mean and standard deviation of the modelled reaction times to each stimulus were calculated from 10,000 simulated trials.

Psychological methods

The participants were third-year psychology undergraduates (age 19–22). All participants reported having normal hearing and normal or corrected to normal vision. Participants were given an instruction sheet and presented with an on-screen summary of the procedure.

315   *Procedure.* On each trial, a variable auditory stimulus (1–4 pips, broadband 300 Hz–3 KHz, 25 ms duration, 100 ms between pips) was followed by visual presentation of a single letter covering the fixation point, either 0 or 1000 ms after the end of the auditory stimulus (i.e., 0 or 1000 ms inter-stimulus-interval from the end of the last pip). The task was to decide (1) whether the number of pips was odd or even and (2) whether the letter was a normal

320   or mirror image irrespective of the letter orientation. Data from dual tasks involving mental transformation can be difficult to interpret. For example, some of the processing associated with mental rotation may occur in parallel with tone-frequency discrimination (Van Selst & Jolicoeur 1994; Ruthruff & Miller 1995; Ruthruff et al. 1995; Heil et al. 1999). There is also evidence that aspects of the processing of stimuli and the processing of behavioural responses

325   can occur in parallel despite the serial nature of the two tasks (Hommel 1998). Thus, the preparation of the behavioural response following tone discrimination may interfere with the mental transformation process even if the two tasks were actually being processed serially. Participants were therefore instructed to withhold the response to the auditory stimulus until after making the response to the visual stimulus. Note this also means that if there was

330   an output switching cost, it would affect the responses to the auditory task rather than the visual recognition task. Responses to the stimuli were to be made as quickly as possible using the appropriate key (mirror letter/odd number of pips = '\' key using the left hand; normal letter/even number of pips = '/' key using the right hand). The delay between trials varied uniformly 1000–1500 ms.

335   Each participant underwent a practice phase with on screen messages indicating incorrect responses. The practice phase consisted of only high contrast large stimuli with 1000 ms inter-stimulus-interval. Participants continued the practice phase until achieving criterion level (70% correct, chance = 25%). All participants reached criteria within a 10 minute period. Although subsequent performance increased in the subsequent experimental phase with

340   reaction time steadily decreasing (37 ms every 10 trials, $R^2 = 0.85$) and accuracy increasing (0.006 every 10 trials, $R^2 = 0.72$), this improvement is not considered further here other than noting that reaction times increased with increasing rotation from upright and decreasing contrast at the end as well as the start of the experimental phase. During the experimental phase, responses were made as in the practice phase except no error feedback was given.

345   Blocks of 192 trials (letter R/G by normal/mirror by 8 orientations by 0/1000 ms delay by contrast (high/low)) were presented with a participant controlled pause between blocks. As the practice phase consisted of only upright high contrast stimuli, the rotated and/or low contrast stimuli were novel to the participant.

Stimuli. The auditory and visual stimuli were presented on PCs equipped with 17″ colour
350   monitors and speakers. *Auditory stimuli.* The auditory stimuli were computer generated 'pips' presented well above auditory threshold. For each trial a random number of pips (1–4) were presented at a delay of 100 ms between pips. *Visual stimuli.* The visual stimuli were black (arial font) uppercase letters (R and G) presented in the middle of the monitor on a grey background, rotated $0°$, $±45°$, $±90°$, $±135°$, $180°$ from the upright as either its normal or
355   mirror image. Two contrast levels of each stimulus were generated (high = 100%, low = 6%).

Data analysis. Initial data sorting discarded any data in which the participant's response to the visual or auditory stimulus was incorrect or when the recognition reaction time was more than four standard deviations from that participant's mean reaction time. All data was log-transformed to establish homogeneity of variance (Snedecor & Cochran 1980). As these
360   analyses showed the same pattern of results obtained using the raw data, the results using the raw data values are presented. Analysis was performed using the mean reaction time

of each participant. Due to the experimental time constraints, data were collapsed across letter and direction of rotation (i.e., R/G data and ± rotation angles were combined). Data was analysed using within subjects analysis of variance (ANOVA) (factors: 2 (contrast) × 2 (delay) × 5 (rotation)) using SPSS with each participant's mean reaction times as the independent variable. Type III error terms and Greenhouse–Geisser correction was used for all analyses (see Snedecor & Cochran (1980) for introduction to ANOVA and its relationship to multiple linear regression).

Experiment 2

The participants, drawn from the same student group, were different from those in experiment 1. The methods for experiment 2 were analogous to the first experiment (10 minute practice phase followed by 20 minute experimental phase, see above) with the following changes in the *stimuli*. Two sizes of each stimulus were generated (large = 72 point, small = 18 point) at high (100%) contrast instead of two levels of stimulus contrast.
Procedure. Blocks of 192 trials with the size/orientation changed stimuli rather than the contrast/orientation manipulated images (R/G by normal/mirror by 8 orientations by 0/1000 ms inter-stimulus-interval by size (large/small)). (3) *Analysis.* The three way within subjects ANOVA, after collapsing across letter and direction of rotation had the design (2 (size) × 2 (delay) × 5 (rotation)).

Results

Integration of neuronal codes into the accumulator model (see also Oram et al. (2002)) predicts that if changes in perceptual decision reaction times are driven by changes in the response latency of the underlying neuronal representation (experimentally induced by changes in stimulus contrast) then the recognition reaction time is influenced only with long inter-stimulus-intervals. On the other hand, changes in predicted recognition reaction time due to change in response magnitude (experimentally induced by changes in stimulus orientation) are predicted to be independent of inter-stimulus-interval (Figure 5).
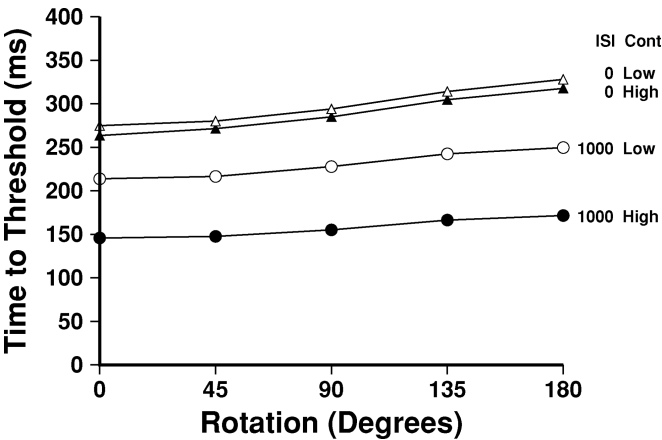

