

**Original citation:**

Rolls, Edmund T. (2016) Cortical coding. *Language, Cognition and Neuroscience*. pp. 1-14.

**Permanent WRAP URL:**

<http://wrap.warwick.ac.uk/80516>

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

**Publisher's statement:**

"This is an Accepted Manuscript of an article published by Taylor & Francis Language, Cognition and Neuroscience on 05/07/2016 available online:

<http://www.tandfonline.com/10.1080/23273798.2016.1203443>

**A note on versions:**

The version presented here may differ from the published version or, version of record, if you wish to cite this item you are advised to consult the publisher's version. Please see the 'permanent WRAP URL' above for details on accessing the published version and note that access may require a subscription.

For more information, please contact the WRAP Team at: [wrap@warwick.ac.uk](mailto:wrap@warwick.ac.uk)

## **Cortical coding**

Edmund T Rolls

*Oxford Centre for Computational Neuroscience, Oxford, England and  
University of Warwick, Department of Computer Science, Coventry CV4 7AL, England*

Email: [Edmund.Rolls@oxcns.org](mailto:Edmund.Rolls@oxcns.org)

Url: <http://www.oxcns.org> where publications referred to are available

Running head: Cortical coding

\papers\visrev\Coding16\Coding16c.docx

For special issue entitled “Cognitive and neurophysiological evidence for and against localist ‘grandmother cell’ representations” to be published in *Language, Cognition and Neuroscience*, 2016.

## Summary

The encoding of information in the primate inferior temporal visual cortex, hippocampus, orbitofrontal cortex, and insula, is described. All these areas investigated have sparse distributed graded firing rate representations. The firing rate probability distribution is close to exponential. The information increases approximately linearly with the number of neurons. Consistent with this relative independence, there is little extra information that is available because of stimulus-dependent synchrony, and little redundancy. The code can be read very fast, in 20-50 ms, by dot-product, biologically plausible decoding. The advantages of this code include high capacity, generalization, graceful degradation, and rapid read-out of the information by biologically plausible dot-product decoding. None of these are properties of local or 'grandmother-cell' representations. Consistent evidence is becoming available for humans. Thus the evidence indicates that information is encoded in many cortical areas by sparse distributed graded firing rate representations, and this type of representation has major computational advantages for the brain including for language that are not met by local or 'grandmother-cell' representations.

In this paper, I describe the cortical neuronal encoding of information in primates including humans. The analysis is based on datasets of neurons in many different cortical areas of primates during the normal operation of the brain during behaviour, allowing a unique comparison of neuronal encoding in the inferior temporal visual cortex, the primary taste cortex in the insula, the orbitofrontal cortex, and the hippocampus and parahippocampal gyrus. The coding was analysed with rigorous information theoretic methods that we developed for application to single and multiple single neuron data and which are described by Rolls and Treves (2011), Rolls (2016), and in the original papers (Franco, Rolls, Aggelopoulos, & Treves, 2004; Rolls, Aggelopoulos, Franco, & Treves, 2004; Rolls, Franco, Aggelopoulos, & Perez, 2006; Rolls, Franco, Aggelopoulos, & Reece, 2003a; Rolls, Treves, & Tovee, 1997a; Rolls, Treves, Tovee, & Panzeri, 1997b).

I start with some definitions, then summarize some evidence that shows the type of encoding used in some cortical regions, and then show how the representation found is advantageous

### **Definitions of types of representation**

A *local representation* is one in which all the information that a particular stimulus or event occurred is provided by the activity of one of the neurons. In a famous example, a single neuron might be active only if one's grandmother was being seen, and this is sometimes called grandmother cell encoding. (The term was coined by Jerry Lettvin in about 1969 — see Charles Gross (2002).) An implication is that most neurons in the brain regions where objects or events are represented would fire only very rarely (Barlow, 1972; Barlow, 1995). A problem with this type of encoding is that a new neuron would be needed for every object or event that has to be represented. Another disadvantage is that this type of coding does not generalize easily to similar inputs, so that similarities between perceptions or memories would not be apparent. Another disadvantage is that the system is rather sensitive to brain damage: if a single neuron is lost, the representation may be lost. Another disadvantage of local encoding is that the storage capacity in a memory system in the brain (the number of stimuli that can be stored and recalled) may not be especially high (in the order of the number of synapses onto each neuron) (Rolls, 2016).

A *fully distributed representation* is one in which all the information that a particular stimulus or event occurred is provided by the activity of the full set of neurons. If the neurons are binary (e.g. either active or not), the most distributed encoding is when half the neurons are active (i.e. firing fast) for any one stimulus or event, and half are inactive. Different stimuli are represented by different subsets of the neurons being active.

A *sparse distributed representation* is a distributed representation in which a small proportion of the neurons is active at any one time. In a sparse representation with binary neurons (i.e. neurons with firing rates that are either high or low), less than half of the neurons are active for any one stimulus or event. For binary neurons, we can use as a measure of the sparseness the proportion of neurons in the active state. For neurons with real, continuously variable, values of firing rates, the sparseness  $a$  of the representation is defined below. A low value of the sparseness  $a$  indicates that few neurons are firing for any one stimulus, and its maximum value is 1.0.

### **Sparse distributed graded firing rate encoding of face and object identity in the inferior temporal visual cortex; of taste and related stimuli in the insula and orbitofrontal cortex; and of space in the hippocampus**

#### *Sparseness of the representation, and an approximately exponential firing rate probability distribution*

Barlow (1972) speculated that a particular object (or face) is represented in the brain by the firing of one or a few gnostic (or "grandmother") cells. We showed that this is not the case, and that although a face-selective cell may respond only to faces, its firing rate is graded to a set of faces with some faces producing large responses, and more and more producing lower and lower responses, with each neuron having a different profile of responses to each of the different faces with an approximately exponential firing rate probability distribution (Baddeley et al., 1997; Baylis, Rolls, & Leonard, 1985; Franco, Rolls, Aggelopoulos, & Jerez, 2007; Rolls & Tovee, 1995; Treves, Panzeri, Rolls, Booth, & Wakeman, 1999) (see Fig. 1; and also Figs. 3 and 4).

The sparseness of the representation provided by a single neuron  $a_s$  can be defined as

$$a_s = (\sum_{s=1,S} r_s/S)^2 / \sum_{s=1,S} (r_s^2/S)$$

where  $r_s$  is the mean firing rate of the neuron to stimulus  $s$  in the set of  $S$  stimuli. The average sparseness of the representation provided by this type of neuron was 0.65 in the study of inferior temporal cortex neurons by Rolls and Tovee (1995). This is the sparseness calculated from a single neuron across its responses to a large set of 68 stimuli, 23 of which were faces and 45 of which were non-face stimuli (see Fig. 2a). The values for the sparseness of each neuron are of interest, and the distribution shown in Fig. 2a shows that none of the neurons had very sparse grandmother cell types of representation (for which the sparseness would be in this study  $1/68=0.015$ ) in this or any other of our many studies (Rolls & Treves, 2011).

Of course it is not possible to exclude the possibility that some of the neurons in the inferior temporal cortex might be grandmother cells, for there are neurons that are rather unresponsive in the inferior temporal cortex as well as the primary taste cortex and orbitofrontal cortex (Kadohisa, Rolls, & Verhagen, 2005a; Rolls, Critchley, Verhagen, & Kadohisa, 2010) that might respond to just one stimulus if it could ever be found, but neurons with anything like this selectivity have never been found in our studies, and if present would be not part of the rather continuous distribution of sparseness values found in our studies (see e.g. Fig. 2). Moreover, the information encoding and transmission by the neurons that we have discovered and analysed are highly efficient with many useful properties as described below. **Further, cortical neurons tend to have low spontaneous firing rates, so can usually be detected, and then tested for responses to a wide range of visual stimuli including junk objects, to investigate whether they might show any indication of being responsive to one or a few of a large number of stimuli. (The reason for the low spontaneous firing rate is that neurons are held close to threshold, so that when an input is received, some of the neurons are ready to respond very quickly without having to charge up the membrane from a resting potential. This enables the cortex to respond quickly, but at the same time means that neurons occasionally emit spikes that are referred to as spontaneous activity, because of the stochastic dynamics as described elsewhere (Rolls, 2008, 2016; Rolls & Deco, 2010).) Further, in longitudinal single neuron recordings over many days, there is so far no indication that a highly sparse “grandmother cell”-like representation is found in the inferior temporal visual cortex (McMahon, Bondar, Afuwape, Ide, & Leopold, 2014).**

The same value for the sparseness is obtained when one calculates it as the responses of a large set of neurons to a single stimulus, which is termed the population sparseness  $a_p$ , and this indicates that the representation is weakly ergodic, that is that the response profiles of the different neurons are uncorrelated (Franco et al., 2007).

These values for  $a$  do not seem very sparse. But these values are calculated using the raw firing rates of the neurons, on the basis that these would be what a receiving neuron would receive as its input representation. However, neocortical neurons have a spontaneous firing rate of several spikes/s (with a lower value of 0.75 spikes/s for hippocampal pyramidal cells), and if this spontaneous value is subtracted from the firing rates to yield a ‘response sparseness’  $a_r$ , this value is considerably lower. For example, if the spontaneous firing rate was subtracted from the firing rate of the neuron to each stimulus, so that the changes of firing rate, i.e., the responses of the neurons, were used in the sparseness calculation, then the ‘response sparseness’ for the set of neurons in had a lower value, with a mean of  $a_r=0.33$  for the population of neurons as shown in Fig. 2b (Rolls & Tovee, 1995). Further, the true sparseness of the representation is probably less than this, for this is calculated only over the neurons that had responses to some of these stimuli. There were many more neurons that had no response to the stimuli. At least 10 times the number of inferior temporal cortex neurons had no responses to this set of 68 stimuli. So the true sparseness would be much lower than this value of 0.33. Further, it is important to remember the relative nature of sparseness measures, which (like the information measures to be discussed below) depend strongly on the stimulus set used (Rolls, 2016; Rolls & Treves, 2011).

