Representation of Number in Animals and Humans: A Neural Model

Tom Verguts and Wim Fias

Abstract

■ This article addresses the representation of numerical information conveyed by nonsymbolic and symbolic stimuli. In a first simulation study, we show how number-selective neurons develop when an initially uncommitted neural network is given nonsymbolic stimuli as input (e.g., collections of dots) under unsupervised learning. The resultant network is able to account for the distance and size effects, two ubiquitous effects in numerical cognition. Furthermore, the properties of the network units conform in detail to the characteristics of recently discovered number-selective neurons. In a second study, we simulate symbol learning by presenting symbolic and nonsymbolic input simultaneously. The same number-selective neurons learn to represent the numerical meaning of symbols. In doing so, they show properties reminiscent of the originally available number-selective neurons, but at the same time, the representational efficiency of the neurons is increased when presented with symbolic input. This finding presents a concrete proposal on the linkage between higher order numerical cognition and more primitive numerical abilities and generates specific predictions on the neural substrate of number processing. ■

INTRODUCTION

Basic numerical abilities have been demonstrated in a variety of animal species, with highly similar characteristics across species (Dehaene, Dehaene-Lambertz, & Cohen, 1998). First, when two numbers have to be compared, a distance effect, meaning that two numerosities are easier to discriminate when the distance between them is larger, is obtained robustly. Second, a size effect is obtained in the same tasks: For a given distance, comparison is easier when numerosities are smaller. Both effects hold for a variety of species (for a review, see Dehaene et al., 1998). Moreover, in humans, the distance and size effects are obtained with symbolic (Moyer & Landauer, 1967) and nonsymbolic stimuli (Buckley & Gillman, 1974). Hence, the distance and size effect do not depend on species or on knowledge of a symbolic system, such as language. This observation favors the hypothesis that a core system for numerical processing (a ''number sense''; e.g., Dehaene, 2001) has been internalized in the brains of various animal species under evolutionary pressure. The fact that these properties prevail when humans use symbols to express quantitative information has led to the hypothesis that high-level human numerical abilities are rooted in these biologically determined mechanisms by linking symbolic representational systems to evolutionary basic and abstract number representations (Dehaene, 2002).

Different hypotheses regarding the nature of this number sense have been put forward. Gallistel and Gelman (1992) proposed that internal number representations code quantity in a linear way: Equal distances between numbers are represented by equal distances in the representation. This assumption can explain the distance effect, because numbers that are close to each other will have overlapping distributions of activation, hence, it will be difficult to discriminate between them. To account for the size effect, Gallistel and Gelman assume scalar variability in the mapping from sensory input to this linear representation. This means that for larger numbers, variability is larger, leading to more noisy representations of larger numbers, and thus, the size effect follows.

As an alternative to scalar variability, compressed coding has been proposed, according to which numbers are represented on a scale that is more compressed for larger numbers (e.g., a logarithmic scale; Dehaene, 2003). This leads to a less accurate representation and a worse discriminability of large numbers relative to small numbers, and hence to a size effect. At a behavioral level, scalar variability and compressed coding lead essentially to the same predictions (Dehaene, 2003), but recent single-cell recording data (Nieder & Miller, 2003; Nieder, Freedman, & Miller, 2002) provide support for compressed coding.

Nieder et al. (2002) recorded from neurons in the lateral prefrontal cortex of rhesus monkeys during a visual delayed match-to-numerosity task (*numerosity* refers to the case where numerical information is conveyed by the number of objects in a display rather than by a symbol). The authors found that many of these neurons consistently preferred (responded most