Figure 5. Predicted patterns of mean reaction times in a dual task paradigm. The mean reaction times predicted using the diffusion model at long and short inter-stimulus-intervals are plotted when stimulus rotation from upright and contrast (High/Low) are varied. The pattern of results predicts an almost constant linear increase in reaction time as the stimulus is rotated. The effect of changing stimulus contrast on reaction times is predicted to be larger with long than with short inter-stimulus-interval.
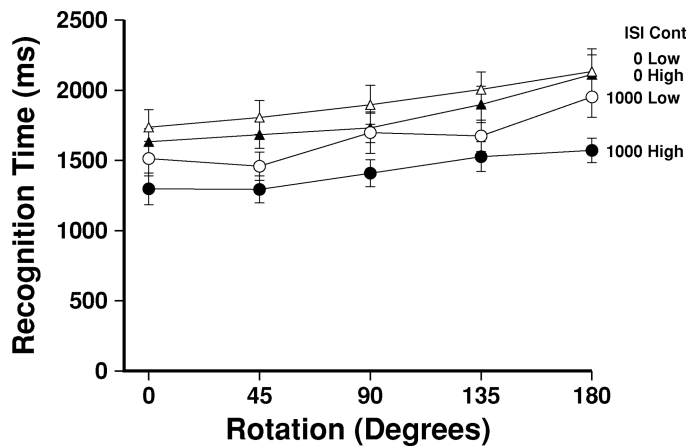
Figure 6. Mean reaction times in a dual task paradigm with changes in stimulus orientation and contrast. The mean reaction times from 23 participants (error bars = standard error of the mean) are plotted. See text for details. Results of ANOVA: effect of contrast: mean $F_{[1,22]} = 25.9$, $p < 0.00005$; effect of inter-stimulus-interval: $F_{[1,22]} = 64.5$, $p < 0.00005$; effect of stimulus rotation: $F_{[4,88]} = 37.3$, $p < 0.00005$. Critically, the effect of stimulus contrast interacted with inter-stimulus-interval $F_{[1,22]} = 28.4$, $p = 0.001$. No other significant interactions were observed in the experimental data (each $p > 0.1$).

Consistent with previous dual task studies examining the impact of changes in orientation (Ruthruff et al. 1995; Pashler et al. 2001; Oram et al. 2002; Ruthruff et al. 2003) it is predicted that the impact of inversion on RT will be constant across inter-stimulus-interval. On the other hand, the impact of changes in stimulus contrast on RT are predicted to decrease at short inter-stimulus-intervals (Van Selst et al. 1999; Oram et al. 2002; Oriet & Jolicoeur 2003). The mean reaction times in the experimental study (Figure 6) were in agreement with the predicted pattern of results. The mean reaction time at short inter-stimulus-intervals (mean ± sem: 1864 ± 115 ms) was significantly ($p < 0.00005$) longer than the mean reaction time with long inter-stimulus-intervals (1540 ± 103 ms). Mean reaction time increased significantly with the increase of rotation angle ($p < 0.00005$). Critically, the effect of stimulus rotation on mean reaction time was statistically the same at both inter-stimulus-intervals ($p > 0.7$). The effect of reducing stimulus contrast on mean reaction time with a 0 ms inter-stimulus-interval between presentation of the auditory and visual stimuli was smaller (104 ms) than with a 1000 ms inter-stimulus-interval (240 ms, $p = 0.001$). Consistent with the predictions from the diffusion model, no other significant interactions were observed in the experimental data (each $p > 0.1$).

It is possible that results obtained in the first experiment occur because visual processing of rotated stimuli interacts with other visual processes such that they are dependent on the inter-stimulus-interval. Therefore, experiment 2 used stimuli of different orientations and sizes rather than different orientations and contrasts. Changes in stimulus size reduce neuronal response magnitude without change in neuronal response latency (see Introduction). Thus, like changes in stimulus orientation, the effect on reaction time is predicted to be independent of the inter-stimulus-interval. Figure 7 plots the mean reaction times to two image sizes at 0 and 1000 ms inter-stimulus-intervals at different stimulus orientations. As with the first experimental study, mean reaction times at short inter-stimulus-intervals were longer than at long inter-stimulus-intervals and there was a main effect of stimulus rotation (each $p < 0.00005$). The increase in mean reaction time with increasing rotation from upright was statistically equivalent at both delays ($p > 0.5$). Mean reaction times to small images (1725 ± 82 ms) were significantly ($p < 0.00005$) greater than reaction times to large images
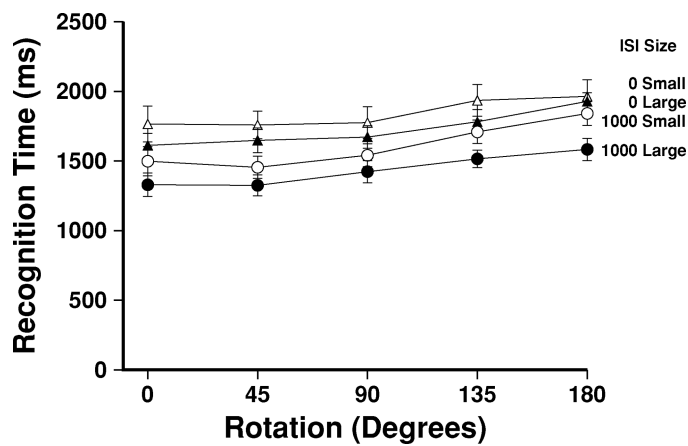
Figure 7. Mean reaction times in a dual task paradigm with changes in stimulus orientation and size. The mean reaction times from 15 participants (error bars = standard error of the mean) are plotted. See text for details. Results of ANOVA: effect of image size $F_{[1,14]} = 42.3$, $p < 0.00005$; effect of inter-stimulus-interval $F_{[1,14]} = 43.0$, $p < 0.00005$; effect of stimulus rotation $F_{[4,56]} = 13.9$, $p < 0.00005$; no significant interactions were found (each $p > 0.2$).

(1582 ± 67 ms). No significant interactions were found (each $p > 0.2$), indicating that the effect of change in stimulus size on mean reaction time was independent of the inter-stimulus-interval (see also Tombu & Jolicoeur (2002b)).

The model used here predicts not only the mean but also the distribution of the trial-by-trial reaction times. While the model provides estimates of the reaction time distributions for incorrect responses, the average accuracy of the subjects' responses was over 95%, providing insufficient trials to analyse the reaction time distributions for incorrect trials (see Smith & Ratcliffe (2004) and references therein for review of diffusion and accumulator models predicting incorrect trial reaction time distributions). Figure 8 plots the distributions of reaction times from the human experimental (top row) and modelling studies (bottom row). The reaction time distributions from both the modelling and human studies show a 'skewing' of the distribution with presentation of stimuli away from upright (left column) or decreasing size (right column) and a rightward 'shift' with decreasing stimulus contrast (middle column).

The reaction time distributions predicted from the model show a linear mean:standard deviation relationship over the neurophysiologically plausible range of activity levels and response latencies. With a delay before the start of accumulation of spike counts the mean time-to-thresholds increase but the slope of the mean~standard deviation relationship remains nearly constant. Consistent with this, the coefficient of variation (ratio of mean to standard deviation (see Figure 9)) of human reaction times with long inter-stimulus-interval is greater than that seen with short inter-stimulus-interval ($F_{[1,80]} = 36.2$, $p < 0.0005$, Figure 9 bottom left). Similarly, an increase in response latency is predicted to cause an increase in mean reaction time with little change in the trial-by-trial variability. The predicted increase in the coefficient of variation as stimulus contrast decreased (increased neuronal response latency) was also observed ($F_{[1,49]} = 10.1$, $p = 0.002$, see Figure 8 middle column and Figure 9 bottom right).