Another way to analyse this is to examine the firing rate probability distribution of neurons. A grandmother cell would have one peak in this distribution at a firing rate of zero, and a second peak with a high firing rate with a very low probability. Franco, Rolls, Aggelopoulos and Jerez (2007) showed that while the firing rates of some single inferior temporal cortex neurons (tested in a visual fixation task to a set of 20 face and non-face stimuli) fit

an exponential distribution, and others with higher spontaneous firing rates have a gamma distribution with the mode shifted above zero rate if they have a higher than typical spontaneous firing rate, it turns out that there is a very close fit to an exponential distribution of firing rates if all spikes from all the neurons are considered together. This interesting result is shown in Fig. 3. Consistent with this (and the concept of weak ergodicity (Franco et al., 2007)), if the activity of a single inferior temporal cortex neuron is measured to a very large number of visual stimuli (while the monkey is watching a video), then the firing rate distribution is again frequently close to exponential, as illustrated in Fig. 4 (Baddeley et al., 1997).

It is helpful to note that the sparseness of the representation provided by a neuron with an exponential probability distribution of firing rates is 0.5, which is close to the values measured for the sparseness in different cortical areas. The possible utility of such a representation in terms of metabolic efficiency by having relatively few high firing rates is considered elsewhere (Rolls, 2016; Rolls & Treves, 2011; Treves et al., 1999).

*Comparisons of sparseness between areas: the inferior temporal visual cortex, hippocampus, insula, orbitofrontal cortex, and amygdala*

In the study of Franco, Rolls et al (2007) on inferior temporal visual cortex neurons, the selectivity of individual cells for the set of stimuli, or single cell sparseness  $a_s$ , had a mean value of 0.77. This is close to a previously measured estimate, 0.65, which was obtained with a larger stimulus set of 68 stimuli (Rolls & Tovee, 1995).

In contrast, the representation in some memory systems may be more sparse. For example, in the hippocampus in which spatial view cells are found in macaques, further analysis of data from macaque spatial view cells analyzed by Rolls et al (1998) shows that for the representation of 64 locations around the walls of the room, the mean single cell sparseness  $a_s$  was  $0.34 \pm 0.13$  (sd), and the mean population sparseness  $a_p$  was  $0.33 \pm 0.11$ . The more sparse representation is consistent with the view that the hippocampus is involved in storing memories, and that for this, more sparse representations than in perceptual areas are relevant (Kesner & Rolls, 2015; Rolls, 2016). Nevertheless, sparse distributed graded encoding is still used (Rolls et al., 1998).

Evidence is now available on sparseness, ergodicity and information encoding in three further brain areas, the macaque insular primary taste cortex, the orbitofrontal cortex, and the amygdala (Kadohisa et al., 2005a; Rolls, 2016; Rolls et al., 2010). In all these brain areas sets of neurons were tested with an identical set of 24 oral taste, temperature, and texture stimuli. (The stimuli were: Taste - 0.1 M NaCl (salt), 1 M glucose (sweet), 0.01 M HCl (sour), 0.001 M quinine HCl (bitter), 0.1 M monosodium glutamate (umami), and water; Temperature - 10°C, 37°C and 42°C; flavour - blackcurrant juice; viscosity - carboxymethyl-cellulose 10 cPoise, 100 cPoise, 1000 cPoise and 10000 cPoise; fatty / oily - single cream, vegetable oil, mineral oil, silicone oil (100 cPoise), coconut oil, and safflower oil; fatty acids - linoleic acid and lauric acid; capsaicin; and gritty texture.) Further analysis of data (Verhagen, Kadohisa, & Rolls, 2004) for the primary taste cortex showed that the mean value of  $a_s$  across 58 neurons was 0.745 and of  $a_p$  was 0.708. Further analysis of data for the orbitofrontal cortex (Kadohisa, Rolls, & Verhagen, 2004; Kadohisa et al., 2005a; Rolls, Verhagen, & Kadohisa, 2003b; Verhagen, Rolls, & Kadohisa, 2003) showed that the mean value of  $a_s$  across 30 neurons was 0.625 and of  $a_p$  was 0.611. Further analysis of data for the amygdala (Kadohisa, Rolls, & Verhagen, 2005b) showed that the mean value of  $a_s$  across 38 neurons was 0.811 and of  $a_p$  was 0.813. The values of  $a$  are relatively high, implying the importance of representing large amounts of information in these brain areas about this set of stimuli by using a very distributed code, and also perhaps about the stimulus set, some members of which may be rather similar to each other. Further, in all these cases, the mean value of  $a_s$  is close to that of  $a_p$ , and weak ergodicity is implied, providing further evidence on an important aspect of cortical encoding, that at least up to reasonable numbers of neurons, the coding by different neurons is relatively independent, that is, the response profiles of the neurons to the set of stimuli are relatively uncorrelated (Rolls, 2016).

Overall, we have seen that in many primate cortical areas and in the amygdala (which receives from the

cortex), a sparse distributed graded representation is found, with no evidence for grandmother cells (Rolls, 2016; Rolls & Treves, 2011).

### *Single cell information*

Complementary evidence comes from applying information theory to analyse how information is represented by neurons in the cortex. The information required to identify which of  $S$  equiprobable events occurred (or stimuli were shown) is  $\log_2 S$  bits. (Thus 1 bit is required to specify which of two stimuli was shown, 2 bits to specify which of 4 stimuli was shown, 3 bits to specify which of 8 stimuli was shown, etc.)

We are interested in measuring the information that we gain from the neuronal response  $r$  to a stimulus  $s$  in a set of  $S$  stimuli. The (Shannon) mutual information is the average information across all stimuli from the set  $S$  and all responses from the set  $R$ , as follows (Shannon, 1948, Rolls, 2016 #6387)

$$I(S, R) = \sum_{s,r} P(s, r) \log_2 \frac{P(s, r)}{P(s)P(r)}$$

where  $P(s, r)$  is the joint probability of the pair of results  $s$  and  $r$ .

We are also interested in the information specifically conveyed about each stimulus

$$I(s, R) = \sum_r P(r|s) \log_2 \frac{P(r|s)}{P(r)}$$

which is a direct quantification of the variability in the responses elicited by that stimulus, compared to the overall variability across all stimuli  $S$ . We term this the stimulus-specific information (Rolls, 2016). For a grandmother cell, a neuron would convey a large amount of information about only one of the stimuli. Thus this measure is useful in analysing encoding. A feature of the single cell information theoretic analyses that we have performed is that rigorous corrections have been made for the finite sampling effect, that is, that the number of trials of data is limited (Rolls, Critchley, & Treves, 1996; Rolls et al., 1997b; Tovee, Rolls, Treves, & Bellis, 1993).

Figure 5 shows the stimulus-specific information  $I(s, R)$  available in the neuronal response about each of 20 face stimuli calculated for the neuron (am242) whose firing rate response profile to the set of 68 stimuli is shown in Fig. 1. This is the information obtained on a single trial by the number of spikes in a 500 ms period. It is shown in Fig. 5 that 2.2, 2.0, and 1.5 bits of information were present about the three face stimuli to which the neuron had the highest firing rate responses. The neuron conveyed some but smaller amounts of information about the remaining face stimuli. The average information  $I(S, R)$  about this set ( $S$ ) of 20 faces for this neuron was 0.55 bits. The average firing rate of this neuron to these 20 face stimuli was 54 spikes/s. It is clear from Fig. 5 that little information was available from the responses of the neuron to a particular face stimulus if that response was close to the average response of the neuron across all stimuli. At the same time, it is clear from Fig. 5 that information was present depending on how far the firing rate to a particular stimulus was from the average response of the neuron to the stimuli. Of particular interest, it is evident that information is present from the neuronal response about which face was shown if that neuronal response was below the average response, as well as when the response was greater than the average response.

One intuitive way to understand the data shown in Fig. 5 is to appreciate that low probability firing rate responses, whether they are greater than or less than the mean response rate, convey much information about which stimulus was seen. This is of course close to the definition of information. Given that the firing rates of neurons are always positive, and follow an asymmetric distribution about their mean, it is clear that deviations above the mean have a different probability to occur than deviations by the same amount below the mean. One may attempt to capture the relative likelihood of different firing rates above and below the mean by computing a  $z$  score obtained

by dividing the difference between the mean response to each stimulus and the overall mean response by the standard deviation of the response to that stimulus. The greater the number of standard deviations (i.e. the greater the  $z$  score) from the mean response value, the greater the information might be expected to be. I therefore show in Fig. 6 the relation between the  $z$  score and  $I(s,R)$ . This results in a C-shaped curve in Figs. 5 and 6, with more information being provided by the neuron the further its response to a stimulus is in spikes per second or in  $z$  scores either above or below the mean response to all stimuli (which was 54 spikes/s for this neuron).

The very clear conclusion from this single cell information theoretic analysis is that neurons have a sparse distributed representation, in which considerable information is conveyed about a small proportion of the stimuli (in cortex those producing the highest firing rates), with some information about many other stimuli (depending on how far the response of a neuron is from its average response across all stimuli), and very little information about very many stimuli. (Of course if a neuron has no response to a stimulus, with this type of encoding one learns a small amount, for a response below the average response to all stimuli does convey a little information, as is made clear in Figs. 5 and 6.)

Similar results, providing clear evidence for a sparse distributed graded representation, were obtained from all neurons that we analysed in several datasets for neurons in the inferior temporal visual cortex responding to objects or faces (Franco et al., 2007; Rolls et al., 1997b). The same type of information representation was found for neurons responding to taste and odour in the orbitofrontal cortex (Rolls et al., 1996; Rolls et al., 2010), and to spatial view in the hippocampus (Rolls et al., 1998).

#### *The information available from multiple cells*

The analysis of the encoding of information by multiple single cells, for which we have developed methods (Franco et al., 2004; Rolls et al., 2004; Rolls et al., 2003a; Rolls et al., 1997a), also has implications for understanding how local or distributed the encoding is by cortical neurons. The important point for the present purposes is that if the encoding was local (or grandmother cell-like), the number of stimuli encoded by a population of neurons would be expected to rise approximately linearly with the number of neurons in the population. In contrast, with distributed encoding, provided that the neuronal responses are sufficiently independent, the information might be expected to rise linearly with the number of neurons, and the number of stimuli encodable by the population of neurons might be expected to rise exponentially as the number of neurons in the sample was increased (as information is a log measure) (Rolls, 2016).