Ghent University

strongly to) a specific number of objects. The critical properties of these number-selective neurons are the following. First, they act like filters over numerosity: Neurons that are most responsive to a particular numerosity x also react somewhat weaker to numerosities $x - 1$ and $x + 1$, still somewhat weaker to $x - 2$ and $x + 2$ and so on. This property can account for the distance effect, because if numbers are farther apart, this will lead to less overlap in the distribution of activation and thus to better discrimination. A second property is that tuning curves are increasingly broader for larger numerosities. This increasing bandwidth is consistent with the size effect, because larger numbers will have broader tuning curves, leading to more representational overlap, and thus will be more difficult to discriminate than smaller numbers. Third, these tuning curves are positively skewed: For example, for a neuron-preferring quantity x, its response rate was stronger when quantity $x + 1$ was presented rather than $x - 1$. As discussed in the authors' follow-up study (Nieder & Miller, 2003), the latter property is critical in differentiating between scalar variability and compressed coding. In particular, a scalar variability account predicts symmetric tuning curves, whereas compressed coding predicts the observed asymmetric positive skewing. In this way, Nieder and Miller (2003) provided evidence in favor of a compressed coding of numerosity.

Despite the significant progress that has been made by the discovery of these numerosity detectors and the detailed description of their processing characteristics, many questions remain unanswered. First, what kind of input is needed for these numerosity detectors? For example, neurons in parietal cortex are known to encode the spatial location of objects. Can numerosity detectors work on this information, or is an intermediate processing stage necessary? Second, can such numerosity detectors be learned? The fact that the ability to detect numerosity is present in a wide range of species and early in life (see Dehaene et al., 1998, for a review) seems to point to an innate sense of number (Dehaene, 2002). An alternative is that this ability is easily and quickly learned and is therefore omnipresent. Third, what is the relation with symbolic cognition? Several authors (e.g., Spelke, 2002; Dehaene, 2001) have suggested that human numerical abilities are built on a primitive number sense shared with other animals. If so, how is this accomplished?

In the present article, an answer is sought to these three questions using insights obtained from a modeling study. First, we argue for a specific type of intermediate representational layer between sensory input and numerosity detectors. Using this insight, we describe how numerosity detectors may arise in an initially uncommitted neural network. The network was equipped with an unsupervised competitive learning rule, and different numerosities (stimuli consisting of a number of objects) were repeatedly presented to this network. We found that the network developed numerosity detectors that are able to account for the distance and size effect, and the detectors exhibited the properties reported by Nieder and Miller (2003) and Nieder et al. (2002). This is described in Simulation 1. After the numerosity detectors had developed, we presented symbolic input (e.g., verbal number labels) in conjunction with the nonsymbolic numerosities in Simulation 2. The same network nodes that represented a given numerosity also learned to represent the value of the corresponding symbol. Similarities and differences in the neural coding of symbolic and nonsymbolic stimuli were observed. Implications for symbolic and nonsymbolic numerical cognition are discussed.

INPUT TO NUMEROSITY DETECTORS: EVIDENCE FROM BACKPROPAGATION

Individuation and enumeration of elements in a visual display depends on spatial processing. For instance, Trick and Pylyshyn (1994) showed that the ability to immediately extract numerosity from a visual display (i.e., subitizing) is hampered when the to-be-enumerated objects are presented concentrically at the same spatial position. Therefore, we take a simple spatial location coding system as input, in which one unit corresponds to one location. With regard to the numerosity detectors, a detector tuned to numerosity 1 should be activated when exactly one object in one location is presented; a detector tuned to numerosity 2 should be activated when two objects in two different locations are presented; and so on. However, it is impossible to obtain a direct linear mapping from such a spatial location coding system to a set of numerosity detectors without assuming additional processing mechanisms. The reason is that whether an input unit should or should not activate a specific numerosity detector depends on the activation of other input units. For example, activation of an input unit should lead to activation of a detector for numerosity 2, but only if exactly one other input unit is active as well. In this sense, numerosity detection is a generalization of the classical exclusive—or problem (Minsky & Papert, 1969). The fact that these problems are not linearly separable makes it impossible to solve them with a direct linear mapping from input to output values. The most straightforward solution (e.g., Minsky & Papert, 1969) is to assume an intermediate processing layer.

The next question, then, is what kind of intermediate layer would constitute a computationally efficient way to allow spatially individuated elements to be mapped onto numerosity detectors. To answer this question, we trained a backpropagation network that was presented the locations of objects in the input layer and was required to give the corresponding numerosity as output.

