

Encoding of Categories by Noncategory-Specific Neurons in the Inferior Temporal Cortex

Elizabeth Thomas¹, Marc M. Van Hulle², and Rufin Vogels²

Abstract

■ In order to understand how the brain codes natural categories, e.g., trees and fish, recordings were made in the anterior part of the macaque inferior temporal (IT) cortex while the animal was performing a tree/nontree categorization task. Most single cells responded to exemplars of more than one category while other neurons responded only to a restricted set of exemplars of a given category. Since it is still not known which type of cells contribute and what is the nature of the code used for categorization in IT, we have performed an analysis on single-cell data. A Kohonen self-organizing map (SOM), which uses an unsupervised (competitive) learning algorithm, was used to study the single cell responses to tree and nontree images. Results from the Kohonen SOM indicated that the collected neuronal data consisting of spike counts was sufficient to account for a good level of categorization success (approximately 83%) when

categorizing a group of 200 trees and nontrees. Contrary to intuition, the results of the investigation suggest that the population of category-specific neurons (neurons that respond only to trees or only to nontrees) was unimportant to the categorization. Instead, a large majority of the neurons that were most important to the categorization was found to belong to a class of more broadly tuned cells, namely, cells that responded to both categories but that favored one category over the other by seven or more images. A simple algebraic operation (without the Kohonen SOM) between the above-mentioned noncategory-specific neurons confirmed the contribution of these neurons to categorization. Thus, the modeling results suggest (1) that broadly tuned neurons are critical for categorization, and (2) that only one additional layer of processing is required to extract the categories from a population of IT neurons. ■

INTRODUCTION

The inferior temporal (IT) cortex has been identified as an area that may be involved in object categorization. Neurons in that area produce responses that are shape- and color-selective (Tanaka, 1993; Gross, Rocha-Miranda, & Bender, 1972). Furthermore, the shape selectivity is largely size invariant (Sary, Vogels, & Orban, 1993; Schwartz, Desimone, Albright, & Gross, 1983; Sato, Kawamura, & Iwai, 1980) and position invariant (Schwartz et al., 1983). Sary et al. (1993) found that the overall shape preference of neurons in the IT cortex remained the same whether the shapes were defined by luminance, relative motion, or texture. In addition, the shape selectivity of these neurons remained the same with partial occlusion (Vogels & Orban, 1996; Kovacs, Vogels, & Orban, 1995). In studies done on monkeys this invariance in shape selectivity and the detrimental affect of lesions of the area IT on object discrimination (Ungerleider & Mishkin, 1982) suggest that this area plays an important role in object recognition and categorization. Human neurophysiological studies also support the conclusion that the IT cortex is involved in

categorization. Indeed, category-specific impairments have been linked to temporal lobe injuries (Damasio, Grabowski, Tranel, Hichwa, & Damasio, 1996).

We studied the nature of the neuronal coding of categories in IT by recording the responses of IT neurons in monkeys while they were performing a categorization of natural images. We utilized a categorization task in which a monkey categorized colored images of approximately 200 trees and 200 nontrees (for details see Vogels, 1999a). The images of the tree category were similar to the nontree images in average size and mean luminance. There was a large variation in the form, texture, and color of the tree exemplars. Some of the nontree stimuli resembled particular trees in their shape, texture, or color. In this article, we have included black and white pictures of some of the original colored images used as stimuli for the monkey (Figure 1). The visual categorization task was designed to minimize the rote learning of individual exemplars, i.e., to minimize the possibility that the animal learned to associate individual stimuli with a particular response. This was done by limiting the number of times a particular stimulus was presented to the monkey. The monkey was found to be capable of performing this task with greater than 95% success. Features like color, texture,