The information available about which of 20 equiprobable faces had been shown that was available from the responses of different numbers of inferior temporal cortex neurons is shown in Fig. 7 (Rolls et al., 1997a). First, it is clear that some information is available from the responses of just one neuron - on average approximately 0.34 bits. Thus knowing the activity of just one neuron in the population does provide some evidence about which stimulus was present. This evidence that information is available in the responses of individual neurons in this way, without having to know the state of all the other neurons in the population, indicates that information is made explicit in the firing of individual neurons in a way that will allow neurally plausible decoding, involving computing a sum of input activities each weighted by synaptic strength, a dot product, to work (see below). Second, it is clear (Fig. 7) that the information rises approximately linearly, and the number of stimuli encoded thus rises approximately exponentially (Fig. 8), as the number of cells in the sample increases (Abbott, Rolls, & Tovee, 1996; Rolls & Treves, 1998; Rolls et al., 1997a), confirmed with simultaneous recordings from different neurons (Panzeri, Schultz, Treves, & Rolls, 1999a; Rolls et al., 2004; Rolls et al., 2006). Consistently, Gawne and Richmond (1993) showed that even adjacent pairs of neurons recorded simultaneously from the same electrode carried information that was approximately 80% independent. The same type of approximately linear increase in the information with the number of neurons in the sample is also found in the orbitofrontal cortex for olfactory and taste stimuli (Rolls et al., 2010), and in the hippocampus for spatial view neurons (Rolls et al., 1998). (These effects can only be seen with large stimulus sets, as only then are more than one or a very few neurons needed to encode the information, and even then the curves of the type illustrated in Fig. 7 tend to asymptote as the total



information required to represent the set of stimuli is approached as analysed by Rolls et al (1997a.)

The conclusion from the multiple cell information analyses then is that the signature of a sparse distributed representation is found, with information encoded relatively independently by different neurons.

Some have postulated that there might be information available if neurons became temporally synchronised to some but not other stimuli in a set (Engel, Konig, Kreiter, Schillen, & Singer, 1992; Singer, 1999). With rigorous information theoretic techniques (Rolls, 2008), we showed that for static faces and objects most of the information is available in the firing rates of the neurons (the number of spikes in a short time period), and that there is little additional information (<5% of the total) in the relative time of firing of simultaneously recorded neurons (Franco et al., 2004; Panzeri et al., 1999a; Rolls et al., 2004; Rolls et al., 2003a). This has been shown to apply to natural vision in natural scenes in which two test images had to be segmented from a complex background, the features of each object had to be bound together, and the monkey had to use top-down attention to search for one of two images in a complex scene (Aggelopoulos, Franco, & Rolls, 2005).

### **The Speed of Information Processing in the Temporal Cortical Visual Areas**

Given that there is a whole sequence of visual cortical processing stages including V1, V2, V4, and the posterior inferior temporal cortex to reach the anterior temporal cortical areas, and that the response latencies of neurons in V1 are about 40-50 ms, and in the anterior inferior temporal cortical areas approximately 80-100 ms, each stage may need to perform processing for only 15-30 ms before it has performed sufficient processing to start influencing the next stage (Rolls, 2012; Rolls, 2016). Consistent with this, response latencies between V1 and the inferior temporal cortex increase from stage to stage (Thorpe & Imbert, 1989).

In a first approach to this issue, we measured the information available in short temporal epochs of the responses of temporal cortical face-selective neurons about which face of a set of faces had been seen. We found that if a period of the firing rate of 50 ms was taken, then this contained 84.4% of the information available in a much longer period of 400 ms about which of four faces had been seen. If the epoch was as little as 20 ms, the information was 65% of that available from the firing rate in the 400 ms period (Tovee et al., 1993). We were able to extend this finding to the case when a much larger stimulus set, of 20 faces, was used. Again, we found that the information available in short (e.g. 50 ms) epochs was a considerable proportion (e.g. 65%) of that available in a 400 ms long firing rate analysis period (Tovee & Rolls, 1995). We extended these results by showing that although there is considerable information in the first spike of each neuron that arrives after a stimulus has been shown, there is more information if the number of spikes in a short window of for example 20 ms is used, and that the order of arrival of the spikes from different neurons is not an important factor, whereas the number of spikes in a short window is an important factor (Rolls et al., 2006).

The next approach has been to use a visual backward masking paradigm. In this paradigm there is a brief presentation of a test stimulus which is rapidly followed (within 1-100 ms) by the presentation of a second stimulus (the mask), which impairs or masks the perception of the test stimulus. It has been shown (Rolls & Tovee, 1994) that when there is no mask, inferior temporal cortex neurons respond to a 16 ms presentation of the test stimulus for 200-300 ms, far longer than the presentation time. It is suggested that this reflects the operation of a short term memory system implemented in cortical circuitry, which we propose is important in learning invariant representations (Rolls, 2008). If the pattern mask followed the onset of the test face stimulus by 20 ms (a stimulus onset asynchrony of 20 ms), face-selective neurons in the inferior temporal cortex of macaques responded for a period of 20-30 ms before their firing was interrupted by the mask (Rolls & Tovee, 1994; Rolls, Tovee, & Panzeri, 1999). We went on to show that under these conditions (a test-mask stimulus onset asynchrony of 20 ms), human observers looking at the same displays could just identify which of 6 faces was shown (Rolls, Tovee, Purcell, Stewart, & Azzopardi, 1994).

These results provide evidence that a cortical area can perform the computation necessary for the recognition of a visual stimulus in 20-30 ms (although it is true that for conscious perception, the firing needs to

occur for 40-50 ms, see Rolls, 2003). This provides a fundamental constraint which must be accounted for in any theory of cortical information representation, encoding, transmission, and computation. The results emphasise just how rapidly cortical circuitry can operate, a topic that is treated elsewhere (Panzeri, Rolls, Battaglia, & Lavis, 2001; Rolls, 2008, 2016; Rolls & Treves, 1998; Treves, 1993; Treves, Rolls, & Tovee, 1996).

These results from the primate cortex provide a clear answer to whether cortical neurons are grandmother cells: they are not, in the sense that each neuron has a graded set of responses to the different members of a set of stimuli, with the prototypical distribution similar to that of the neuron illustrated in Fig. 1. On the other hand, each neuron does respond very much more to some stimuli than to many others, and in this sense is tuned to some stimuli. This type of representation is not found in some models of invariant object representation in the visual cortical areas such as HMAX, but is approximated in VisNet (Robinson & Rolls, 2015). **(HMAX C layer neurons are extremely broadly tuned (Robinson & Rolls, 2015). When interpreted by a machine learning tool such as support vector machine learning, the representation that is required by the machine learning is just that specified by the experimenter, which might be for one object, or typically for any example of a single class of object such as hats or bears (Robinson & Rolls, 2015).)**

### **Advantages of the sparse distributed graded representation of objects for brain processing**

The advantages of the distributed encoding found include the following explained in more detail elsewhere (Rolls, 2007, 2008, 2014, 2016; Rolls & Treves, 1998, 2011), with a full analysis of how information theory has helped in the understanding of neural representations in the brain provided by Rolls (2016):

#### *1. Exponentially high coding capacity*

This property arises from a combination of the encoding being sufficiently close to independent by the different neurons (i.e. factorial), and sufficiently distributed, and is illustrated by the evidence shown in Figs. 7 and 8.

#### *2. Ease with which the code can be read by receiving neurons*

For brain plausibility, it is also a requirement that neurons should be able to read the code. This is why when we have estimated the information from populations of neurons, we have used in addition to a probability estimating measure (optimal, in the Bayesian sense), also a dot product measure, which is a way of specifying that all that is required of decoding neurons would be the property of adding up postsynaptic potentials produced through each synapse as a result of the activity of each incoming axon (Abbott et al., 1996; Rolls et al., 1997a). It was found that with such a neurally plausible algorithm (the Dot Product, DP, algorithm), which calculates which average response vector the neuronal response vector on a single test trial was closest to by performing a normalised dot product (equivalent to measuring the angle between the test and the average vector), the same generic results were obtained, with only a 40% reduction of information compared to the more efficient (Bayesian) algorithm. This is an indication that the brain could utilise the exponentially increasing capacity for encoding stimuli as the number of neurons in the population increases.

#### *3. Higher Resistance to Noise*

Because the information is decoded from a large population of neurons by inner product multiplication with the synaptic weight vector, there is less dependence on the random (almost Poisson) firing times for a given mean rate of single neurons, and thus there is resistance to noise inherent in the activity of single neurons (Rolls & Deco, 2010).

#### *4. Generalization*

Generalization to similar stimuli is again a property that arises in neuronal networks if distributed but not if local encoding is used. The generalization arises as a result of the fact that a neuron can be thought of as

computing the inner or dot product of the stimulus representation expressed as the firing rate on the set of input neurons with its synaptic weight vector (see further Rolls, 2008, 2016; Rolls & Deco, 2002; Rolls & Treves, 1998). With distributed representations, a neuron can generalize because the dot product of a related input will produce a similar (related) output (Rolls, 2008, 2016). With localist encoding, where a single neuron represents an object, there is no generalization of that neuron to related inputs. An input either fires the neuron if the object is present, or does not fire it if the object is not present. (Of course, those who advocate “grandmother cell” or “localist” encoding in the cortex might relax their definition of what they mean: and if so, it will be interesting see how close they reach to the sparse distributed encoding described in this paper, which is based on experimental evidence from thousands of single neuron recordings.)

### 5. Completion

Completion occurs in associative memory networks by a similar process. Completion is the property of recall of the whole of a pattern in response to any part of the pattern (Rolls, 2016).

### 6. Graceful degradation or fault tolerance

Again, because the information is decoded from a large population of neurons by inner product multiplication with the synaptic weight vector, there is less dependence on the firing of any one neuron or on any particular subset of neurons, so that if some neurons are damaged, the performance of the system only gradually degrades, and is in this sense fault tolerant (Rolls, 2016).