The model predicts qualitatively different changes in the distribution of reaction times when neuronal response magnitudes are changed. The probability of detection threshold being reached rapidly (early reaction time) is decreased with decreasing response magnitude with corresponding increases in the probability of longer reaction times. This skewing of the reaction time distribution, rather than a shift, is reflected in the constant coefficient
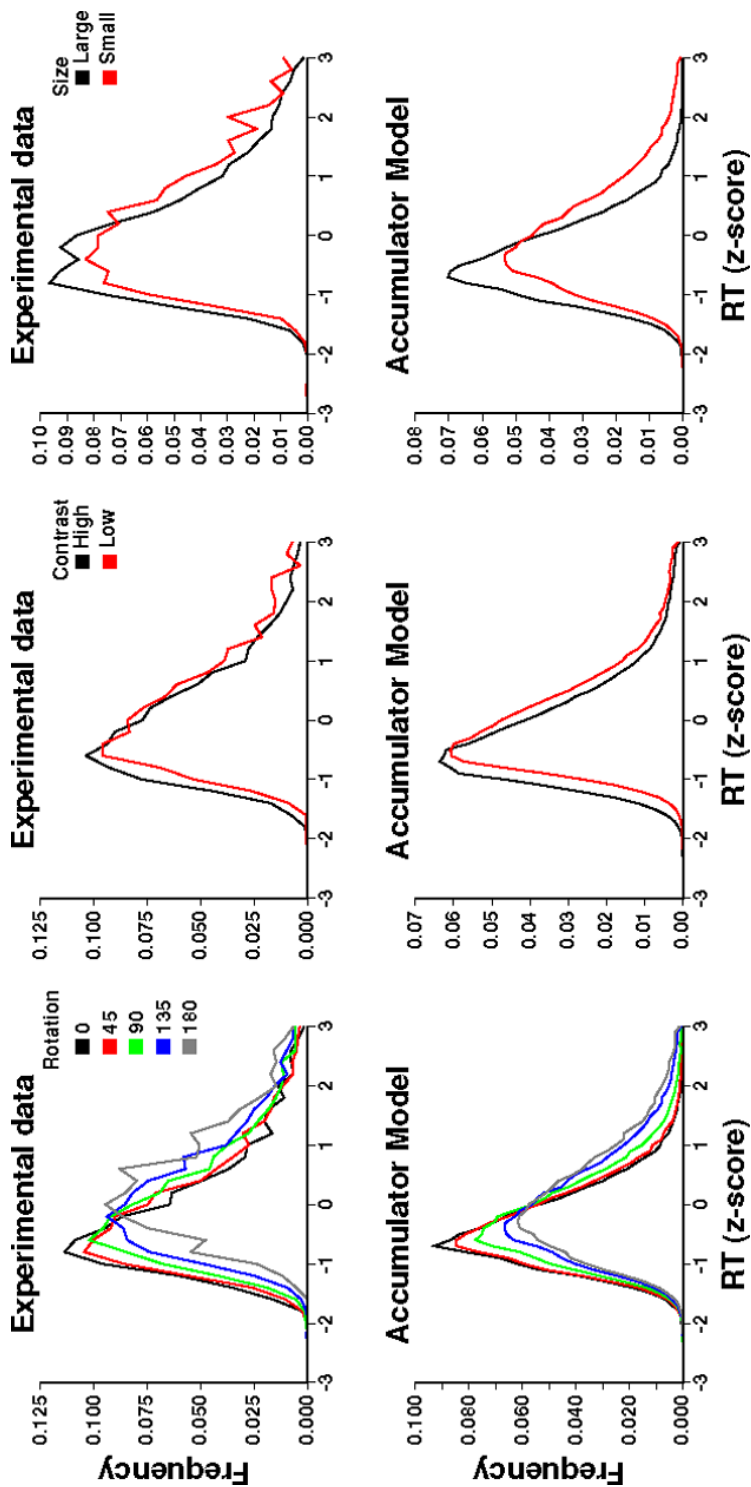
Figure 8. Changes in reaction time distributions with change in stimulus orientation, contrast and size. The trial-by-trial reaction time data was converted into a z-score ((trial RT—mean)/SD) for each participant. Left column: the distributions when stimulus orientation changes (upright, 45,90,135 and inverted). Middle column: the distributions when stimulus contrast was changed (high = 100%, low = 6%). Right column: the distributions when stimulus size change (large = 72 point, small = 12 point). Top row: reaction time distributions from human experimental studies. Bottom row: reaction time distributions from the diffusion model. The main features to note are that (1) the reaction time distributions have similar shape; (2) changes in stimulus orientation (left) or size (right) induce skewing of the distributions; (3) changes in stimulus contrast (centre) and short inter-stimulus-interval (not shown) induce shifts of the distributions as well as slight skewing.
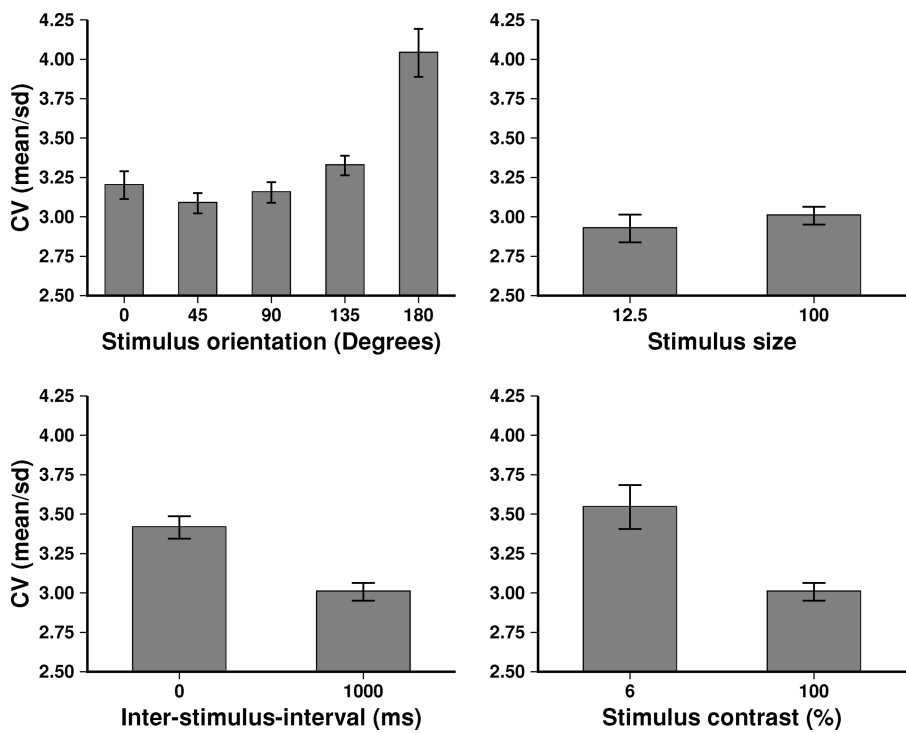
4C/Art