¹ Université de Liège, ² Katholieke Universiteit Leuven

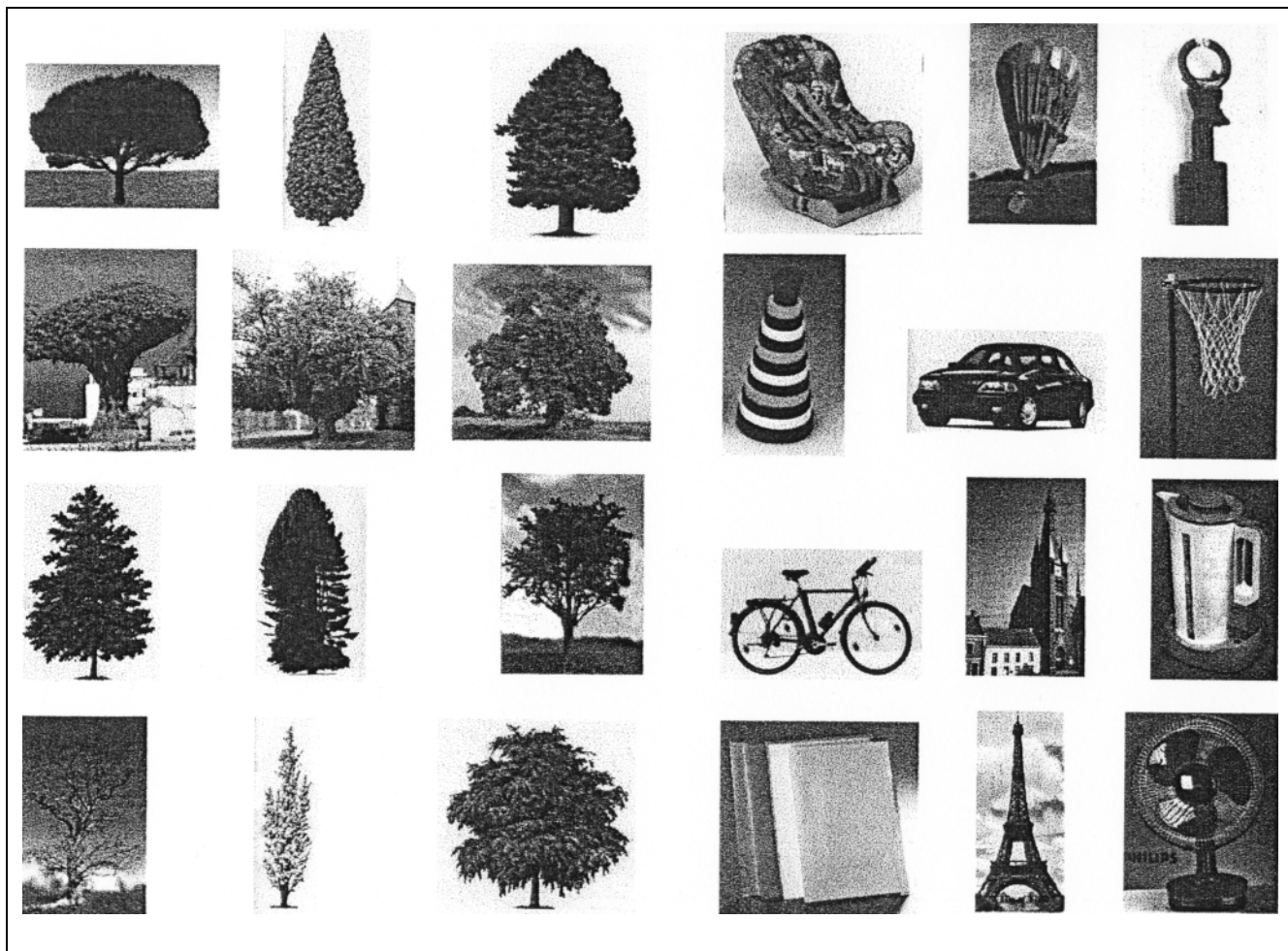


Figure 1. Some black and white examples of the original colored images used as tree and nontree images for the monkey. The three left-hand columns contain the tree exemplars while the three right-hand columns contain the nontree exemplars.

and shape were manipulated in order to rule out the possibility that the monkey used a single visual feature as cue.

Single unit recordings were made in the anterior part of the temporal cortex, including areas of the superior temporal sulcus and parts of the perirhinal cortex while the monkey performed the task (Vogels, 1999b). As mentioned above, in order to limit rote learning, only a limited subset of the entire set of tree and nontree stimuli was presented to the monkey for each daily recording session. Only 60 stimuli (half of them tree stimuli) were randomly selected from the stimulus set for each daily recording session. As a consequence, the responses of the neurons to all the stimuli were not measured and the recordings for all neuron were not done with the same stimulus set. The response of each neuron to a stimulus was defined as the average response over four to five stimulus presentations. Responses were recorded for 226 neurons.

The responses of the neurons were found to be highly image selective (Vogels, 1999b). Most of the neurons responded to both trees and nontrees. About one fourth

of the neurons responded in a category-specific manner, i.e., only responded to the members of one category. However, their response was only limited to particular exemplars of the chosen category.

Based on these recordings, a neuronal data matrix was constructed, consisting of spike counts computed with a 200-msec time bin starting 50 msec after stimulus onset (See Methods). For each neuron, the responses to 60 stimuli were obtained in the manner described above. The neuronal response matrix was more than 80% incomplete. As previously mentioned, this was due to the fact that not all the stimuli were tested for all the neurons and not all the neurons were tested with the same stimuli.

The first purpose of our investigation was to determine whether there was sufficient information in the responses of this population of IT neurons to account for the behavioral success of the monkey. A severe limitation in the ability to answer this question was the incomplete neuronal response matrix. A Kohonen self-organizing map (SOM), which employs an unsupervised (competitive) learning rule, was used to deter-

mine if the IT population activity is sufficient to classify the images into two groups, namely trees and nontrees (Van Hulle, 2000; Kohonen, 1982, 1989). A Kohonen SOM is a network in which every node of the network receives the same stimulus (i.e., network input). The output node that responds the best to the network input is called the “winner” and its synaptic weights (i.e., the weights from the input layer to the output layer) will be modified in such a manner as to respond better in the future to the same input. This is called competitive learning and is an example of unsupervised learning. It is intended to identify the presence of clusters in the input. In our case, each input pattern corresponds to the responses of 226 neurons to each stimulus. The clusters then correspond to the tree and nontree classes. Any missing data in the neuronal response matrix were treated by simply ignoring the missing entry. This method had been reported by Samad and Harp (1992) as one that had been successful in dealing with the problem of classification in cases where only partial data were available.

All of our cells were stimulus specific, i.e., the average cell only responded to approximately 17% of the images. In addition, most neurons were not category specific, i.e., they responded to both trees and nontrees. It was therefore not a logical necessity that the network would be able to classify the stimuli. Success in this endeavor would indicate that even though individual neuronal responses could not distinguish between tree and nontree categories, the correct categorization was available at the level of the neuronal population. Indeed, the Kohonen SOM can be viewed as a sort of clustering method that determines, in an unsupervised manner, to what degree a tree category is present in the neuronal response matrix. Thus the Kohonen SOM was used as a statistical tool and its architecture does not need to correspond to that of IT or other parts of the nervous system.

A second purpose of the Kohonen SOM analysis was to identify the neurons most important to the classification. The first class of neurons whose importance was investigated was the class of category-specific neurons. This group of neurons only responded to exemplars of one category. This class of neurons was found not to be critical in the categorization by the Kohonen SOM. Instead, the neurons critical for the categorization were found to be the more broadly tuned cells, most of which responded to both categories but which favored one category over the other by seven or more test stimuli (D7 neurons).

Finally, we attempted to identify an algorithm with which the responses of the D7 neurons could be used to distinguish a tree and nontree stimulus without the use of the Kohonen SOM.

A report of some preliminary results from this investigation has been presented (Thomas, Van Hulle, & Vogels, 1998).

RESULTS

The first test we applied was designed to determine whether the collected neuronal data were sufficient to account for the behavioral responses of the monkey. Since none of the recorded neurons responded to all the presented members of either category, the encoding of a category had to be accomplished as the result of the activity of a population of neurons. The ability of the collective neuronal response to encode for categorization was determined with the help of the Kohonen SOM. Next, we determined whether the population of category-specific neurons was critical to the categorization. The importance of any class of neurons was determined by the performance of the Kohonen SOM after the selective elimination of the responses of the concerned class of neurons from the response matrix. Controls for tests of this sort were performed by randomly eliminating the same number of cells from the response matrix, and evaluating the performance of the Kohonen SOM. The neurons most crucial to the performance of the network were then identified by noting the elements in the Kohonen node weight vectors that contributed the most to the difference in the activities of the two Kohonen nodes. The importance of these neurons is demonstrated and their properties described. We then characterized the important neurons by their difference index and demonstrate the importance of a subset of these neurons in the categorization by the Kohonen SOM. Finally, we identified a simple additive operation between this subset of neurons that is capable of giving a good amount of categorization success.