### 7. Speed of readout of the information

The information available in a distributed representation can be decoded by an analyzer more quickly than can the information from a local representation, given comparable firing rates. Within a fraction of an interspike interval, with a distributed representation, much information can be extracted (Panzeri, Treves, Schultz, & Rolls, 1999b; Rolls et al., 2006; Rolls et al., 1997a; Treves, 1993; Treves, Rolls, & Simmen, 1997; Treves et al., 1996). In effect, spikes from many different neurons can contribute to calculating the angle between a neuronal population and a synaptic weight vector within an interspike interval (Franco et al., 2004; Rolls, 2008, 2016). With local encoding, the speed of information readout depends on the exact model considered, but if the rate of firing needs to be taken into account, this will necessarily take time, because of the time needed for several spikes to accumulate in order to estimate the firing rate.

### 8. Distributed representations support attractor representations in the brain

Another advantage of distributed representations is that attractor states are very likely to be used to hold information on-line for short-term memory, attention, long-term memory, decision-making etc, and attractors can only be supported in the brain by neuronal systems with distributed not local representations because positive feedback from many neurons (not from a single neuron) is required to keep an attractor state active in the brain (Rolls, 2016; Rolls & Deco, 2010). (A single neuron providing input to itself would be insufficient to maintain its activity in a typical cortical pyramidal cell with 10,000 recurrent collateral synapses.)

My view is that even in systems such as those involved in language, the same sparse distributed graded encoding will be used, though the coding may be more sparse. My reasoning is that the identity of a person or object is encoded by a sparse distributed representation of neurons that reflect the main attributes. The more those attributes match, the more likely I am to decode the input as, for example, my grandmother. This type of representation can be built by competitive learning (Rolls, 2016). If on one occasion my grandmother is wearing glasses, that is then an attribute represented by some members of the neurons in the sparse distributed representation. If the glasses are absent (but are usually present), the input may look a little less like my grandmother, because the dot product decoding used by the receiving neurons is a little lower. This type of decoding, which is neuronally plausible, is described in more detail by Rolls (2016). A local or grandmother cell

representation does not cope at all with this scenario, for I might need a separate grandmother cell for every possible appearance of my grandmother, and there would be no useful generalization between the different grandmother cells all for my same grandmother but with different hats, glasses, etc.

Another advantage of sparse distributed representations for language is that attractor states are very likely to be used to hold on-line the parts of a sentence for syntactical operations (Rolls & Deco, 2015), and attractors can only be supported in the brain by neuronal systems with distributed not local representations as noted above (Rolls, 2016; Rolls & Deco, 2010).

What happens in this situation if I need to remember that at a particular occasion, say her 70<sup>th</sup> birthday, my grandmother was wearing a red hat? My view is that this case where particular attributes have to be remembered is then the role for the hippocampus, which is involved in episodic memory, and which can associate together the sparse representation received from the inferior temporal visual cortex that it is my grandmother, with other neurons that sparsely encode the red hat (and with other neurons that encode the time and place) (Kesner & Rolls, 2015; Rolls, 2016).

### **Evidence from humans**

Interesting findings are now becoming available about how neurons respond in medial temporal lobe regions (Fried, Rutishauser, Cerf, & Kreiman, 2014; Rolls, 2015). Mormann, Ison, Quiroga, Koch, Fried and Kreiman describe some neurons with responses that appear quite selective, with one neuron responding for example to Jennifer Aniston but much less to other individuals, and responding multimodally, for example not just to the sight of Jennifer Aniston, but also to the sound of her voice (Fried et al., 2014; Rey et al., 2015). At first, the tuning of single neurons might on the basis of a few striking examples in humans be thought to be more selective than those in the macaque temporal lobe, but on the basis of many such recordings, those who have recorded these neurons argue that the code is sparsely distributed (Quiroga, 2012, 2013; Quiroga, Kreiman, Koch, & Fried, 2008; Rey et al., 2015), and therefore somewhat similar to that of neurons in macaques (Rolls & Treves, 2011).

Some neurons recorded in the human medial temporal lobe areas such as the hippocampus are described as being ‘concept neurons’, for not only are they multimodal, but they can also respond to imagery of for example Jennifer Aniston, as in one famous case (Fried et al., 2014; Rey et al., 2015). How does this fit with concepts of hippocampal function? The hippocampus is thought to be involved in episodic or event memory, for example the memory of a particular person in a particular place (Kesner & Rolls, 2015; Rolls, 1989, 2008). Each memory must be as separate as possible from other memories, and the evidence is that single neurons in the macaque CA3 respond to combinations of for example a particular place being viewed, or a particular object, or a combination of these (Rolls, Xiang, & Franco, 2005). Indeed the theory is that the CA3 region with its recurrent collateral associatively modifiable synaptic connections enables any object or person to be associated with any place by this associativity, to form a unique episodic memory (Kesner & Rolls, 2015; Rolls, 1989, 2016). Human neurons in the hippocampus that respond to ‘concepts’, for example with quite selective tuning for a person, appear to be consistent with this theory. Of course, the nature of the sparsely distributed encoding is that no single neuron does need to be selective for just one person or object, for it is across a population of such neurons with sparsely distributed encoding that a particular individual is represented (Rolls & Treves, 2011) and becomes part of the autoassociation or attractor memory of a particular object or person in a particular place (Kesner & Rolls, 2015; Rolls, 1989, 2016).

A possible difference of single neurons in the human medial temporal lobe is that many seem to have rather low firing rate responses compared to those in macaques. However, the firing rates of neurons in different cortical areas are very different. In the inferior temporal visual cortex neurons with peak firing rates of 100 spikes/s to the most effective stimulus are common, whereas in the hippocampus neuronal responses have much lower firing rates, typically reaching a peak of 10-15 spikes/s to the most effective stimulus, from a spontaneous rate of

less than 1 spike/s (Rolls & Xiang, 2006). Thus when interpreting temporal lobe recordings in humans, it is important to take into account as much as possible the recording site, for what neurons respond to, and how much they respond, differs greatly between cortical areas. In this context, any information such as MNI coordinates of recorded single neurons in humans is important information to provide, and moreover will help the single neuron studies to be related to the activations found in human imaging studies, which of course reflect the average activity of hundreds of thousands of neurons, so provide little evidence about how the information is encoded by the neurons.

What are unique to humans are the findings on neuronal responses related to human language, described by Ojemann (Fried et al., 2014) (Chapter 14). Many of these recordings were made in lateral temporal cortex, and not from areas that are essential for language. One interesting finding has been of single neurons that change their activity when naming objects in one language, but not in another language. This suggests that the neuronal networks for different languages may be at least partly separate in terms of how they operate. Another interesting finding is that some temporal lobe neurons are involved in perception, and others in production, and indeed neurons with mirror-like properties are described as being rare in the superior temporal lobe cortical areas. In a more recent study, Halgren, Cash and colleagues described in the left anterior superior temporal gyrus of a right-handed man that 45% of units robustly fired to some spoken words with little or no response to pure tones, noise-vocoded speech, or environmental sounds (Chan et al., 2014). The tuning to words might be described as sparsely distributed. Many units were tuned to complex but specific sets of phonemes, which were influenced by local context but invariant to speaker, and suppressed during self-produced speech. The firing of several units to specific visual letters was correlated with their response to the corresponding auditory phonemes, providing direct neural evidence for phonological recoding during reading. A fundamental issue is how syntax is encoded in the cortex. A recent hypothesis is that place coding might be used, with for example a neuron responding to the word “cat” when it is the subject of a sentence, and not when it is the object (Rolls & Deco, 2015). If this hypothesis was found to be the case in single neuron recordings in future, this would greatly simplify concepts about how language and in particular syntax are implemented in the cortex (Rolls, 2016; Rolls & Deco, 2015).

## FIGURE LEGENDS

Figure 1. Firing rate distribution of a single neuron in the temporal visual cortex to a set of 23 face (F) and 45 non-face images of natural scenes. The firing rate response ( $\pm$  the standard error) to each of the 68 stimuli is shown, i.e. the spontaneous firing rate has been subtracted so that the 0 baseline is the spontaneous firing rate. P indicates a face profile stimulus, a B a body part stimulus such as a hand. Rather few stimuli produce high firing rates (e.g. above 60 spikes/s), and increasingly large numbers of stimuli produce lower and lower firing rates. The spontaneous firing rate of this neuron, the rate when no stimuli were being shown, was 20 spikes/s (Rolls and Tovee 1995). The histogram bars indicate the change of firing rate from the spontaneous value produced by each stimulus. The neuron responded best to three of the faces (profile views), had some response to some of the other faces, and had little or no response, and sometimes had a small decrease of firing rate below the spontaneous firing rate, to the non-face stimuli. (After Rolls and Tovee, 1995.) (fratedist.eps)

Figure 2. a. Distribution of sparseness value for the population of 14 inferior temporal cortex neurons for which firing rates were measured to a set of 23 face and 45 non-face images by Rolls and Tovee (1995). The mean sparseness value was  $0.65 \pm 0.16$  (mean  $\pm$  SD). b. The same data as (a), but now shown as response sparseness  $a_r$ , by subtracting the spontaneous firing rate. The mean of the response sparseness was  $0.33 \pm 0.22$  (mean  $\pm$  SD). (After Rolls and Tovee, 1995.) (sparseness95.eps)

Fig. 3. An exponential firing rate probability distribution obtained by pooling the firing rates of a population of 41 inferior temporal cortex neurons tested to a set of 20 face and non-face stimuli. The firing rate probability distribution for the 100-300 ms interval following stimulus onset was formed by adding the spike counts from all 41 neurons, and across all stimuli. The fit to the exponential distribution (dashed line) was high. (After Franco, Rolls, Aggelopoulos and Jerez 2007.) (Fig4-41\_2.eps)