Model Description

The model is depicted in Figure 1A: In the input layer, there are five nodes, each corresponding to a particular spatial location. In the output layer, there are also five nodes, one for each possible numerosity. For example, if two input units are active, the required output is activation of the numerosity—two detectors in the output layer, with no activation for any other output unit. The hidden layer contains five nodes. Each hidden node receives a linear combination of the values in the input layer, which is then transformed according to a sigmoid function $1/[1 + \exp(-in - \theta)]$, where "in" represents the input to that node and θ is a threshold value. The output layer receives a linear combination of these hidden layer values, and it transforms its input with the same sigmoid function. For simplicity, no threshold was used for the output units, but results were similar with thresholds.

Learning Procedure

The backpropagation algorithm adapted all weights (input to hidden and hidden to output) and thresholds by minimizing a least-squares error function. After each adaptation step, weights were decreased by a fraction of 0.02%, which amounts to a weight decay procedure (Plaut, McClelland, Seidenberg, & Patterson, 1996). Weights were initially set at random values. The network was presented 100,000 trials.

Results and Discussion

After training, the output units learned to selectively respond to the presented numerosity with 100% accuracy (see Figure 1C for a graphical illustration).

An interesting property emerged at the level of the hidden field: All connection weights to a given hidden unit were approximately equally strong (mean standard deviation of the weight values to the different nodes was 0.05). This implies that each of the five hidden units responded in a monotonous manner to numerosity, four in an increasing way (with stronger activation for increasing number of objects, see Figure 1B for graphical illustration), and one in a decreasing way (with weaker activations for increasing number of objects). Based on these two properties, the hidden units can be called *summation units*, because each unit sums the activation of the input units without distinguishing between input units. Some hidden units add in an additive manner (more objects, more activation), and others "add" in a subtractive manner (more objects, less activation). In this way, summation units are sensitive to numerosity, but not selective to numerosity: They are a type of analog magnitude representation (Moyer & Landauer, 1967), with the amount of activation reflecting the presented numerosity. Hence, summation coding seems to provide an efficient intermediate step to generate the required numerosity-selective output. An intermediate layer with similar properties was assumed in the neuronal models of Ahmad, Casey, and Bale (2002) and Dehaene and Changeux (1993). Recently, Romo and Salinas (2003) have described neurons in the somatosensory, medial premotor, and prefrontal cortex, whose firing rate was monotonously related to the frequency of vibrotactile stimulation. This supports the biological plausibility of neuronal summation coding.

We do not assume that the backpropagation learning mechanism is a biologically plausible way for the creation of summation neurons; the simulation does, however, show that summation coding is a computationally natural solution to solve the problem of numerosity detection. A similar distinction between featural sensitivity and selectivity, as used here, has been demonstrated for other features, for example, speed processing: As