Categorization Success Achieved with Neuronal Responses

The neuronal responses were first analyzed to determine if they could account for the behavioral ability to categorize between tree and nontree stimuli. The analysis was done with the use of a two-node Kohonen SOM in which the missing data were ignored. A training phase with a total of 200 tree and nontree stimuli was followed by a testing phase with 200 previously unseen tree and nontree stimuli. The stimuli in each case were randomly picked. The mean categorization success obtained by using this method was $83.66 \pm 2.47\%$ (mean \pm SD). The highest categorization success obtained using the method was 88%.

The same tests were repeated with the use of a larger number of nodes in the Kohonen SOM. In this case, an additional layer that provided supervised learning was included in the network (i.e., a perceptron, see Hertz, Krogh, & Palmer, 1991 and the Methods section). The maximum number of nodes used in the unsupervised layer was 25. The use of the additional number of nodes was not found to improve the performance of the network.

The results from the use of the Kohonen SOM demonstrate that much of the information necessary to account for the behavioral ability of the monkey is available in the spike counts of the IT neurons. A mean categorization success of 83.66% could be accounted for without the use of more detailed information such as the temporal pattern of the spike discharges. Since all the neuronal recordings were obtained with the use of single unit recordings, the results also demonstrate that a precise synchronization of the neuronal firings in the IT cortex is not necessary to achieve this level of categorization.

The categorization success with the use of the Kohonen SOM was obtained with a response matrix that had more than 80% of its entries missing. These results also suggest the use of a code in the IT cortex that is redundant and very robust in the presence of noise.

Are the Category-Specific Neurons Critical to Categorization?

Further tests were done in order to determine how critical the category-specific neurons were to categorization. A neuron was counted as having responded to an image only if the response was at least a third of the maximal cell response. Using this criterion, a category-specific neuron, as the name suggests, responded only to images of one category with at least a third of its maximum response.

The question of whether the category-specific neurons were critical to categorization was analyzed by eliminating the responses of this class of neurons from the input vectors applied to the Kohonen SOM. The number of category-specific neurons was 57. The average categorization success achieved with this pruned response matrix was $79 \pm 1.89\%$ ($n = 30$). Although significantly different (using t test, $p < .01$), the categorization success following the elimination of the category-specific neurons, was only marginally less than the original categorization success of $83.66 \pm 2.47\%$ reported in the Results section, Categorization Success Achieved with Neuronal Responses. In addition, the same tests done with the random elimination of an equivalent number of neurons yielded an average categorization success rate of $80.9 \pm 2.14\%$ ($n = 30$) (Figure 2). These results demonstrate that the population of category-specific neurons does not play a crucial role in the encoding of the tree and nontree categories. Little deterioration in the performance of the Kohonen SOM was observed with the deletion of the category-specific neurons. A similar deterioration was obtained by the random elimination of an equivalent number of neurons from the response matrix.

Although the population of category-specific neurons was not critical to the categorization success of the Kohonen SOM, this does not mean that individual neurons in the group, when present, did not contribute

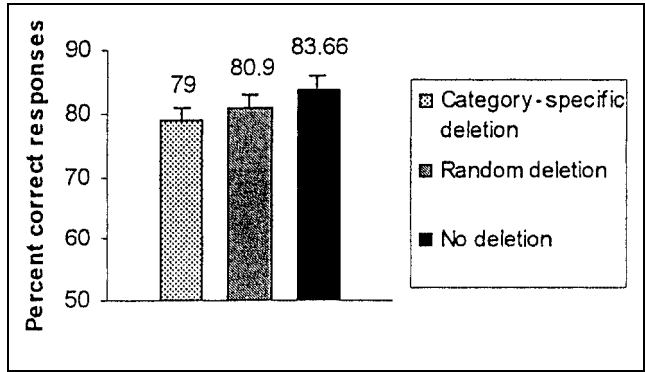


Figure 2. The deletion of the category-specific neurons from the neuronal response matrix did not deteriorate the categorization performance of the Kohonen SOM much more than the original performance of the Kohonen SOM with no deletions. A random deletion of the same number of neurons was also found to slightly decrease the categorization success of the Kohonen SOM. There were 57 category-specific neurons.

to the categorization. The results in the section below will demonstrate that some neurons of the category-specific group could indeed make important contributions to the categorization.

Category and Stimulus Selectivity of the Neurons Important to the Categorization

In the Methods section we have described how the Kohonen SOM gradually adjusts its weights in order to best match the incoming input. This eventually leads to the development of a tree and nontree node in the Kohonen SOM. An understanding of which neurons played the most crucial role in the categorization was obtained by noting which neurons contributed most to the difference in the weight vectors of the tree and nontree node in the Kohonen network. The following operation was performed on the average weight vectors of the tree and nontree nodes of the Kohonen network, once the training phase was over. The average weight vector \mathbf{w} for each node was computed after 30 trials. The “importance” value, I_i , for each neuron was then computed in the following manner

$$I_i = (\mathbf{w}_i^T - \mathbf{w}_i^{NT})^2 \quad (1)$$

where i could vary from 1 to 226. The vector \mathbf{w}^T was the weight vector for the tree node while \mathbf{w}^{NT} was the weight vector for the nontree node. The value I_i was found to lie between 0.00011 and 0.12255. The value 0.03 was chosen as a cutoff point, i.e., if $I_i > 0.03$, the neuron was identified as “important.” At cutoff points above 0.03, the number of “important” neurons was too small for their elimination to affect the categorization process very much.

Using this method, 47 neurons were identified as contributing critically to the categorization performance of the Kohonen SOM. The deletion of these neurons

from the neuronal response matrix was found to lead to a great deterioration in the performance of the network: the categorization success of the Kohonen SOM fell to $64.5 \pm 5.5\%$ ($n = 30$). A random deletion of the same number of neurons from the response matrix however, did not lead to any big deterioration in the performance of the SOM: the categorization performance of the Kohonen SOM remained at the good level of $82.7 \pm 2.6\%$ ($n = 30$).