Fig. 4. The probability of different firing rates measured in short (e.g. 100 ms or 500 ms) time windows of a temporal cortex neuron calculated over a 5 min period in which the macaque watched a video showing natural scenes, including faces. An exponential fit (+) to the data (diamonds) is shown. (After Baddeley, Abbott, Booth, Sengpiel, Freeman, Wakeman and Rolls 1997.) (ratenatsscenes.eps)

Fig. 5. The stimulus-specific information  $I(s,R)$  available in the response of the same single neuron as in Fig. 1 about each of the stimuli in the set of 20 face stimuli (abscissa), with the firing rate of the neuron to the corresponding stimulus plotted as a function of this on the ordinate. The horizontal line shows the mean firing rate across all stimuli. (After Rolls, Treves, Tovee and Panzeri 1997.) (scellinfo20f.eps)

Fig. 6. The relation for a single cell between the number of standard deviations the response to a stimulus was from the average response to all stimuli (see text,  $z$  score) plotted as a function of  $I(s,R)$ , the information available about the corresponding stimulus,  $s$ . (After Rolls, Treves, Tovee and Panzeri 1997) (zvsinfo.eps)

Figure 7a. The values for the average information available in the responses of different numbers of inferior temporal cortex neurons on each trial, about which of a set of 20 face stimuli has been shown. The decoding method was Dot Product (DP, x) or Bayesian Probability Estimation (PE, +), and the effects obtained with cross validation procedures utilising 50% of the trials as test trials are shown. The remainder of the trials in the cross-validation procedure were used as training trials. Probability estimation refers to estimating the probability from the neuronal responses for a single trial that each of the 20 faces was shown, thus utilizing the neuronal evidence available about how likely it was that each stimulus was shown, see Rolls et al 1997. The full line indicates the amount of information expected from populations of increasing size, when assuming random correlations within the constraint given by the ceiling (the information in the stimulus set,  $I=4.32$  bits). b. The percent correct for the corresponding data to those shown in Fig. 7a. (After Rolls, Treves and Tovee, 1997.) (multicellinfo20f\_2.eps)

Fig. 8. The number of stimuli (in this case from a set of 20 faces) that are encoded in the responses of different

numbers of neurons in the temporal lobe visual cortex, based on results shown in Fig. 7. The solid circles are the result of the raw calculation, and the open circles correspond to the cross-validated case. (After Abbott, Rolls and Tovee 1996.) (multicellexprot.eps)

## ACKNOWLEDGEMENTS

The author has worked on some of the investigations described here with L.Abbott, N.Aggelopoulos, P.Azzopardi, G.C.Baylis, R.Baddeley, M.Booth, A.S.Browning, H.Critchley, G.Deco, L.Franco, M.Kadohisa, C.M.Leonard, S.Panzeri, D.I.Perrett, S.Reece, L.Robinson, M.J.Tovee, A.Treves, S.J.Thorpe, and E.A.Wakeman, and their collaboration is sincerely acknowledged. Different parts of the research described were supported by the Medical Research Council, PG8513790; by The Wellcome Trust; by a Human Frontier Science Program grant; by an EC Human Capital and Mobility grant; by the MRC Oxford Interdisciplinary Research Centre in Cognitive Neuroscience; and by the Oxford McDonnell Centre in Cognitive Neuroscience. A description of the computational implications of some of the findings described here is provided in Rolls (2016).

Some of the publications referred to here are available at <http://www.oxcns.org> .





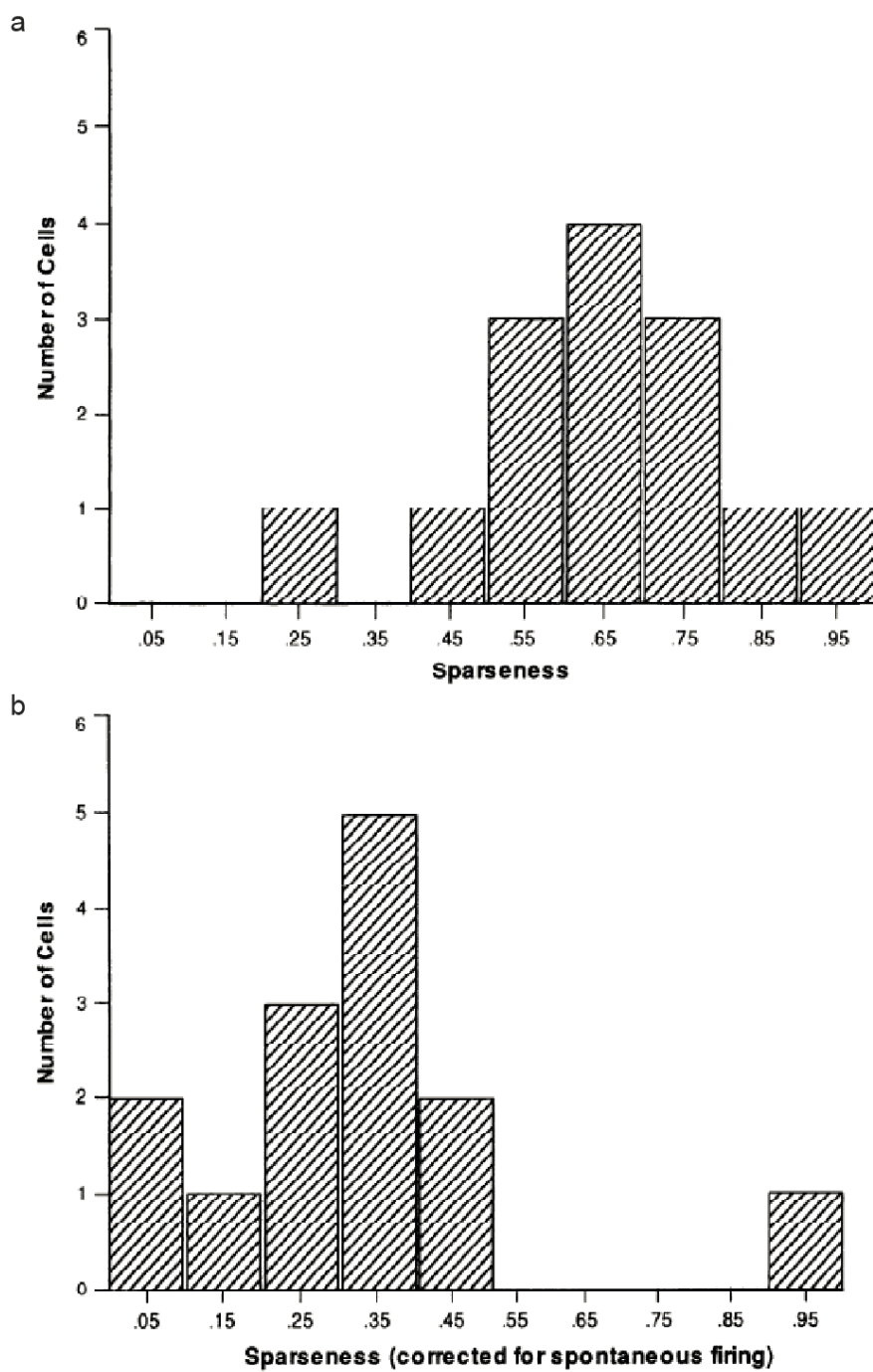


Fig. 2.

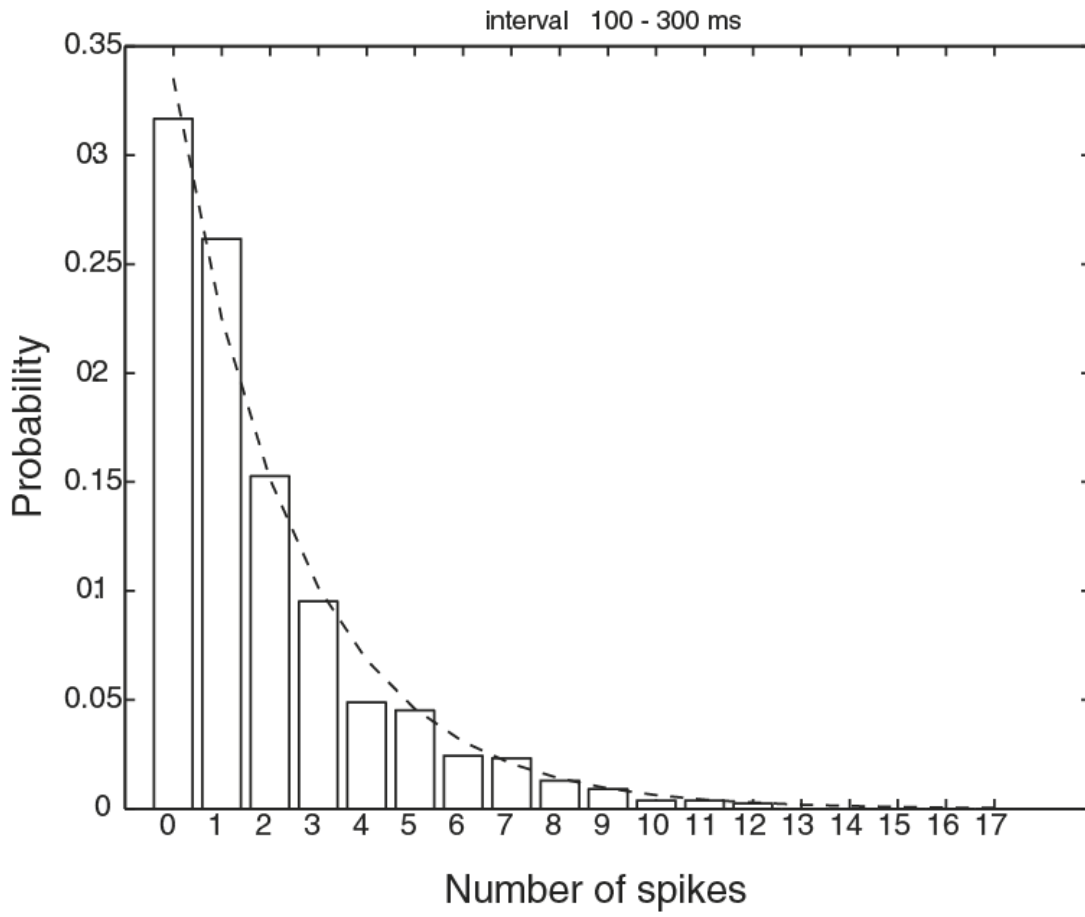


Fig. 3.