16   *M. W. Oram*

Figure 9. Coefficient of variation for changes in experimental condition. The coefficient of variation averaged across participants is shown for different experimental manipulations (stimulus orientation, stimulus size, inter-stimulus-interval and stimulus contrast).

of variation seen with changes in stimulus size ($F_{[1,30]} = 1.2$, $p > 0.2$, see Figure 8 right and Figure 9 top right) and with changes in the stimulus orientation (Figure 8 left and Figure 9 top left) up to rotations of 135 ($p > 0.05$ each comparison) although there is an unpredicted increase in the CV when the stimulus is inverted (Figure 9 top left). However, given the predicted low probability of observing the shortest reaction times when the stimulus is inverted this may be due to under-sampling due to experimental time constraints. 450

Discussion

Summary of results

A model was outlined using accumulation of spike counts with known relationships between 455 neuronal codes and image transformations to predict changes in reaction time distributions under a variety of conditions. Changes in either neuronal response magnitude or response latency of an encoded visual stimulus predict changes in mean reaction times to that stimulus. A dual task paradigm was used to tease apart the predicted effects of changes in neuronal response latency from changes in response magnitude on mean reaction times. As predicted 460 (Figure 5), effects of changing stimulus contrast (neuronal response latency) depended on inter-stimulus-interval, whereas changes in stimulus orientation and size (neuronal response magnitude) were independent of inter-stimulus-interval (Figures 6 and 7).

The model also predicted that the changes in the mean reaction times would be due to different types of change in the underlying reaction time distributions. Changes in mean 465

reaction times due to changes in neuronal response magnitude (changing stimulus orientation and size) were predicted and generally consistent with the predicted skewing of the reaction time distributions (Figures 8 and 9). The analysis of the co-efficient of variation indicated that the change in reaction time distribution of inverted stimuli may involve a 'shift' as well as
470 a skewing (Figure 9). However, this can be attributed to the very low probability of observing the shortest reaction times to inverted images and the limited number of experimental trials that can be collected. Changes in mean reaction times due to changes in neuronal response latency (changing stimulus contrast) and the impact of changing the inter-stimulus-interval were, as predicted, due to shifts of the reaction time distributions (Figures 8 and 9).

475 *Use of neuronal codes in models*

It is noteworthy that fitting model parameters to spike counts from a completely different task still leads to a good match between the predicted response time distributions and the ones measured in the psychophysical experiments. Studies that have examined the impact of behavioural tasks on neuronal responses have reported modulation of the response: the
480 attentional 'gain' (Morane & Desimone 1985; McAdams & Maunsell 1999) is thought to be implemented by biasing inhibitory interactions between different neuronal representations (Reynolds et al. 1999, 2003). When actively searching for a visual stimulus in a complex scene, the size of the neurone's functional receptive field is reduced but without change in stimulus selectivity (Sheinberg & Logothetis 2001; Aggelopoulos et al. 2005; see also
485 Morane & Desimone 1985). These results indicate that the impact of task specific effects is a modulation of neuronal responses rather than a change in the selectivity. While it will be important to examine neuronal responses under appropriate task conditions, the data to date suggest that task induced changes will influence the absolute values of the predictions of accumulator models rather than the pattern of the predictions.
490 Neurophysiological studies on the relationship between behavioural reaction time and neuronal response latency have shown that the variability of response latency in the inferior temporal cortex is typically significantly smaller than the variability of behavioural reaction time (DiCarlo & Maunsell 2005). This does not mean, however, that neuronal activity in the visual system and behavioural reaction times are unrelated. In a binocular rivalry task, the
495 time of reported target perception is correlated with the onset of activity of inferior temporal cortical neurones (Sheinberg & Logothetis 1997). During a visual search, the magnitude of responses from neurones is inversely correlated with time-to-detection (i.e., the larger a neuronal response the shorter the time before target detection (Sheinberg & Logothestis 2001)). Thus, analysis of response magnitude and response latency is needed to determine
500 whether there is a relationship between encoded activity and behavioural reaction time. For the model used here, the time for the difference in activity between two populations to reach threshold is related to both the response latency and magnitude and captures much of the variability in behavioural reaction times.

Comparison of cognitive and neurophysiologically based models

505 In line with cognitive models, the neurophysiologically based computational model used here assumed a delay before accumulating the difference of spike counts between two population responses when short inter-stimulus-intervals were used. The neuronal mechanisms underlying this delay were not modelled, nor do the present results shed light on this issue. The assumed delay received at least some experimental support in that the change in mean
510 reaction time between long and short inter-stimulus-intervals was due to a shift rather than a skewing of the underlying reaction time distribution (Figure 9).

Within a cognitive framework, the linear increase in mean recognition reaction times as the image is rotated from upright forms the hallmark of mental rotation (Shepard & Metzler 1971). Similarly, the increase in mean reaction times as stimulus size is decreased is taken as indicating of cognitive process of size normalisation (Tombu & Jolicoeur 2002b). It is shown 515 here that incorporation of neurophysiological data into diffusion or accumulator models provides an alternative explanation of the behavioural results (see also Perrett et al. 1998; Oram et al. 2002). The diffusion model used here does not postulate secondary mechanisms (mental rotation or size normalisation) operating on the neuronal representations and as such gives a more parsimonious explanation of the behavioural data. 520

The studies presented here also suggest an alternative explanation of the dual task experiments. Cognitive frameworks assume that some aspects of sensory processing occur in parallel while others are serial (Ruthruff & Miller 1995; Ruthruff et al. 1995; Pashler et al. 2001; Tombu & Jolicoeur 2003). The model used here suggests that if a stimulus manipulation influences neuronal response latency, the effect of that manipulation on reaction time will 525 depend on inter-stimulus-interval. If the image manipulation influences response magnitude but not response latency, the effect of that manipulation on reaction time will be independent of inter-stimulus-interval. Furthermore, within the outlined cognitive frameworks, the patterns of reaction time data cannot be predicted *a priori* as a dual task experiment is used to define whether the process is serial or parallel. The model presented here allows measures 530 of neuronal activity with changes in stimulus parameters to be used to predict the pattern of behavioural reaction times in a dual task paradigm without invoking serial or parallel processes.