Figure 3 presents the category selectivity and sparseness indices (defined in the Methods section) of these important neurons. The category-selective index of a neuron is a measure of the extent to which a neuron favors one category over the other. The sparseness index is a measure of the overall stimulus selectivity of the neuron, independent of the category membership of the stimulus. Figure 3A presents the characteristics of all the 226 neurons, while Figure 3B presents the characteristics of the neurons for which $I > 0.03$. The majority of these neurons were found to be moderately to highly category selective (Figure 3B). The average category selectivity of the neurons in Figure 3B was 0.81 ± 0.15 . The mean sparseness index for these “important” neurons was 0.47 ± 0.22 . We divided the neurons into four neuronal groups (see Figure 3): the category-selective but sparse group (sector I), the category-selective but nonsparse group (sector II), the noncategory-selective but sparse group (sector III) and, finally, the noncategory-selective and nonsparse group (sector IV). As illustrated in Figure 3, the lines dividing the graph into the four sectors are the halfway point for category selectivity, 0.75, and the halfway point for sparseness, which is 0.5. As is evident from Figure 3 and a χ^2 test, the properties of these neurons that contributed the most to the weight differences in the tree and nontree node, did not simply reflect the general properties of all the recorded 226 neurons ($\chi^2(3) = 14.91$, $p < .01$) (The χ^2 test was carried out by first computing the expected cell numbers in each of the four subdivisions in the case

where representation in the “important” cell population was proportional to that in the total population. A comparison was then made of these expected values and those seen in Figure 3A). For example, in comparing Figure 3A and B, only 3% of the neurons from sector III in Figure 3A were represented in Figure 3B. Figure 3B demonstrates that the neurons contributing the most to the categorization are the category-selective but sparse neurons ($n = 24$). This is followed by the noncategory-selective, nonsparse group ($n = 14$). This is then followed by the category-selective but nonsparse group ($n = 7$). The group that contributes the least is the noncategory-selective but sparse group ($n = 2$).

It should also be noted that a few category-specific neurons were in the group of neurons making important contributions to the difference in the weight vectors of the Kohonen nodes. These are the neurons for which category selectivity is 1.0. This demonstrates what was said in the previous section about individual category-specific neurons contributing to the categorization, even though the population of category-specific neurons is not critical to the categorization.

The Difference Index

Approximately 75% of the abovementioned neurons identified by the Kohonen network as neurons contributing the most to the categorization were found to belong to a group that responded to both categories, but which favored one category over the other by seven or more images. We called this difference in response the difference index of the neuron. The D7 neurons were defined as the class of neurons for which the absolute value of the difference between the number of responses to tree images and the number of responses to nontree images was seven or more.

The crucial role played by the 49 D7 neurons was further confirmed by testing the performance of the Kohonen SOM once these neurons were eliminated

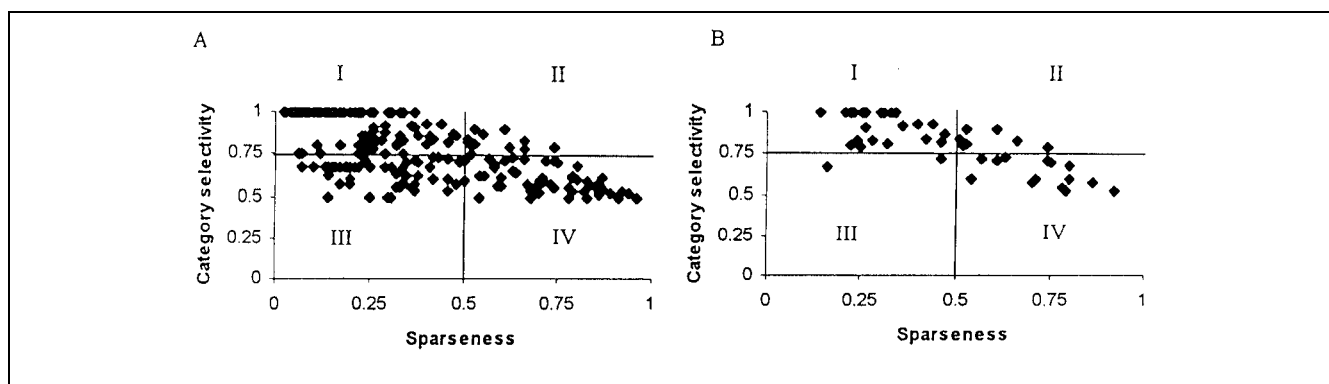


Figure 3. Scattergram describing the category selective and sparseness properties of the neurons. On the x-axis are the sparseness values, while the y-axis specifies the category selectivity. The horizontal line at 0.75, marks the halfway point in category selectivity. The vertical line at 0.5 marks the halfway point in sparseness. (A) Sparseness and category selectivity of all the recorded neurons. (B) Sparseness and category selectivity of neurons contributing the most to the difference between tree and nontree node of the Kohonen SOM.

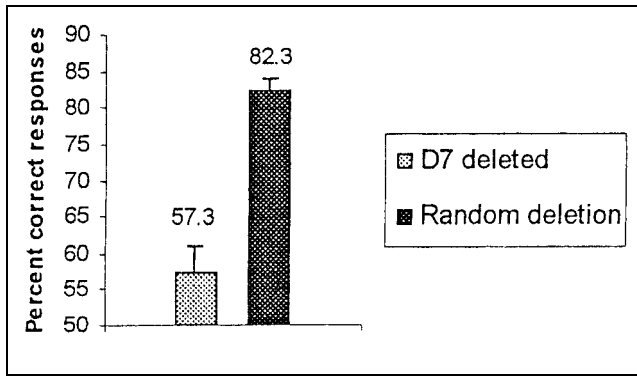


Figure 4. (2)The network performance was found to deteriorate greatly with the deletion of the broadly tuned D7 neurons.

from the response matrix. The performance of the SOM was found to drop to an average of $57.3 \pm 3.9\%$ ($n = 30$) correct. The control for this test was performed by randomly eliminating the same number of neurons in the response matrix. The network performance remained high with a mean of $82.3 \pm 1.67\%$ ($n = 30$) with this random elimination of neurons (Figure 4).

The difference in the role of the category-specific neurons and these more broadly tuned noncategory-specific cells cannot be attributed to a difference in cell numbers. There were, in fact, more category-specific neurons than D7 neurons. There were only 49 D7 neurons while the number of category-specific neurons was 57.