Figure 1. (A) Structure of backpropagation model. Amount of black in network node indicates amount of activation. (B) On the abscissa, the different possible numerosities are shown (1–5). The ordinate shows the mean activation of a hidden unit over all possible input layer configurations leading to that particular numerosity. Tuning curves of two representative hidden field units are shown. (C) See (B), but for two representative number field units.
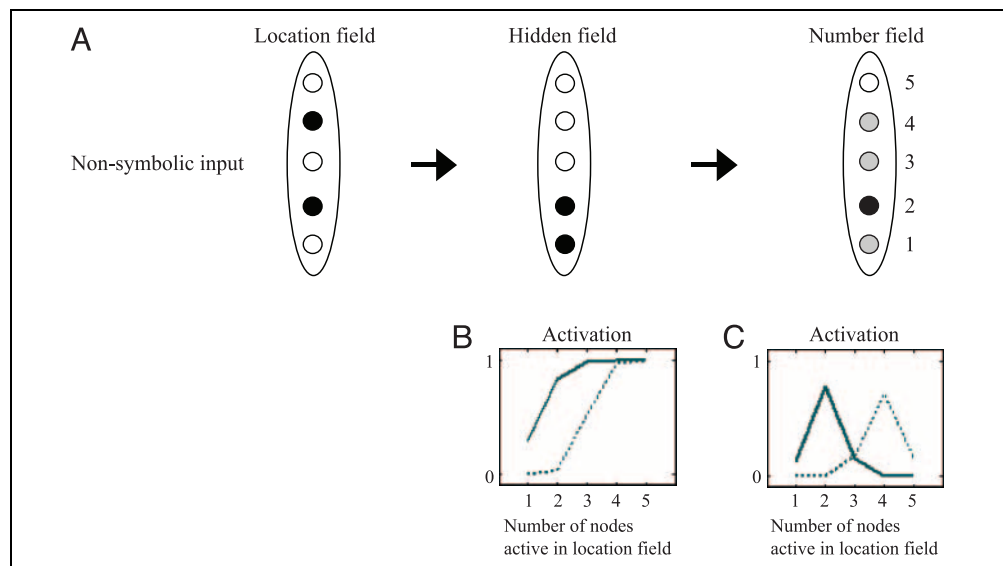
noted by Chey, Grossberg, and Mingolla (1998), speed-sensitive cells are found in the monkey retina and LGN, but speed-selective cells are only found higher up in the processing stream, in the striate cortex (Orban, Kennedy, & Bullier, 1986) and, particularly, in area MT (Maunsell & Van Essen, 1983). This further underlines the fact that a succession of featural sensitivity and featural selectivity is a biologically plausible processing chain to extract and represent gradually more refined information from raw sensory information.

UNSUPERVISED LEARNING MODEL

Given its computational efficiency and biological plausibility, we used summation coding as a starting point for the construction of an unsupervised learning model. With this model, we intended to answer the questions of learnability of number-selective neurons (Simulation 1) and the grounding of symbolic numerical representations (Simulation 2). The unsupervised model architecture is depicted in Figure 2. It consists of a summation field (similar to the hidden field of the previous section), a number field, and a symbol field. The number field receives input from the other two. Because Nieder et al. (2002) obtained neural recording data for numerosities 1–5 only, and because our aim was to model the neural data, we also used numbers from 1 to 5. However, very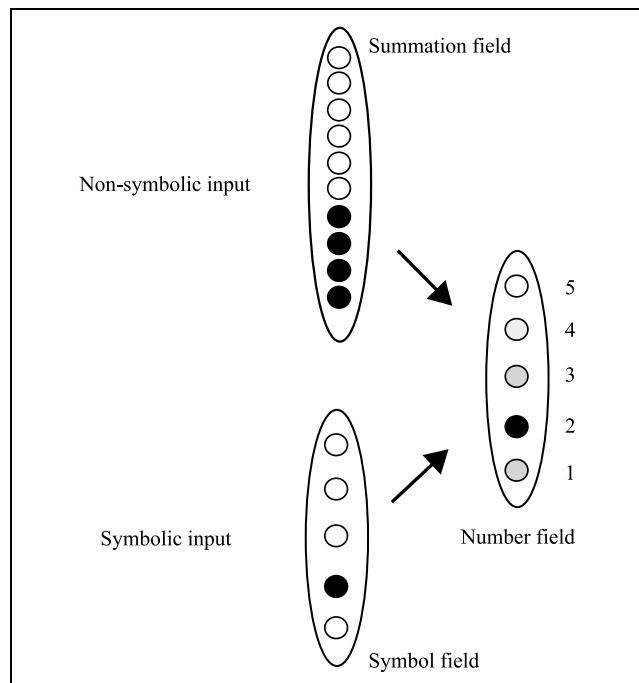 similar results were obtained when the network was also trained with larger numerosities. Each of the three fields is now discussed separately.