Accumulator and diffusion models (Ward & McClelland 1989; Shadlen & Newsome 1996; Perrett et al. 1998; Schall & Thompson 1999; Gold & Shadlen 2001; Usher & McClelland 535 2001; Oram et al. 2002; Schall 2003) predict the patterns of reaction time distributions as well as the mean reaction times under various experimental conditions. The prediction of differential patterns of changes in reaction time distributions with different types of change in stimuli is unavailable from most cognitive models or computational models that do not incorporate differential encoding (changes in magnitude and latency). A change in neuronal 540 coding with changing input can influence behaviour wherever the change in coding occurs: manipulations that change magnitude and latency could occur at the same points in the processing stream yet would give rise to differential dual task performance. Thus, and unlike most models of dual task performance, the model used here does not assume different 'stages' of processing (e.g., sensory coding, response selection and response initiation). Thus, 545 diffusion or accumulator models, when incorporating neurophysiological data, offer a more complete and more parsimonious explanation of behavioural results than is available from traditional cognitive models.

An alternative framework for interpreting results from dual-task experiments

The available neurophysiological data from visual association cortices indicates encoding of 550 visual stimuli is not invariant to image transformations such as changes in orientation, size or contrast (e.g., Gross et al. 1972; Desimone et al. 1984; Rolls & Baylis 1986; Tanaka et al. 1991; Perrett et al. 1998; Oram et al. 2002). The idea that anisotropies in population level encoding of stimuli may underlie perceptual performance is not new. For example, meridian anisotropy—being faster and more accurate at determining upright and horizontal lines than 555 oblique lines—is attributed to more neurones being maximally activated by vertical and horizontal lines than oblique lines (Regan & Beverley 1985; Heeley & Timney 1988) although see (Heeley & Buchanan-Smith 1992). The studies here show that changes in reaction times

Long Inter-Stimulus-Interval

Auditory task

Ongoing auditory processing

Threshold?
Response

Visual stimulus 1: Short RT

Ongoing Visual processing

Threshold?
Response

Visual stimulus 2: Long RT

RT2 > RT1 with Increased latency / Decreased magnitude

Ongoing Visual processing

Threshold?
Response

Short Inter-Stimulus-Interval

Ongoing auditory processing

Threshold?
Response

Ongoing Visual processing (Stim 1)

Threshold?
Response

Ongoing Visual processing (Stim 2)

Threshold?
Response

Increased latency
RT2 = RT1

Ongoing Visual processing (Stim 2)

Threshold?
Response

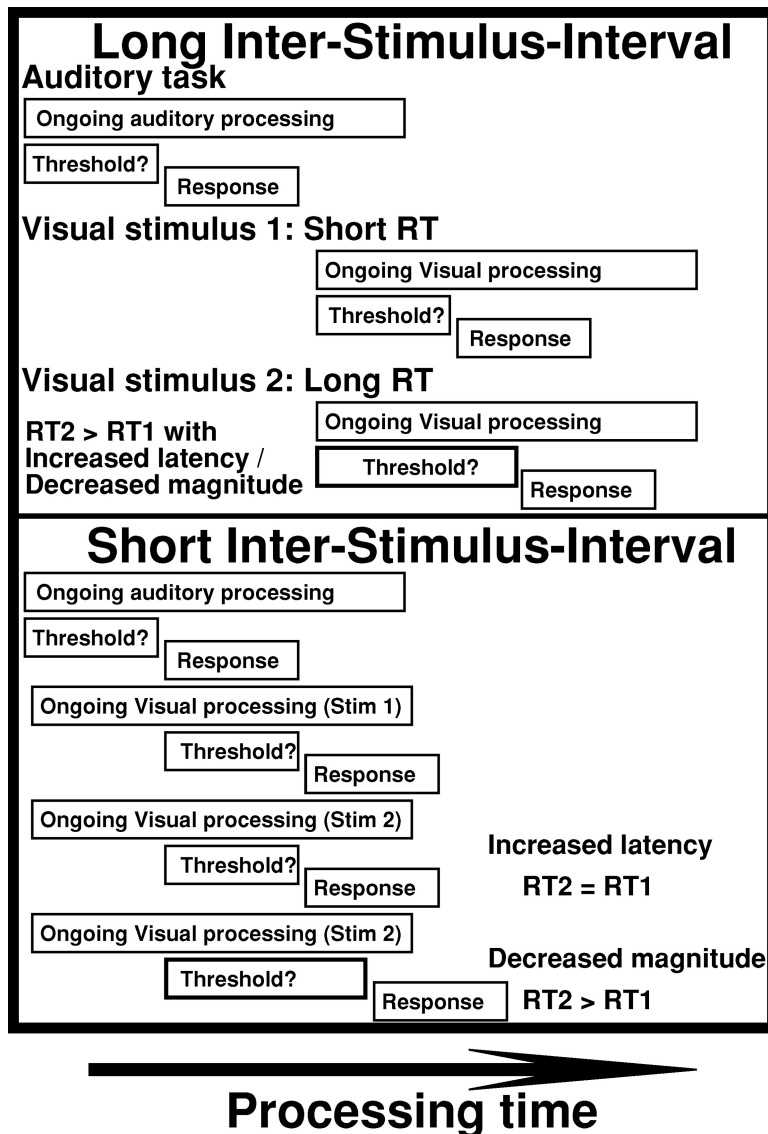**Decreased magnitude**
RT2 > RT1

Processing time

Figure 10. Alternative framework of dual task paradigm performance. As in the bottleneck model (see Figure 1), presentation of an auditory stimulus is followed by presentation of a visual stimulus at either long (Upper) or short (Lower) inter-stimulus-intervals. Upper: visual stimulus 1 leads to short reaction time and visual stimulus 2 which gives rise to long reaction times as a result of either an increase in neuronal response latency (delayed signal) or reduced neuronal response magnitude (reduced signal strength) compared to the neuronal responses elicited by stimulus 1. Lower: with short inter-stimulus-intervals, the serial 'bottle-neck' apparatus is assumed to be processing the first stimulus when the second stimulus is presented. Thus there is a delay before the accumulation of evidence (information) starts. If the change in behavioural reaction time was due to a change in neuronal response latency of the encoded visual stimulus (e.g., change in stimulus contrast), reaction times to stimulus 2 (low contrast) will now be comparable to reaction times to stimulus 1 (high contrast). If the change in behavioural reaction time was due to a change in neuronal response magnitude (e.g., change in stimulus orientation or size), reaction times to stimulus 2 (inverted/small size) will remain longer than reaction times to stimulus 1 (upright/ large size).

predicted here may not require specific cognitive processes (e.g., mental rotation, size normalisation) with specific properties (serial versus parallel) but might be, like explanations of meridian anisotropy, a consequence of the way in which the representation of the stimulus input is encoded.

Finally, the model used here assumes that serial processes are only necessary for response selection and initiation to avoid possibly catastrophic interference. At long inter-stimulus-intervals, monitoring and evidence accumulation, based on neuronal activity, can start at the onset of the visual stimulus. At long inter-stimulus-intervals, therefore, an image manipulation that results in a decrease in neuronal response magnitude or an increase in neuronal response latency will result in increased mean reaction times (Figure 10, upper). At short inter-stimulus-intervals (Figure 10, lower), the accumulation of evidence is delayed until the decision about first task is complete. This will result in differential reaction times only if the image manipulation evokes a change in response magnitude (e.g., rotation away from upright or decrease in stimulus size). If the image manipulation changes neuronal response latency (e.g., decreasing stimulus contrast) differential reaction times will be observed at long but not short inter-stimulus-intervals.