The network performance was not found to deteriorate in a completely discontinuous manner with the removal of the D7 neurons. We also tested the role of neurons that responded to at least more than 8, 9, and 10 more images of one category than the other. We will designate these neurons with the use of the letter D and a number to indicate how much greater the response to one category was compared to the opposite category. So, e.g., if a particular neuron responds to at least 10 more trees than nontrees, it is a D10 neuron. The network performance was found to deteriorate in a rather consistent manner as the neurons D7 to D10 were eliminated. Figure 5 displays the mean categorization performance of the network in each case once the neurons in question have been eliminated from the input response matrix. As is evident from the definition of the difference index, there is a continual increase in cell numbers as we go from D10 to D7. Figure 5 therefore demonstrates the gradual deterioration in the performance of the network with the increasing damage of the network.

Categorization Without the Use of the Kohonen SOM

Once we had determined that the D7 neurons were crucial to the process of categorization, we attempted to

identify an algorithm, using the output of these neurons, which could be used to categorize the images as tree or nontree without the use of a Kohonen SOM. For each neuronal response vector to a stimulus, the following simple additive operation on the D7 neurons of the vector was found to give a good degree of categorization success:

$$\Delta = \sum_{i=1}^N \mathbf{x}_i^{\text{D7-T}} - \sum_{j=1}^M \mathbf{x}_j^{\text{D7-NT}} \quad \Delta > 0 = \text{Tree}$$

$$\Delta < 0 = \text{Nontree} \quad (2)$$

As is evident from the definition of the D7 neurons, some of these neurons responded to more trees than nontrees (tree-favoring neurons) while others responded to a larger number of nontrees (nontree-favoring neurons). The value $\mathbf{x}^{\text{D7-T}}$ therefore is the nonnormalized neuronal response of a tree favoring D7 neuron and $\mathbf{x}^{\text{D7-NT}}$ is the nonnormalized response of a nontree favoring D7 neuron. N is the number of D7 neurons that responded to more trees than nontrees and M is the number of nontree-favoring D7 neurons. This operation was found to correctly classify the tree and nontree neuronal responses for 80% of the stimuli. The use of normalized neuronal responses was not found to increase the number of correct responses.

The same algorithm was applied to the classes of neurons ranging from D1 to D10. It should be recalled that since the class D1, for example, contains all the neurons whose number of responses to the tree class is greater than the number of responses to the nontree class (and vice versa) by at least 1, it also includes all the classes from D2 to D10. Starting from class D10, the categorization performance using Equation 2 was found to undergo relatively large improvements until the class D7 was reached. Although the performance continued to improve further from the class D7 to the class D1, the increase in performance was not as high as in the

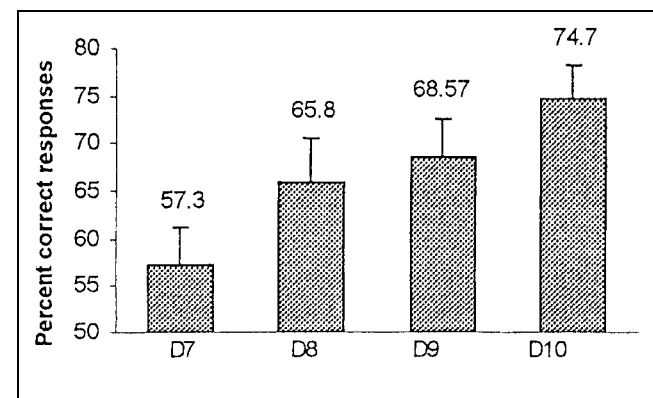


Figure 5. The network performance was found to decrease in a continuous manner as the neurons of category D10–D7 were deleted from the neuronal response matrix.

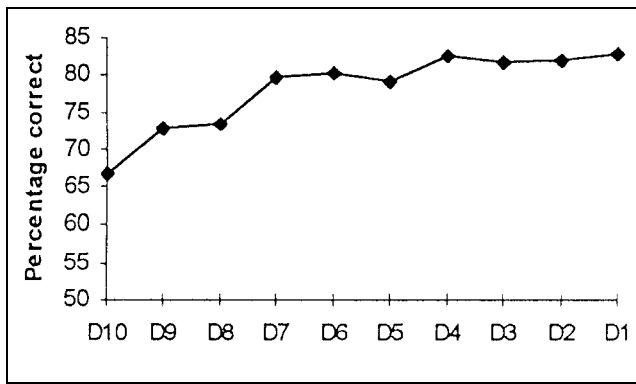


Figure 6. Categorization success with the various classes of neurons D10–D1, utilizing the algorithm described in Equation 2.

transition from D10 to D7 (Figure 6). The slope from D10 to D7 was 3.98, while the slope from D7 to D1 was 0.56. The level of categorization success achieved using the additive algorithm of Equation 2 with the D1 class of neurons was 83%.

DISCUSSION

The discussion is organized as follows. We will first discuss the significance of the finding that a good level of categorization was obtained with data from the area IT. Secondly, we will discuss some possible reasons for the discrepancy in the categorization obtained with the neuronal data versus the behavioral response. Thirdly, we will discuss some possible reasons for the fact that the responses of the category-specific neurons are not critical to the categorization of the network. Following this, we will speculate on why the class of category-selective but nonsparse neurons did not make a more prominent contribution to the categorization. We will then discuss future research that should be done to investigate the importance of the D7 class of neurons in more general cases. Finally, we discuss the gradual degradation in the performance of the Kohonen network as seen in Figure 5 in the light of sparse or semidistributed encoding.

Significance of Categorization Success by the Kohonen SOM

With the use of the Kohonen SOM, we were able to verify that the information in the IT cortex is sufficient to perform the categorization with a mean success of $83.66 \pm 2.47\%$. We could also have considered a network model for classifying trees and nontrees based on information in the striate cortex (V1). This however would have been a very difficult modeling task since the V1 neurons have small, localized receptive fields and respond to much simpler features. The neurons of area V1 are much less

invariant to translation and scale, and hence, even a change in size of the stimulus will lead to large changes in the representations of the trees and the nontrees. The purpose of the model developed in this article was to demonstrate that most of the information necessary to distinguish trees from nontrees was available in the responses of the IT neurons, which are known to show a high degree of invariance and to code for more complex features than V1 neurons. A greater generalization can therefore be expected from the responses of the IT cortex compared to the striate cortex.

As the categorization was achieved by means of only a two-node Kohonen SOM it indicates that only an additional layer is required to extract the categorization information from the IT cortex. The ability to obtain an 83% success in categorization with the use of a response matrix that was more than 80% empty also indicates the use of a code in area IT that is redundant and therefore robust in the presence of noise.