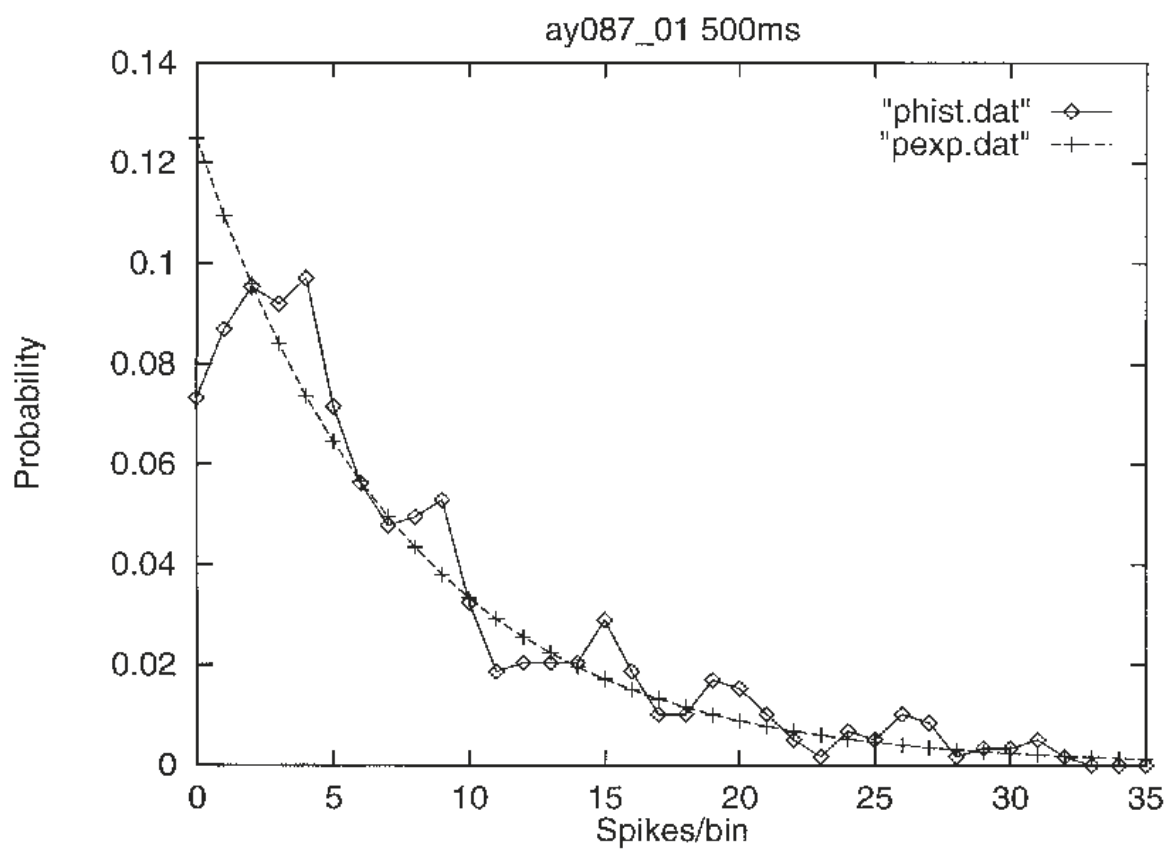


Fig. 4.