Summary

The experiments here were designed to test specific hypotheses generated from consideration of neurophysiological activity within the visual system. The results confirmed that if a stimulus parameter variation produces a change in reaction time that is constant across different inter-stimulus-intervals (orientation and/or size) it is sufficient to attribute this to changes in the neuronal response strength without the need to postulate a separate serial mental process (mental rotation/size normalisation). On the other hand, if a stimulus parameter variation produces a change in reaction time that varies with inter-stimulus-intervals (contrast) it is sufficient to attribute this to changes in the neuronal response latency without the need to postulate a separate parallel mental process (contrast normalisation).

Not only do diffusion or accumulator models incorporating neurophysiological data provide a simple explanation of behavioural phenomena associated with 'mental rotation', 'size normalisation' and 'contrast normalisation' without recourse to specific transformation processes (Perrett et al. 1998; Oram et al. 2002), but also a parsimonious account of results from dual task paradigms. The power of the model extends beyond predicting patterns for mean reaction times, offering an explanation of changes in the distributions of reaction times seen in visual tasks. Thus, these studies indicate a fruitful interchange between mathematical models and neuronal coding that can lead to a more parsimonious explanation of behaviour than is available from current cognitive models.

Acknowledgments

References

Abeles M. 1991. Corticonics. Cambridge: Cambridge University Press.

Ashbridge E, Perrett DI, Oram MW, Jellema T. 2000. Effect of image orientation and size on object recognition: Responses of single units in the macaque monkey temporal cortex. Cog Neuropsych 17:13–34.

Baker Cave C, Kosslyn SM. 1993. The role of parts and spatial relations in object identification. Perception 22:229–248.

Baker SN, Lemon RN. 2000. Precise spatiotemporal repeating patterns in monkey primary and supplementary motor areas occur at chance levels. J Neurophysiol 84:1770–1780.

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

Barlow HB. 1972. Single units and sensation: A neuron doctrine for perceptual psychology? Perception 1:371–394.

Barlow HB. 1985. The twelfth Bartlett memorial lecture: The role of single neurons in the psychology of perception. Q J Exp Psychol A 37:121–145.

610 Besner D. 1983. Visual pattern recognition: Size preprocessing re-examined. Q J Exp Psychol A 35:209–216.

Biederman I. 1987. Recognition-by-components: A theory of human image understanding. Psychol Rev 94:115–147.

Bredfeldt CE, Ringach DL. 2002. Dynamics of spatial frequency tuning in macaque V1. J Neurosci 22:1976–1984.

Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA. 1996. A relationship between behavioral choice
615 and the visual responses of neurons in macaque MT. Vis Neurosci 13:87–100.

Bulthoff HH, Edelman S. 1992. Psychophysical support for a two-dimensional view interpolation theory of object recognition. Proc Natl Acad Sci USA 89:60–64.

Bulthoff HH, Edelman SY, Tarr MJ. 1995. How are three-dimensional objects represented in the brain? Cereb Cortex 5:247–260.

620 Bundesen C, Larsen A. 1975. Visual transformation of size. J Exp Psychol [Hum Percept Perform] 1:214–220.

Busey TA, Loftus GR. 1994. Sensory and cognitive components of visual information acquisition. Psychol Rev 101:446–469.

Busey TA, Loftus GR. 1998. Binocular information acquisition and visual memory [In Process Citation]. J Exp Psychol [Hum Percept Perform] 24:1188–1214.

625 Carpenter RHS, Williams MLL. 1995. Neural computation of log likelihood in control of saccadic eye-movements. Nature 377:59–62.

Chun MM, Wolfe JM. 1996. Just say no: How are visual searches terminated when there is no target present? Cog Psychol 30:39–78.

Desimone R, Albright TD, Gross CG, Bruce C. 1984. Stimulus-selective properties of inferior temporal neurons
630 in the macaque. J Neurosci 4:2051–2062.

Diwadkar VA, McNamara TP. 1997. Viewpoint dependence in scene recognition. Psychol Sci 8:302–307.

Edelman S, Bulthoff HH. 1992. Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. Vis Res 32:2385–2400.

Edelman S, Weinshall D. 1991. A self-organizing multiple-view representation of 3D objects. Biol Cybern 64:209–
635 219.

Ellis R, Allport DA, Humphreys GW, Collis J. 1989. Varieties of object constancy. Q J Exp Psychol A 41:775–796.

Engel AK, Konig P, Kreiter AK, Schillen TB, Singer W. 1992. Temporal coding in the visual cortex: New vistas on integration in the nervous system. Trends Neurosci 15:218–226.

Gauthier I, Tarr MJ. 1997a. Becoming a 'Greeble' expert: Exploring mechanisms for face recognition. Vis Res
640 37:1673–1682.

Gauthier I, Tarr MJ. 1997b. Orientation priming of novel shapes in the context of viewpoint-dependent recognition. Perception 26:51–73.

Gauthier I, Williams P, Tarr MJ, Tanaka J. 1998. Training 'greeble' experts: A framework for studying expert object recognition processes. Vis Res 38:2401–2428.

645 Gawne TJ, Kjaer TW, Richmond BJ. 1996. Latency: Another potential code for feature binding in striate cortex. J Neurophysiol 76:1356–1360.

Geisler WS, Albrecht DG. 1995. Bayesian analysis of identification performance in monkey visual cortex: Nonlinear mechanisms and stimulus certainty. Vis Res 35:2723–2730.

Gold JI, Shadlen MN. 2001. Neural computations that underlie decisions about sensory stimuli. Trends Cog Sci
650 5:10–16.

Gross CG, Rocha-Miranda CE, Bender DB. 1972. Visual properties of neurons in inferotemporal cortex of the Macaque. J Neurophysiol 35:96–111.

Hamm J, Matheson WR, Honig WK. 1997. Mental rotation in pigeons (*Columba livia*)? J Comp Psychol 111:76–81.

Heeley DW, Buchanan-Smith HM. 1992. Orientation acuity estimated with simultaneous and successive proce-
655 dures. Spat Vis 6:1–10.

Heeley DW, Timney B. 1988. Meridional anisotropies of orientation discrimination for sine wave gratings. Vis Res 28:337–344.

Heil M, Rosler F, Link M, Bajric J. 1998. What is improved if a mental rotation task is repeated-the efficiency of memory access, or the speed of a transformation routine? Psychol Res 61:99–106.

660 Heil M, Wahl K, Herbst M. 1999. Mental rotation, memory scanning, and the central bottleneck [In Process Citation]. Psychol Res 62:48–61.

Hietanen JK, Perrett DI, Oram MW, Benson PJ, Dittrich WH. 1992. The effects of lighting conditions on responses of cells selective for face views in the macaque temporal cortex. Exp Brain Res 89:157–171.