Comparison of Performance by Monkey and Kohonen SOM

While in individual cases the level of categorization success by the Kohonen SOM went as high as 88%, the level of behavioral success by the monkey in the categorization was higher than 95% (Vogels, 1999a). A number of possibilities remain open to explain this discrepancy in performance. The largest discrepancies can most likely be explained by the fact that a large amount of data were missing from the response matrix. Indeed, the neuronal response matrix obtained from the monkey recordings was more than 80% empty. Upon analysis of the data, we had found that the class of D7 neurons formed a high percentage of the neurons most important to the categorization. For these 49 neurons that were identified by the Kohonen SOM, a simple additive operation performed on them confirmed their importance since it yielded a categorization success of 80%. Additional recordings from more neurons in this class of neurons might have more closely approximated the observed behavioral success in categorization.

Since this categorization was achieved without the use of any information on the temporal patterns of the neuronal spike trains, or any attention to the precise timing of the neuronal spikes, we may conclude that this information is not necessary for this level of categorization success. There was, however, a discrepancy between the behavioral performance of the monkey and the level of categorization achieved with the neuronal data. While the large amount of missing data is likely to explain a good deal of this discrepancy, another possibility is that information on the temporal patterns of the spike trains is required to make more complete decisions on the presented images. Certain investigators,

e.g., Richmond, Optican, Podell, and Spitzer (1987) have presented evidence for the representation of stimulus information in the temporal modulation of the spike train in the IT cortex. However, others such as Shadlen and Newsome (1998) have argued that cortical neurons are unlikely to transmit information in the temporal arrangement of their spike discharges.

It should be noted here that since neither response variability nor measures of response covariance between different neurons were included in the input to the Kohonen SOM, we cannot determine how many neurons are needed to successfully solve the categorization problem. Other studies (e.g., Gochin, Colombo, Dorfman, Gerstein, & Gross, 1994) suggest that discrimination of a small set of highly different images can be accomplished using a relatively small number of IT neurons.

Questions like whether the network misclassified the same stimuli as the monkey could not be addressed. The neuronal responses were obtained while the monkey was making almost no mistakes. In the behavioral studies, the percentage of correct responses was between 97% and 99%. More seriously, the few mistakes made were not consistent misclassifications (see Vogels, 1999b).

Importance of the Category-Specific Neurons

The deletion of the category-specific neurons was not found to affect the categorization much more than a random deletion of the equivalent number of neurons (Figure 2). A very likely explanation for this may have been the extremely low sparseness index of these neurons (Figure 3A). Approximately 80% of these neurons had a sparseness index less than 0.25. Indeed, a comparison of the category-specific neurons in Figure 3A and B reveals that the category-specific neurons that made the most important contributions to the differences in the tree and nontree nodes for the Kohonen network tended to have higher sparseness indices than the other category-specific neurons. The mean sparseness index for all the category-specific neurons in Figure 3A was 0.147 ± 0.085 ($n = 57$). The average sparseness index for the category-specific neurons in Figure 3B was 0.259 ± 0.06 ($n = 11$). The mean sparseness index of the category-specific neurons that made more important contributions to the categorization was therefore significantly higher (t test, $p < .01$).

The Category-Selective and Nonsparse Neurons

Figure 3B demonstrates how the most important contributions to the categorization were made first by the category-selective but sparse neurons of sector I, followed by the noncategory-selective, nonsparse neurons of sector IV followed by the category-selective but nonsparse neurons of sector II. We might, however, have expected the neurons of sector II, the category-

selective and nonsparse neurons, to contribute the most to the difference in the tree and nontree node of the Kohonen SOM. We found, in fact, that when the percentage of each class contributing to Figure 3B is taken into account, sector II contributes the most. Approximately 58% of the category-selective but nonsparse neurons in sector II of Figure 3A were found in Figure 3B. In contrast, only 33% of the neurons from sector I and 27% of the neurons of sector IV were found in Figure 3B. The category-selective but nonsparse neurons may therefore have been expected to make a more important contribution to the categorization if they were present in greater numbers in the total neuronal population.

Future Research on the D7 Neurons

The class of D7 neurons was found to contribute greatly to the classification of the neuronal input. This is demonstrated both by Figure 5 as well as Figure 6. There is a deterioration in the performance of the Kohonen network to near-chance levels with the deletion of the D7 neurons. In Figure 6, it can be seen that only slight improvements in performance with the use of the additive algorithm (Equation 2) take place with the inclusion of more neurons after the D7 neurons. Future research will be required to determine if the class of D7 neurons are similarly important when variables such as the degree of completion of the response matrix are varied. Another factor that may be important in the crucial role of the D7 neurons is the maximum number of images to which the neurons responded. In the case of these experiments performed with tree and nontree images, the maximum number of images responded to by the D7 neurons was 14. This may change with other stimuli and change the apparent category-selective and sparseness properties of the D7 neurons.

Sparse or Semidistributed Encoding

The average sparseness index for all the recorded neurons in the study was 0.377 ± 0.24 . The average sparseness index of the neurons found to be important in the categorization was 0.47 ± 0.21 . The sort of encoding used for the tree/nontree discrimination therefore lies between a fully distributed code in which all the neurons participate (Churchland & Sejnowski, 1992), or grandmother cell encoding in which single neurons are narrowly tuned for a particular object (Barlow, 1972). Such a sparse or semidistributed encoding (Field, 1994) in the area IT has already been observed for the discrimination of face stimuli (Rolls & Tovee, 1995; Young & Yamane, 1992). Our study indicates that sparse or semidistributed representation in the area IT may be more general and apply in the case of nonface stimuli as well. Another study indicating that this may be the case was also done earlier with

a much smaller sample of nonface stimuli (Gochin et al., 1994). The deletion of neurons from the D10–D7 categories leads to a progressive deterioration in the performance of the Kohonen SOM. As is evident from the definition of the difference index, the D10 neurons are a subset of the D9 neurons, which are in turn a subset of the D8 neurons, and so on. There is, therefore, a gradual increase in cell numbers as we go from D10 to D7. Hence, Figure 6 demonstrates the gradual deterioration in the performance of the network with the elimination of increasing numbers of these important neurons. Such a graceful degradation in the coding abilities of a network has been described as one of the benefits of a sparse or semidistributed encoding (Field, 1994; French, 1992, 1999; Treves & Rolls, 1991; Rolls & Treves, 1990).