Summation Field

Clearly, a visual image first passes through a preprocessor before it is ready for numerical analysis. Because the operation of such a preprocessor is beyond the scope of this article, we did not implement this stage (see Chey et al., 1998; Dehaene & Changeux, 1993, for possible implementations). Rather, the model starts from the presumed result of this preprocessing stage, namely, a pattern of activation over a set of summation coding nodes (see above). This activation pattern is then normalized, meaning that the length of the activation vector equals 1. For example, if the pattern of activation is originally (1, 1, 0, 1, 0), with length $\sqrt{3}$, after normalization, it is $(1, 1, 0, 1, 0)/\sqrt{3} = (0.58, 0.58, 0, 0.58, 0)$, with length 1. Normalization is a common assumption in many computational models (e.g., McNellis & Blumstein, 2001) and can be obtained neurally by lateral inhibition within the summation field (Usher & McClelland, 2001).

The number of summation nodes activated upon presentation of a particular number of objects was equal to two times this number. This is an instance of the "additive" summation coding obtained in the backpropagation network. For example, when a display with one object was shown, two nodes were activated in the summation field; two additional nodes were activated upon presentation of two objects and so on (see Figure 2). Because numerosities could range from 1 to 5 in the simulations, 10 summation field nodes were implemented. Similar results were obtained with different numbers of summation nodes.

Symbolic Field

Symbols are related in an arbitrary manner to their corresponding quantity: The physical appearance of a symbol bears no numerical information. Similarly, two different symbols are related in an arbitrary manner to each other as they give no indication as to the difference between their respective quantities. To incorporate these two key properties, representations in the symbolic field for different numbers were chosen in an arbitrary manner. In particular, number 1 was coded by one arbitrarily chosen symbolic field unit, number 2 by another unit, and so on. Because our simulations used numbers 1 to 5, the symbolic field consisted of five nodes.

Number Field

Activation of a particular number field node is a linear function of the activation in the two input layers. In particular, we used $y = \sum_i w_i x_i$ where y is the activation value of a number field node, index i ranges over input nodes (in summation and symbolic fields), x_i is the

Figure 2. Structure of unsupervised learning model, with summation field, symbol, and number fields. Note that four hidden units are activated for numerosity 2, because in our implementation, we used a coding scheme where each extra object activates two extra summation field units.

activation of input node i, and w_i is the connection weight between node i and the number field node. There were 500 nodes in the number field. A relatively large number was taken here to make sure that averages calculated over nodes with a given number preference would be sufficiently stable.

Learning Rule

Connections between input field (summation or symbolic) and number field nodes were initially uncommitted: half of them were 0 and the other half had a value randomly selected from a uniform distribution over (0, 1). Unsupervised learning was applied on this originally uncommitted network. The following principles are implemented in the learning rule. After activation is passed on from an input field (summation field in Simulation 1; summation and symbolic field in Simulation 2) to the number field, each connection between an input and number field node is changed by an amount that depends on two factors: the activation of the number field node and the activation of the input node. Concerning the first factor, weights attached to more active number field nodes are changed more. The rationale is that active number field nodes already partially code for the input pattern and that therefore the number node's weights should be adapted to make them even more responsive to the input pattern. Weights connected to less active number units are changed to a lesser degree. With respect to the second factor, the principle is that a weight will change more if it deviates more from the input value. If a weight is already close to the activation value of the input node, it is not substantively changed. In particular, if the activation value of an input node equals x and the weight from this node to a number field node is w, the difference $x - w$ will partly determine how strongly the value w should change. If w is already close to x, then $x - w \approx 0$, and the resultant change will be small.

Formally, weight adaptation was implemented as follows. The weight w_{ij} from input node i to number field node j is updated as

$$\Delta w_{ij} = \alpha \times \exp[-\beta(y_{max} - y_j)] \times (x_i - w_{ij}) \qquad (1)$$

where x_i is the activity of input field node i and the parameter α is a fixed learning rate set at 0.25. The factor $\exp[-\beta(y_{max} - y_j)]$ indicates how strongly this particular weight should change relative to the other weights. Here, y_j is the response of number field node j, and y_{max} is the maximal response over the number field nodes. The constant β is a sensitivity parameter that determines the degree of differentiation between weight changes. For large β, only the weights attached to the maximally responding number field nodes will be adapted, whereas for smaller β, other weights will be adapted too. The factor $\exp[-\beta(y_{max} - y_j)]$ ensures an appropriate amount of differentiation between the changeability of number field nodes. In particular, for the most active node (which should be changed most), it equals 1, and for the least active node (which should be changed least), it is about 0.