Hommel B. 1998. Automatic stimulus-response translation in dual-task performance. J Exp Psychol [Hum Percept Perform] 24:1368–1384.

Hopkins WD, Fagot J, Vauclair J. 1993. Mirror-image matching and mental rotation problem solving by baboons (*Papio papio*): Unilateral input enhances performance. J Exp Psychol [Gen] 122:61–72.

Humphrey GK, Khan SC. 1992. Recognizing novel views of three-dimensional objects. Can J Psychol 46:170–190.

Johannes S, Munte TF, Heinze HJ, Mangun GR. 1995. Luminance and spatial attention effects on early visual processing. Cog Brain Res 2:189–205.

Jolicoeur P. 1999. Dual-task interference and visual encoding. J Exp Psychol [Hum Percept Perform] 25:596–616.

Jolicoeur P, Besner D. 1987. Additivity and interaction between size ratio and response category in the comparison of size-discrepant shapes. J Exp Psychol [Hum Percept Perform] 13:478–487.

Jordan K, Huntsman LA. 1995. Repetition of misoriented words reduces lexical-decision response-times at repeated and novel orientations. J Exp Psychol [Hum Percept Perform 21:963–977.

Knierim JJ, Van Essen DC. 1992. Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. J Neurophysiol 67:961–980.

Koenderink JJ, van Doorn AJ. 1979. The internal representation of solid shape with respect to vision. Biol Cybern 32:211–216.

Koriat A, Norman J. 1984. What is rotated in mental rotation? J Exp Psychol [Learn Mem Cogn] 10:421–434.

Koriat A, Norman J. 1988. Frames and images: Sequential effects in mental rotation. J Exp Psychol [Learn Mem Cogn] 14:93–111.

Lawson R, Humphreys GW. 1996. View specificity in object processing: Evidence from picture matching. J Exp Psychol [Hum Percept Perform] 22:395–416.

Lee BB, Elepfandt A, Virsu V. 1981. Phase of responses to sinusoidal gratings of simple cells in cat striate cortex. J Neurophysiol 45:818–828.

Lee D, Port NL, Kruse W, Georgopoulos AP. 1998. Variability and correlated noise in the discharge of neurons in motor and parietal areas of the primate cortex. J Neurosci 18:1161–1170.

Leek EC. 1998a. Effects of stimulus orientation on the identification of common polyoriented objects. Psychon Bull Rev 5:650–658.

Leek EC. 1998b. The analysis of orientation-dependent time costs in visual recognition. Perception 27:803–816.

Loftus GR, Nelson WW, Kallman HJ. 1983. Differential acquisition rates for different types of information from pictures. Q J Exp Psychol A 35:187–198.

Loftus GR, Ruthruff E. 1994. A theory of visual information acquisition and visual memory with special application to intensity-duration trade-offs. J Exp Psychol [Hum Percept Perform] 20:33–49.

Logothetis NK, Pauls J, Bulthoff HH, Poggio T. 1994. View-dependent object recognition by monkeys. Curr Biol 4:401–414.

Mauck B, Dehnhardt G. 1997. Mental rotation in a California sea lion (*Zalophus californianus*). J Exp Biol 200:1309–1316.

Nasanen R, Ojanpaa H, Kojo I. 2001. Effect of stimulus contrast on performance and eye movements in visual search. Vis Res 41:1817–1824.

Newell FN, Findlay JM. 1997. The effect of depth rotation on object identification. Perception 26:1231–1257.

Oram MW, Foldiak P, Perrett DI, Sengpiel F. 1998. The 'Ideal Homunculus': Decoding neural population signals. Trends Neurosci 21:259–265.

Oram MW, Gawne, TJ, Richmond BJ. 1997. Relationship of response latency and magnitude in the LGN of macaque monkey. Society for Neuroscience Abstracts 23:430–177.2.

Oram MW, Hatsopoulos NG, Richmond BJ, Donoghue JP. 2001. Excess synchrony in motor cortical neurons provides redundant direction information with that from coarse temporal measures. J Neurophysiol 86:1700–1716.

Oram MW, Perrett DI. 1992. Time course of neural responses discriminating different views of the face and head. J Neurophysiol 68:70–84.

Oram MW, Perrett DI. 1994. Modeling visual recognition from neurobiological constraints. Neural Networks 7:945–972.

Oram MW, Perrett DI. 1996. Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. J Neurophysiol 76:109–129.

Oram MW, Perrett DI, Hietanen JK. 1993. Directional tuning of motion-sensitive cells in the anterior superior temporal polysensory area of the macaque. Exp Brain Res 97:274–294.

Oram MW, Wiener MC, Lestienne R, Richmond BJ. 1999. Stochastic nature of precisely timed spike patterns in visual system neuronal responses. J Neurophysiol 81:3021–3033.

665

670

675

680

685

690

695

700

705

710

715

720   Oram MW, Xiao D, Dritschel B, Payne KR. 2002. The temporal resolution of neural codes: Does response latency have a unique role? Philos Trans R Soc Lond B 357:987–1001.

Oriet C, Jolicoeur P. 2003. Absence of perceptual processing during reconfiguration of task set. J Exp Psychol [Human Percept Perform 29:1036–1049.

Palmer J, McLean J. 1996. Visual search: Large set-size effect do not reject models based upon independent
725   channels. Invest Ophthalmol Vis Sci 37:62.

Pashler H, Johnston JC, Ruthruff E. 2001. Attention and performance. Annu Rev Psychol 52:629–651.

Perrett DI, Mistlin AJ, Chitty AJ, Smith PA, Potter DD, 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.

730   Perrett DI, Oram MW. 1993. Neurophysiology of shape processing. Image Vision Comput 11:317–333.

Perrett DI, Oram MW, Ashbridge E. 1998. Evidence accumulation in cell populations responsive to faces: An account of generalisation of recognition without mental transformations. Cognition 67:111–145.

Perrett DI, Oram MW, Harries MH, Bevan R, Hietanen JK, Benson PJ, Thomas S. 1991. Viewer-centred and object-centred coding of heads in the macaque temporal cortex. Exp Brain Res 86:159–173.

735   Perrett DI, Smith PA, Potter DD, Mistlin AJ, Head AS, Milner AD, Jeeves MA. 1984. Neurones responsive to faces in the temporal cortex: Studies of functional organization, sensitivity to identity and relation to perception. Hum Neurobiol 3:197–208.

Poggio T, Edelman S. 1990. A network that learns to recognize three-dimensional objects [see comments]. Nature 343:263–266.

740   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.

Ratcliff R, Rouder JN. 2000. A diffusion model account of masking in two-choice letter identification. J Exp Psychol [Hum Percept Perform] 26:127–140.

Regan D, Beverley KI. 1985. Postadaptation orientation discrimination. J Opt Soc Am A-Optics Image Sci Vis
745   2:147–155.

Reich DS, Mechler F, Victor JD. 2001a. Formal and attribute-specific information in primary visual cortex. J Neurophysiol 85:305–318.