GENERAL CONCLUSIONS

The neuronal responses from the IT cortex of monkeys while they performed a visual categorization task were analyzed with the use of a Kohonen SOM. The assigned task was to distinguish tree from nontree stimuli. The neuronal response matrix consisted of the averaged spike counts of 226 neurons to subsets of a set of 200 trees and 200 nontrees. The results of the analysis show that, even though the information necessary to make this distinction was not available at the level of the single neuron, much of the data necessary to make the classification is available at the level of the population. Since the analyzed neuronal responses consisted of averaged spike counts from single-unit recordings, the results indicated that aspects of neuronal firing such as temporal patterns or precise synchrony between the firing patterns of different neurons was unnecessary to achieve a mean 83% categorization success. Since this level of categorization was achieved with a neuronal response matrix that was more than 80% incomplete, the categorization success also indicates an encoding that is redundant and robust to noise.

About one-fourth of the recorded neurons were category specific and only responded either to the tree or to the nontree images. The Kohonen SOM analysis indicated that the responses of the class of category-specific neurons is not critical to the categorization. Instead, the majority of the class of neurons that was critical for categorization responded to both classes but favored one class over the other. These results therefore indicate the importance of the more broadly tuned, feature detecting (almost certainly complex features) neurons over the category-specific neurons in the classification.

The average sparseness index of the neurons critical to categorization was 0.47 ± 0.21 . Since the stimuli consisted of nonface stimuli, the study therefore helps to demonstrate that the use of sparse or semidistributed coding is a general mode of coding in the IT cortex.

METHOD

The Kohonen Self-Organizing Map

A Kohonen unsupervised (competitive) learning algorithm was used in order to analyze the neuronal response matrix. In an unsupervised learning algorithm, the algorithm must discover patterns and categories in the input data by itself, and must code them in the output (Van Hulle, 1997, 2000; Hertz et al., 1991). In particular, a Kohonen self-organizing feature map (Kohonen, 1982, 1989) was constructed for the analysis of the neuronal data. The Kohonen SOM uses an unsupervised, competitive learning rule that makes use of a winner-takes-all approach.

The input to each node of the Kohonen SOM consisted of the responses of 226 neurons to approximately 200 trees and 200 nontrees. Spikes were counted in two 200-msec-long time windows, one before stimulus onset (baseline) and a second one starting 50 msec after stimulus onset. Net responses were calculated for each trial by subtracting the spike counts obtained in these two time windows. Since the behavioral responses of the monkey already occurred with a 250-msec delay (mean 241 msec) (Vogels, 1999b), a longer time window was not used. An extensive analysis using shorter time windows of less than 150 msec (see Vogels) showed highly similar stimulus selectivity profiles to those measured with the 250-msec time window. In fact, the sparseness and category specificity indices obtained with the short time windows of Vogels correlated strongly with those obtained with the longer 200-msec-long time window of the present paper (Pearson correlation coefficients of 0.95 and 0.96 for the sparseness and category specificity indices, respectively). Thus, the selectivity profiles of the neurons are largely independent of the duration of the time window used to compute the spikes, at least within this 150- to 200-msec range. These and other analyses (Vogels) suggest that the behavioral response itself did not contribute in a significant manner to the neural responses measured in these time windows.

The firing rates entered into the response matrix were average, net firing rates from four to five trials. Thus, we did not take into account the trial-to-trial variability of the responses to a stimulus. The latter only becomes critical if one wants to determine how many neurons are needed to solve a task. This, however, was not the purpose of the present study. The aim of this study was to determine whether categorization could be achieved using neurons having the selectivity profile reported by Vogels (1999b) and to investigate which sort of selectivity profile (e.g., degree of stimulus selectivity and category selectivity) might be most important for categorization. Furthermore, the inclusion of response variability would most likely have only affected the classification of stimuli on the border between the tree and nontree classes.

All the neuronal input to the Kohonen SOM was scaled so that the maximum activity of each neuron to all stimuli was 1.0, while the minimum activity was 0.

$$m^k = \sum_{i=1}^n (\mathbf{x}_i - \mathbf{w}_{ki})^2$$

$\forall \mathbf{x}_i$ for which there are input values (3)

The winning unit in the Kohonen SOM was chosen by determining the node that best matched the input stimulus. The degree of match m^k for the node k of the Kohonen SOM was computed by obtaining the squared Euclidian distance between the input stimulus vector \mathbf{x} and the Kohonen SOM weight vector \mathbf{w}_k of the node k . Since the input to the network consisted of the neuronal responses from 226 neurons, n had the value 226. As indicated in the equation above, all computations were only performed for the case where \mathbf{x}_i had been measured (Samad & Harp, 1992). The network was first trained on a set of 200 trees and nontrees. The winning node c was the node with the closest match to the neuronal input vector \mathbf{x} , i.e., the least Euclidian distance from the input stimulus vector

$$m^c = \min(m^1, m^2, \dots, m^n) \quad (4)$$

Weight modification was done not only for the winning node but also for the surrounding nodes based on the distance $d(k, c)$ between the network coordinates of the winner c and the unit k under consideration

$$\Delta \mathbf{w}_{ki}(t) = \eta(t) \exp[-d^2(k, c)/2\sigma^2(t)] \times (\mathbf{x}_i(t) - \mathbf{w}_{ki}(t))$$

$\forall \mathbf{x}_i$ for which there are input values (5)

Weight alteration, also, only applied in the cases for which \mathbf{x}_i had been measured.

The initial weights of the Kohonen SOM nodes were randomly drawn between 0 and 1.0. The parameters η and σ were reduced over the course of the training. The parameter η , e.g., was reduced in the following manner

$$\eta(t) = \eta_i (\eta_f / \eta_i)^{t/T} \quad 0 \leq t \leq T \quad (6)$$

The parameters used were $\eta_i = 1.0$, $\eta_f = 0.03$, $\sigma_i = 10.0$, $\sigma_f = 0.1$ (Samad & Harp, 1992). The value of T was 2000. If there was sufficient data for the classification, in a two-node Kohonen SOM, the weight vectors would evolve so that one node would eventually become a “tree” node while the other became a “nontree” node.

The network was trained on a sample of 200 trees and nontree images. The number of cycles used in training were fixed at 2000. This was sufficient to ensure that the activities of the nodes in the Kohonen SOM did not change by more than 10^{-4} in the immediately succeeding cycle. Once the network was trained on a sample of 200 tree and nontree images, it was tested on a set of 200 previously unseen images.