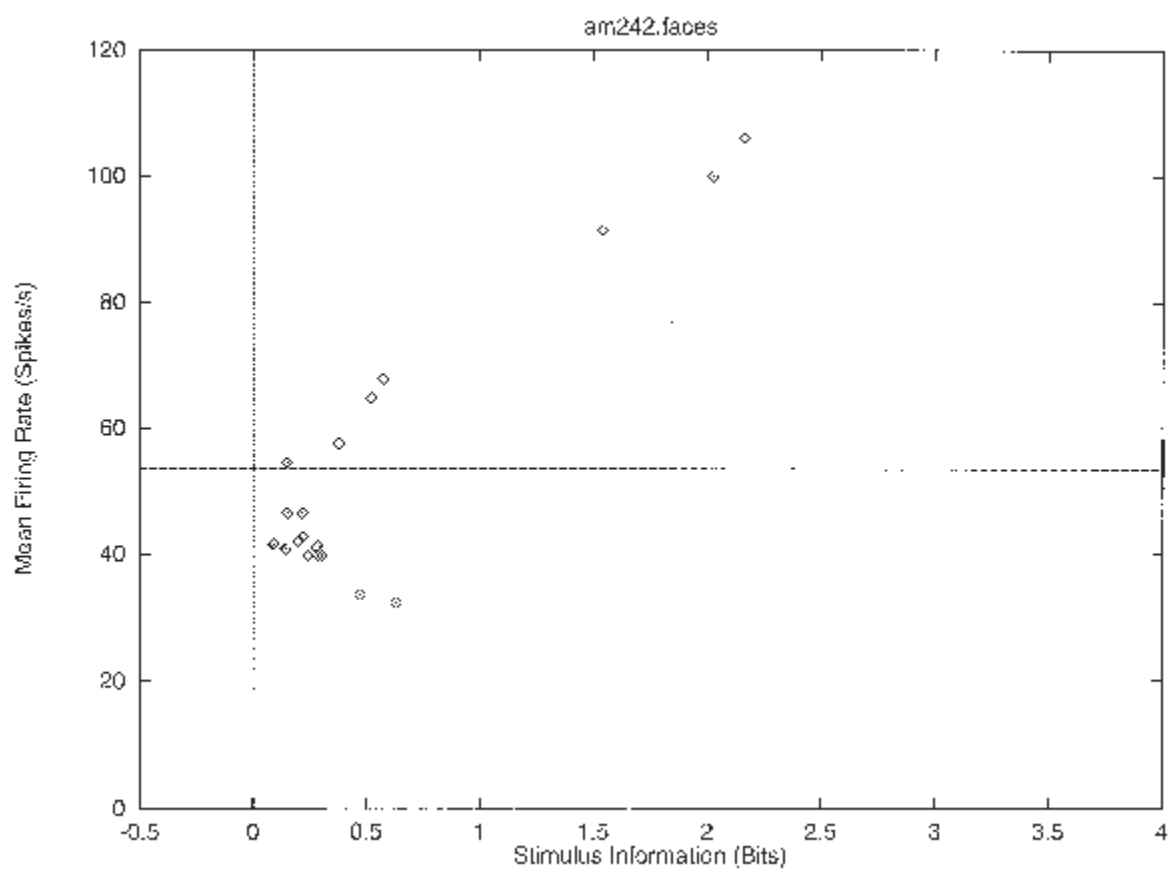


Fig. 5

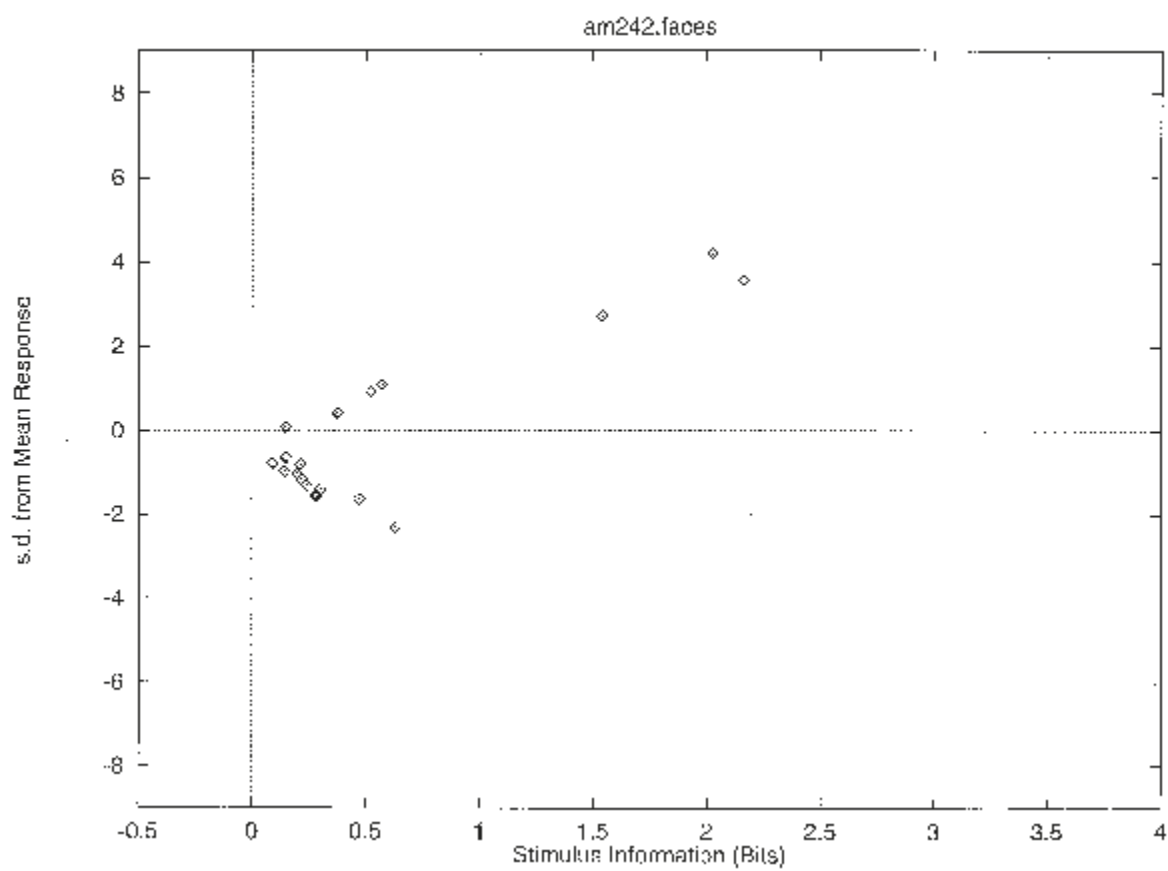


Fig. 6

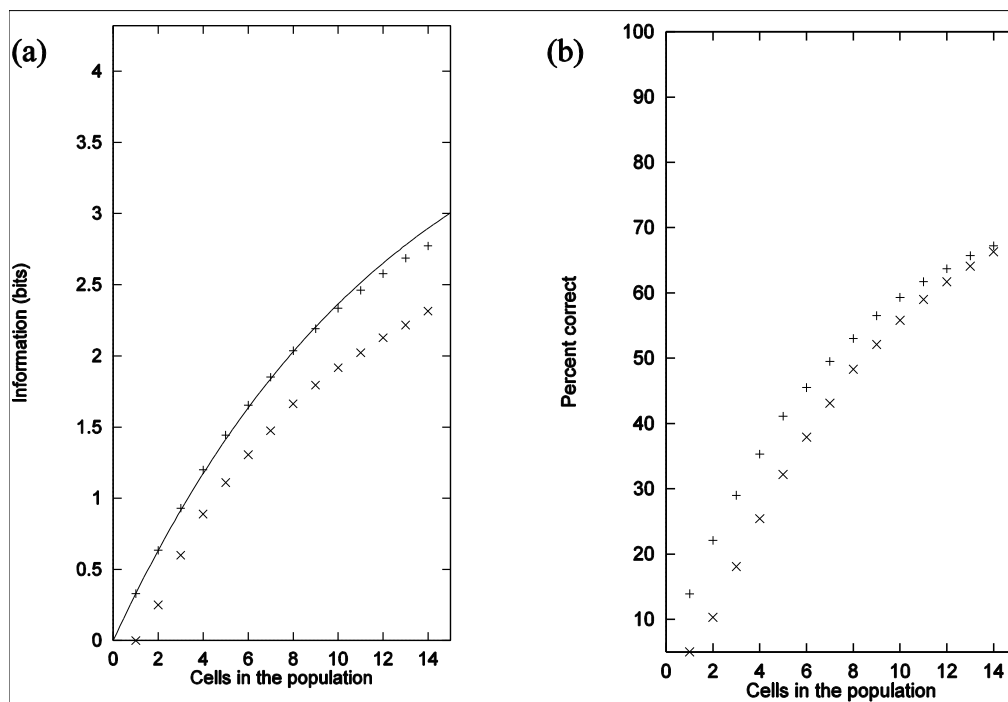


Fig. 7

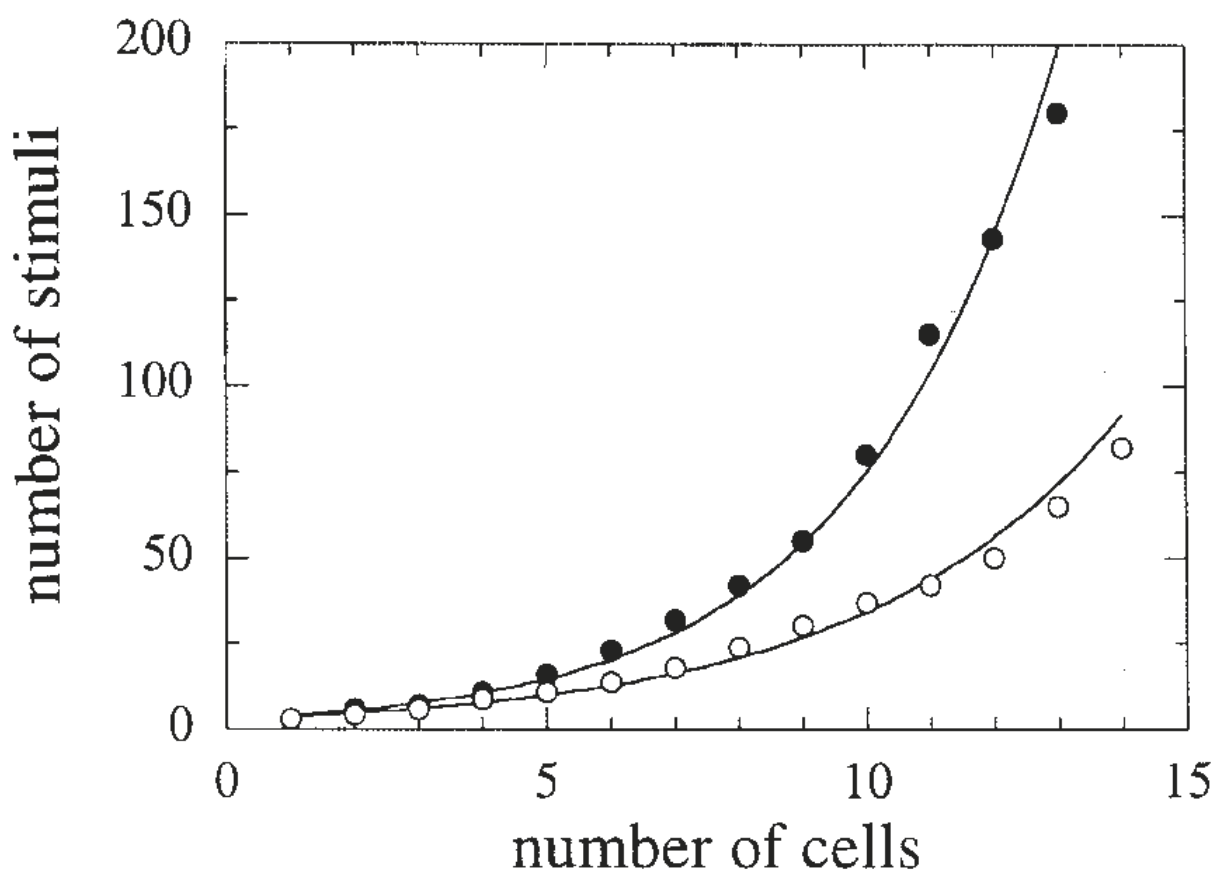


Fig. 8.



## References

- Abbott, L. F., Rolls, E. T., & Tovee, M. J. (1996). Representational capacity of face coding in monkeys. *Cerebral Cortex*, 6(3), 498-505. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8670675>
- Aggelopoulos, N. C., Franco, L., & Rolls, E. T. (2005). Object perception in natural scenes: encoding by inferior temporal cortex simultaneously recorded neurons. *Journal of Neurophysiology*, 93(3), 1342-1357. doi:10.1152/jn.00553.2004
- Baddeley, R. J., Abbott, L. F., Booth, M. J. A., Sengpiel, F., Freeman, T., Wakeman, E. A., & Rolls, E. T. (1997). Responses of neurons in primary and inferior temporal visual cortices to natural scenes. *Proceedings of the Royal Society of London B*, 264(1389), 1775-1783. doi:10.1098/rspb.1997.0246
- Barlow, H. B. (1972). Single units and sensation: a neuron doctrine for perceptual psychology? *Perception*, 1(4), 371-394. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/4377168>
- Barlow, H. B. (1995). The neuron doctrine in perception. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences* (pp. 415-435). Cambridge, Mass.: MIT Press.
- Baylis, G. C., Rolls, E. T., & Leonard, C. M. (1985). Selectivity between faces in the responses of a population of neurons in the cortex in the superior temporal sulcus of the monkey. *Brain Research*, 342, 91-102.
- Chan, A. M., Dykstra, A. R., Jayaram, V., Leonard, M. K., Travis, K. E., Gygi, B., . . . Cash, S. S. (2014). Speech-specific tuning of neurons in human superior temporal gyrus. *Cerebral Cortex*, 24(10), 2679-2693. doi:10.1093/cercor/bht127
- Engel, A. K., Konig, P., Kreiter, A. K., Schillen, T. B., & Singer, W. (1992). Temporal coding in the visual system: new vistas on integration in the nervous system. *Trends in Neurosciences*, 15(6), 218-226. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/1378666>
- Franco, L., Rolls, E. T., Aggelopoulos, N. C., & Jerez, J. M. (2007). Neuronal selectivity, population sparseness, and ergodicity in the inferior temporal visual cortex. *Biological Cybernetics*, 96(6), 547-560. doi:10.1007/s00422-007-0149-1
- Franco, L., Rolls, E. T., Aggelopoulos, N. C., & Treves, A. (2004). The use of decoding to analyze the contribution to the information of the correlations between the firing of simultaneously recorded neurons. *Experimental Brain Research*, 155, 370-384.
- Fried, I., Rutishauser, U., Cerf, M., & Kreiman, G. (2014). *Single Neuron Studies of the Human Brain: Probing Cognition*: MIT Press.
- Gawne, T. J., & Richmond, B. J. (1993). How independent are the messages carried by adjacent inferior temporal cortical neurons? *Journal of Neuroscience*, 13, 2758-2771.
- Gross, C. G. (2002). Genealogy of the "grandmother cell". *Neuroscientist*, 8(5), 512-518. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12374433>
- Kadohisa, M., Rolls, E. T., & Verhagen, J. V. (2004). Orbitofrontal cortex neuronal representation of temperature and capsaicin in the mouth. *Neuroscience*, 127, 207-221.
- Kadohisa, M., Rolls, E. T., & Verhagen, J. V. (2005a). Neuronal representations of stimuli in the mouth: the primate insular taste cortex, orbitofrontal cortex, and amygdala. *Chemical Senses*, 30(5), 401-419. doi:10.1093/chemse/bji036
- Kadohisa, M., Rolls, E. T., & Verhagen, J. V. (2005b). The primate amygdala: neuronal representations of the viscosity, fat texture, temperature, grittiness and taste of foods. *Neuroscience*, 132(1), 33-48. doi:10.1016/j.neuroscience.2004.12.005
- Kesner, R. P., & Rolls, E. T. (2015). A computational theory of hippocampal function, and tests of the theory: new developments. *Neuroscience and Biobehavioral Reviews*, 48, 92-147. doi:10.1016/j.neubiorev.2014.11.009
- McMahon, D. B., Bondar, I. V., Afuwape, O. A., Ide, D. C., & Leopold, D. A. (2014). One month in the