Reich DS, Mechler F, Victor JD. 2001b. Temporal coding of contrast in primary visual cortex: When, what, and why. J Neurophysiol 85:1039–1050.

750   Robertson LC, Palmer SE, Gomez LM. 1987. Reference frames in mental rotation. J Exp Psychol [Learn Mem Cogn] 13:368–379.

Roitman JD, Shadlen MN. 2002. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. J Neurosci 22:9475–9489.

Rolls ET, Aggelopoulos NC, Franco L, Treves A. 2004. Information encoding in the inferior temporal visual cortex:
755   Contributions of the firing rates and the correlations between the firing of neurons. Biol Cybern 90:19–32.

Rolls ET, Baylis GC. 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.

Rolls ET, Franco L, Aggelopoulos NC, Reece S. 2003. An information theoretic approach to the contributions of the firing rates and the correlations between the firing of neurons. J Neurophysiol 89:2810–2822.

760   Ruthruff E, Johnston JC, Van Selst M, Whitsell S, Remington R. 2003. Vanishing dual-task interference after practice: Has the bottleneck been eliminated or is it merely latent? J Exp Psychol [Hum Percept Perform] 29:280–289.

Ruthruff E, Miller J. 1995. Can mental rotation begin before perception finishes? Mem Cognit 23:408–424.

Ruthruff E, Miller J, Lachmann T. 1995. Does mental rotation require central mechanisms? J Exp Psychol [Hum
765   Percept Perform] 21:552–570.

Salzman CD, Britten KH, Newsome WT. 1990. Cortical microstimulation influences perceptual judgements of motion direction. Nature 346:174–177.

Schall JD. 2002. The neural selection and control of saccades by the frontal eye field. Philos Trans R Soc Lond B 357:1073–1082.

770   Schall JD. 2003. Neural correlates of decision processes: Neural and mental chronometry. Curr Opin Neurobiol 13:182–186.

Schall JD, Thompson KG. 1999. Neural selection and control of visually guided eye movements. Annu Rev Neurosci 22:241–259.

Shadlen MN, Britten KH, Newsome WT, Movshon JA. 1996. A computational analysis of the relationship between
775   neuronal and behavioral responses to visual motion. J Neurosci 16:1486–1510.

Shadlen MN, Newsome WT. 1994. Noise, neural codes and cortical organization. Curr Opin Neurobiol 4:569–579.

Shadlen MN, Newsome WT. 1995. Is there a signal in the noise? Curr Opin Neurobiol 5:248–250.

24   *M. W. Oram*

Shadlen MN, Newsome WT. 1996. Motion perception: Seeing and deciding. Proc Natl Acad Sci USA 93:628–633.

Shadlen MN, Newsome WT. 1998. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. J Neurosci 18:3870–3896.   780

Shadlen MN, Newsome WT. 2001. Neural basis of a perceptual decision in the parietal cortex. area LIP) of the rhesus monkey. J Neurophysiol 86:1916–1936.

Shepard RN, Metzler J. 1971. Mental rotation of three-dimensional objects. Science 171:701–703.

Sheinburg DL, Logothetis NK. 1997. The role of temporal cortical areas in perceptual organization. Proc Natl Acad Sci USA 94:3408–3413   785

Sheinburg DL, Logothetis NK. 2001. Noticing familiar objects in real world scenes: The role of temporal cortical neurons in natural vision. J Neurosci 21:1430–1450

Singer W, Gray CM. 1995. Visual feature integration and the temporal correlation hypothesis. Annu Rev Neurosci 18:555–586.

Smith PL, Ratcliffe R. 2004. Psychology and neurobiology of simple decisions. Trends Neurosci 27:161–168   790

Snedecor GW, Cochran WG. 1980. Statistical Methods. Ames: Iowa State University Press.

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

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.   795

Tarr MJ, Bulthoff HH. 1995. Is human object recognition better described by geon structural descriptions or by multiple views? Comment on Biederman & Gerhardstein 1993. J Exp Psychol [Hum Percept Perform] 21:1494–1505.

Tarr MJ, Gauthier I. 1998. Do viewpoint-dependent mechanisms generalize across members of a class? Cognition 67:73–110.   800

Tarr MJ, Pinker S. 1989. Mental rotation and orientation-dependence in shape recognition. Cognit Psychol 21:233–282.

Tombu M, Jolicoeur P. 2002a. All-or-none bottleneck versus capacity sharing accounts of the psychological refractory period phenomenon. Psychol Res 66:274–286.

Tombu M, Jolicoeur P. 2002b. Does size rescaling require central attention? Can J Exp Psychol 56:10–17.   805

Tombu M, Jolicoeur P. 2003. A central capacity sharing model of dual-task performance. J Exp Psychol [Hum Percept Perform] 29:3–18.

Tovee MJ, Rolls ET, Azzopardi P. 1994. Translation invariance in the responses to faces of single neurons in the temporal visual cortical areas of the alert macaque. J Neurophysiol 72:1049–1060.

Usher M, McClelland JL. 2001. The time course of perceptual choice: The leaky, competing accumulator model. Psychol Rev 108:550–592.   810

Van Selst M, Jolicoeur P. 1994. Can mental rotation occur before the dual-task bottleneck? J Exp Psychol [Hum Percept Perform] 20:905–921.

Van Selst M, Ruthruff E, Johnston JC. 1999. Can practice eliminate the psychological refractory period effect? J Exp Psychol [Hum Percept Perform] 25:1268–1283.   815

Verfaillie K, Boutsen L. 1995. A corpus of 714 full-color images of depth-rotated objects. Percept Psychophys 57:925–961.

Verghese P. 2001. Visual search and attention: A signal detection theory approach. Neuron 31:523–535.

Vetter T, Poggio T, Bulthoff HH. 1994. The importance of symmetry and virtual views in three-dimensional object recognition. Curr Biol 4:18–23.   820

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.

Ward R, McClelland JL. 1989. Conjunctive search for one and two identical targets. J Exp Psychol [Hum Percept Perform] 15:664–672.

Wiener MC, Oram MW, Liu Z, Richmond BJ. 2001. Consistency of encoding in monkey visual cortex. J Neurosci 21:8210–8221.   825

Wiener MC, Oram M W, Richmond B J. 1998. Response latency is related to stimulus contrast but not to response strength. Soc Neuroscience Abstracts 24:1258–497.6.

Wiener MC, Richmond BJ. 1999. Using response models to estimate channel capacity for neuronal classification of stationary visual stimuli using temporal coding. J Neurophysiol 82:2861–2875.   830

Williams JM, Lit A. 1983. Luminance-dependent visual latency for the hess effect, the pulfrich effect, and simple reaction-time. Vis Res 23:171–179.

Queries

Q1. Au: Please provide received/accepted dates for manuscript.
Q2. Au: Please provide 3–5 article keywords.
Q3. Au: Please provide author's telephone and fax number.
Q4. Au: Pls. provide DOI.
Q5. Au: Please provide legend for Table I.