In the cases where more than two nodes were used in the Kohonen SOM, an additional layer was added for the final tree, nontree decision. This additional layer consisted of a single decision node using backpropagation in order to alter the weights between it and the unsupervised layer (Hertz et al., 1991).

The number of trials for each test performed using the Kohonen SOM in the Results section was 30.

Category-Selective Index

The category selectivity index of a neuron indicated the degree of preference of the neuron for a tree or nontree class. A response to a stimulus was counted if the neuron had responded to that stimulus with at least one third of its maximal response. Thus, e.g., if a neuron had responded to five trees and seven nontrees, it had a category selectivity of 0.58 in favor of the nontrees. The category selectivity was therefore the ratio of the number of responses of the neuron to the preferred class, to the total number of responses of the neuron.

Sparseness Index

The sparseness index using n stimuli was computed as follows:

$$\text{Sparseness} = \left[\sum_{i=1, n} (R_i/n) \right]^2 / \left[\sum_{i=1, n} (R_i^2/n) \right] \quad (7)$$

with R_i being the net response to stimulus i (with negative net responses clipped to zero, as described by Rolls and Tovee, 1995). The sparseness index can have a maximum value 1.0 in the situation where a neuron responds to all presented stimuli.

Acknowledgments

This research was supported in part by the Geneeskundige Stichting Koningin Elizabeth, GOA, and FNRS of Belgium. Thanks to Robert French for many suggestions. M. M. V. H. is a Research Associate of the Fund for Scientific Research, Flanders (Belgium) and is supported by grants from the Fund for Scientific Research (G.0185.96), the National Lottery (Belgium) (9.0185.96), the Flemish Ministry of Education (GOA 95/99-06), and the Flemish Ministry for Science and Technology (VIS/98/012). Rufin Vogels is a research associate of the Fund for Scientific Research, Flanders (Belgium).

Reprint requests should be sent to Marc M. Van Hulle, Katholieke Universiteit Leuven, Faculteit Geneeskunde, Laboratorium voor Neuro- en Psychofysiologie, Campus Gathuisberg, Herestraat, B-3000 Leuven, Belgium.

REFERENCES

- Barlow, H. B. (1972). Single units and sensation: A neuron doctrine for perceptual psychology? *Perception*, 1, 371–394.
- Churchland, P. S., & Sejnowski, T. J. (1992). *The computational brain*. Cambridge, MA: MIT Press.
- Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., &

- Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature*, 380, 499–505.
- Field, D. J. (1994). What is the goal of sensory coding? *Neural Computation*, 6, 559–601.
- French, R. M. (1992). Semidistributed representations and catastrophic forgetting in connectionist networks. *Connection Science*, 4, 365–377.
- French, R. M. (1999). Catastrophic forgetting in connectionist networks. *Trends in Cognitive Science*, 3, 128–136.
- Gochin, P., Colombo, M., Dorfman, G., Gerstein, G., & Gross, C. (1994). Neural ensemble coding in inferior temporal cortex. *Journal of Neurophysiology*, 71, 2325–2337.
- Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in the inferior temporal cortex of the macaque. *Journal of Neurophysiology*, 35, 96–111.
- Hertz, J., Krogh, A., & Palmer, R. G. (1991). *Introduction to the theory of neural computation*. Redwood City, CA: Addison-Wesley.
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43, 59–69.
- Kohonen, T. (1989). *Self organization and associative memory*, 3rd ed. Berlin: Springer Verlag.
- Kovacs, G., Vogels, R., & Orban, G. (1995). Selectivity of macaque inferior temporal neurons for partially occluded shapes. *Journal of Neuroscience*, 15, 1984–1997.
- Richmond, B. J., Optican, L., Podell, M., & Spitzer, H. (1987). Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex: I. Response characteristics. *Journal of Neurophysiology*, 57, 132–146.
- Rolls, E. T., & Tovee, M. J. (1995). Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *Journal of Neurophysiology*, 73, 713–726.
- Rolls, E. T., & Treves, A. (1990). The relative advantage of sparse versus distributed encoding for associative neuronal networks in the brain. *Network*, 1, 407–421.
- Samad, T., & Harp, S. A. (1992). Self organization with partial data. *Network*, 3, 205–212.
- Sary, G., Vogels, R., & Orban, G. A. (1993). Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science*, 260, 995–997.
- Sato, T., Kawamura, T., & Iwai, E. (1980). Responsiveness of inferior temporal single units to visual pattern stimuli in monkeys performing discrimination. *Experimental Brain Research*, 38, 313–319.
- Schwartz, E. L., Desimone, R., Albright, T. D., & Gross, C. G. (1983). Shape recognition and inferior temporal neurons. *Proceedings of the National Academy of Sciences U.S.A.*, 80, 5776–5778.
- Shadlen, M. N., & Newsome, W. (1998). The variable discharge of cortical neurons: Implications for connectivity, computation and information coding. *Journal of Neuroscience*, 18, 3870–3896.
- Tanaka, K. (1993). Neuronal mechanisms of object recognition. *Science*, 262, 685–688.
- Thomas, E., Van Hulle, M., & Vogels, R. (1998). The study of the neuronal encoding of categorization with the use of a Kohonen network. *Proceedings, 5th neural computation and psychology workshop. Connectionist models in cognitive neuroscience*. Birmingham, England: Springer Verlag.
- Treves, A., & Rolls, E. T. (1991). What determines the capacity of autoassociative memories in the brain? *Network*, 2, 371–397.
- Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle (Ed.), *Analysis of visual behavior* (pp. 549–586). Cambridge: MIT Press.
- Van Hulle, M. (1997). The formation of topographic maps that maximize the average mutual information of the output response to noiseless input signals. *Neural Computation*, 9, 595–606.
- Van Hulle, M. M. (2000). *Faithful representations and topographic maps: From distortion-to information-based self-organization*. New York: Wiley.
- Vogels, R. (1999a). Categorization of complex visual images by rhesus monkeys: 1. Behavioral study. *European Journal of Neuroscience*, 11, 1223–1238.
- Vogels, R. (1999b). Categorization of complex visual images by rhesus monkeys: 2. Single cell study. *European Journal of Neuroscience*, 11, 1239–1255.
- Vogels, R., & Orban, G. (1996). Coding of stimulus invariance by inferior temporal neurons. *Progress in Brain Research*, 112, 195–211.
- Young, M., & Yamane, S. (1992). Sparse population coding of faces in the inferior temporal cortex. *Science*, 256, 1327–1331.