- life of a neuron: longitudinal single-unit electrophysiology in the monkey visual system. *Journal of Neurophysiology*, 112(7), 1748-1762. doi:10.1152/jn.00052.2014
- Panzeri, S., Rolls, E. T., Battaglia, F., & Lavis, R. (2001). Speed of feedforward and recurrent processing in multilayer networks of integrate-and-fire neurons. *Network: Computation in Neural Systems*, 12(4), 423-440. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11762898>
- Panzeri, S., Schultz, S. R., Treves, A., & Rolls, E. T. (1999a). Correlations and the encoding of information in the nervous system. *Proceedings of the Royal Society of London B*, 266(1423), 1001-1012. doi:10.1098/rspb.1999.0736
- Panzeri, S., Treves, A., Schultz, S., & Rolls, E. T. (1999b). On decoding the responses of a population of neurons from short time windows. *Neural Computation*, 11, 1553-1577.
- Quiroga, R. Q. (2012). Concept cells: the building blocks of declarative memory functions. *Nature Reviews: Neuroscience*, 13(8), 587-597. doi:10.1038/nrn3251
- Quiroga, R. Q. (2013). Gnostic cells in the 21st century. *Acta Neurobiologiae Experimentalis*, 73(4), 463-471. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/24457638>
- Quiroga, R. Q., Kreiman, G., Koch, C., & Fried, I. (2008). Sparse but not 'grandmother-cell' coding in the medial temporal lobe. *Trends in Cognitive Sciences*, 12(3), 87-91. doi:10.1016/j.tics.2007.12.003
- Rey, H. G., Ison, M. J., Pedreira, C., Valentin, A., Alarcon, G., Selway, R., . . . Quiroga, R. (2015). Single-cell recordings in the human medial temporal lobe. *Journal of Anatomy*, 227(4), 394-408. doi:10.1111/joa.12228
- Robinson, L., & Rolls, E. T. (2015). Invariant visual object recognition: biologically plausible approaches. *Biological Cybernetics*, 109(4-5), 505-535. doi:10.1007/s00422-015-0658-2
- Rolls, E. T. (1989). Functions of neuronal networks in the hippocampus and neocortex in memory. In J. H. Byrne & W. O. Berry (Eds.), *Neural Models of Plasticity: Experimental and Theoretical Approaches* (pp. 240-265). San Diego: Academic Press.
- Rolls, E. T. (2003). Consciousness absent and present: a neurophysiological exploration. *Progress in Brain Research*, 144, 95-106.
- Rolls, E. T. (2007). The representation of information about faces in the temporal and frontal lobes. *Neuropsychologia*, 45, 125-143.
- Rolls, E. T. (2008). *Memory, Attention, and Decision-Making: A Unifying Computational Neuroscience Approach*. Oxford: Oxford University Press.
- Rolls, E. T. (2012). Invariant visual object and face recognition: neural and computational bases, and a model, VisNet. *Frontiers in Computational Neuroscience*, 6, 35, 1-70. doi:10.3389/fncom.2012.00035
- Rolls, E. T. (2014). *Emotion and Decision-Making Explained*. Oxford: Oxford University Press.
- Rolls, E. T. (2015). The neuronal representation of information in the human brain. Review. *Brain*, 138, 3459-3462. doi:10.1093/brain/awv242
- Rolls, E. T. (2016). *Cerebral Cortex: Principles of Operation*. Oxford: Oxford University Press.
- Rolls, E. T., Aggelopoulos, N. C., Franco, L., & Treves, A. (2004). Information encoding in the inferior temporal cortex: contributions of the firing rates and correlations between the firing of neurons. *Biological Cybernetics*, 90, 19-32. doi:10.1007/s00422-003-0451-5
- Rolls, E. T., Critchley, H. D., & Treves, A. (1996). The representation of olfactory information in the primate orbitofrontal cortex. *Journal of Neurophysiology*, 75(5), 1982-1996. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8734597>
- Rolls, E. T., Critchley, H. D., Verhagen, J. V., & Kadohisa, M. (2010). The representation of information about taste and odor in the orbitofrontal cortex. *Chemosensory Perception*, 3, 16-33. doi:10.1007/s12078-009-9054-4
- Rolls, E. T., & Deco, G. (2002). *Computational Neuroscience of Vision*. Oxford: Oxford University Press.
- Rolls, E. T., & Deco, G. (2010). *The Noisy Brain: Stochastic Dynamics as a Principle of Brain Function*.

- Oxford: Oxford University Press.
- Rolls, E. T., & Deco, G. (2015). Networks for memory, perception, and decision-making, and beyond to how the syntax for language might be implemented in the brain. *Brain Research*, 1621, 316-334. doi:10.1016/j.brainres.2014.09.021
- Rolls, E. T., Franco, L., Aggelopoulos, N. C., & Perez, J. M. (2006). Information in the first spike, the order of spikes, and the number of spikes provided by neurons in the inferior temporal visual cortex. *Vision Research*, 46(25), 4193-4205. doi:10.1016/j.visres.2006.07.026
- Rolls, E. T., Franco, L., Aggelopoulos, N. C., & Reece, S. (2003a). An information theoretic approach to the contributions of the firing rates and correlations between the firing of neurons. *Journal of Neurophysiology*, 89, 2810-2822. doi:10.1152/jn.01070.2002
- Rolls, E. T., & Tovee, M. J. (1994). Processing speed in the cerebral cortex and the neurophysiology of visual masking. *Proceedings of the Royal Society of London B*, 257(1348), 9-15. doi:10.1098/rspb.1994.0087
- Rolls, E. T., & Tovee, M. J. (1995). Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *Journal of Neurophysiology*, 73(2), 713-726. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7760130>
- Rolls, E. T., Tovee, M. J., & Panzeri, S. (1999). The neurophysiology of backward visual masking: information analysis. *Journal of Cognitive Neuroscience*, 11(3), 335-346. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10402257>
- Rolls, E. T., Tovee, M. J., Purcell, D. G., Stewart, A. L., & Azzopardi, P. (1994). The responses of neurons in the temporal cortex of primates, and face identification and detection. *Experimental Brain Research*, 101(3), 473-484. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7851514>
- Rolls, E. T., & Treves, A. (1998). *Neural Networks and Brain Function*. Oxford: Oxford University Press.
- Rolls, E. T., & Treves, A. (2011). The neuronal encoding of information in the brain. *Progress in Neurobiology*, 95(3), 448-490. doi:10.1016/j.pneurobio.2011.08.002
- Rolls, E. T., Treves, A., Robertson, R. G., Georges-François, P., & Panzeri, S. (1998). Information about spatial view in an ensemble of primate hippocampal cells. *Journal of Neurophysiology*, 79, 1797-1813.
- Rolls, E. T., Treves, A., & Tovee, M. J. (1997a). The representational capacity of the distributed encoding of information provided by populations of neurons in the primate temporal visual cortex. *Experimental Brain Research*, 114(1), 177-185. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9125461>
- Rolls, E. T., Treves, A., Tovee, M. J., & Panzeri, S. (1997b). Information in the neuronal representation of individual stimuli in the primate temporal visual cortex. *Journal of Computational Neuroscience*, 4(4), 309-333. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9427118>
- Rolls, E. T., Verhagen, J. V., & Kadohisa, M. (2003b). Representations of the texture of food in the primate orbitofrontal cortex: neurons responding to viscosity, grittiness and capsaicin. *Journal of Neurophysiology*, 90(6), 3711-3724. doi:10.1152/jn.00515.2003
- Rolls, E. T., & Xiang, J.-Z. (2006). Spatial view cells in the primate hippocampus, and memory recall. *Reviews in the Neurosciences*, 17(1-2), 175-200. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16703951>
- Rolls, E. T., Xiang, J.-Z., & Franco, L. (2005). Object, space and object-space representations in the primate hippocampus. *Journal of Neurophysiology*, 94(1), 833-844. doi:10.1152/jn.01063.2004
- Shannon, C. E. (1948). A mathematical theory of communication. *AT&T Bell Laboratories Technical Journal*, 27, 379-423.
- Singer, W. (1999). Neuronal synchrony: A versatile code for the definition of relations? *Neuron*, 24(1), 49-65. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10677026>
- Thorpe, S. J., & Imbert, M. (1989). Biological constraints on connectionist models. In R. Pfeifer, Z. Schreter,

- & F. Fogelman-Soulie (Eds.), *Connectionism in Perspective* (pp. 63-92). Amsterdam: Elsevier.
- Tovee, M. J., & Rolls, E. T. (1995). Information encoding in short firing rate epochs by single neurons in the primate temporal visual cortex. *Visual Cognition*, 2(1), 35-58.
- 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. *Journal of Neurophysiology*, 70(2), 640-654. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8410164>
- Treves, A. (1993). Mean-field analysis of neuronal spike dynamics. *Network*, 4, 259-284.
- Treves, A., Panzeri, S., Rolls, E. T., Booth, M., & Wakenan, E. A. (1999). Firing rate distributions and efficiency of information transmission of inferior temporal cortex neurons to natural visual stimuli. *Neural Computation*, 11(3), 601-631. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10085423>
- Treves, A., Rolls, E. T., & Simmen, M. (1997). Time for retrieval in recurrent associative memories. *Physica D*, 107, 392-400.
- Treves, A., Rolls, E. T., & Tovee, M. J. (1996). On the time required for recurrent processing in the brain. In V. Torre & F. Conti (Eds.), *Neurobiology: Ionic Channels, Neurons, and the Brain* (pp. 325-353). New York: Plenum.
- Verhagen, J. V., Kadohisa, M., & Rolls, E. T. (2004). The primate insular/opercular taste cortex: neuronal representations of the viscosity, fat texture, grittiness, temperature and taste of foods. *Journal of Neurophysiology*, 92(3), 1685-1699. doi:10.1152/jn.00321.2004
- Verhagen, J. V., Rolls, E. T., & Kadohisa, M. (2003). Neurons in the primate orbitofrontal cortex respond to fat texture independently of viscosity. *Journal of Neurophysiology*, 90(3), 1514-1525. doi:10.1152/jn.00320.2003