After application of Equation 1, weights from the input field feeding into number field unit j are (approximately) normalized. In particular, a weight vector $\mathbf{w}$ feeding into a number field node is transformed as $\mathbf{w}/(\|\mathbf{w}\| + 0.001)$, with 0.001 added to avoid division by 0. The results to be presented in Simulations 1 and 2 are stable with respect to changes in parameter settings.

Our learning rule can be described as a Hebbian-like learning rule (see Wallis & Rolls, 1997). It also resembles the well-known Kohonen learning rule (Kohonen, 1995), with the exception that we assume no topographic ordering among the output nodes. In our algorithm, the amount of change of a nonwinning output node is not determined depending on its distance to the winning node, as is the case in a Kohonen rule, but rather depending on the activation value of the nonwinning node itself (similar to the neural gas algorithm; Martinetz & Schulten, 1991). The significance of this change is clarified later.

SIMULATION 1: THE DEVELOPMENT OF NUMBER-SELECTIVE NEURONS

In the first simulation, we investigated whether neuronal number-selectivity could be achieved spontaneously, and if so, whether it would be characterized by the same properties as those reported by Nieder and Miller (2003) and Nieder et al. (2002). For this purpose, only nonsymbolic summation-coded input was given. For 1000 trials, numerosities from 1 to 5 were randomly presented to the model. The sensitivity parameter β was set to 20. After learning, the model was tested by presenting each pattern (numerosities 1–5) again. The activation value of each number field node was then recorded for further analysis. Fifteen different replications with different random starting values were run, and we describe the mean values over these 15 replications. The trends to be described were obtained in individual replications as well.

Results and Discussion

The results clearly show that unsupervised learning led to the spontaneous development of nodes that are tuned to a specific numerosity. Importantly, these number-selective nodes exhibit the same properties as those described by Nieder and Miller (2003) and Nieder et al. (2002). We first report results on filtering, increasing bandwidth, and positive skewing, followed by Nieder and Miller's analyses to distinguish linear from logarithmic coding. Details of the statistical tests are provided in the Methods. In the text, we just report the p values of these tests.

Filter Property

Figure 3A shows the mean activation value of the number field nodes as a function of the distance to each node's preferred numerosity. The activation value for numerosities other than the preferred numerosity gradually diminishes as the distance increases ($p < .001$).

Why does the model show the filter property? Numerosities are represented as vectors in the 10-dimensional space of the summation field (because there are 10 summation field units), and vectors that code for similar numerosities are represented as vectors that are close together also. In particular, the distance between vectors increases monotonously as a function of the distance in numerosity (with or without normalization in the input field). Because the mapping from the summation to the number field is smooth (linear), this closeness will be preserved, and units that prefer a given numerosity will

also tend to respond relatively strongly to a nearby numerosity. From this property, the distance effect follows in a straightforward manner. Numerosities that are close to each other, lead to strongly overlapping distributions of activation, and therefore will be harder to discriminate than numerosities that are farther apart.

Increasing Bandwidth

The mean bandwidth equals 2.92. Figure 3B shows that bandwidth size is strongly correlated with preferred numerosity (see also Figure 3C). A nonparametric test shows that increasing bandwidth is a robust model property ($p < .001$). It follows from the fact that larger numbers are encoded by smaller activation values, because the summation field vectors are normalized. From this property, it follows that a "disagreement" between

Figure 3. (A–E) Results for nonsymbolic input. (F–J): Results for nonsymbolic input. (A, F) Normalized responses (as defined by Nieder et al., 2002) as a function of the distance between sample number and preferred number. (B, G) Bandwidth of units as a function of preferred number. (C, H) Normalized response as a function of sample number for units preferring a particular number. (D, I) Goodness-of-fit values (r^2) for the four fitted models (linear, power function $x^{1/2}$, power function $x^{1/3}$, and logarithmic). (E, J) Estimated standard deviations for the four models. Dashed lines denote fitted regression lines.
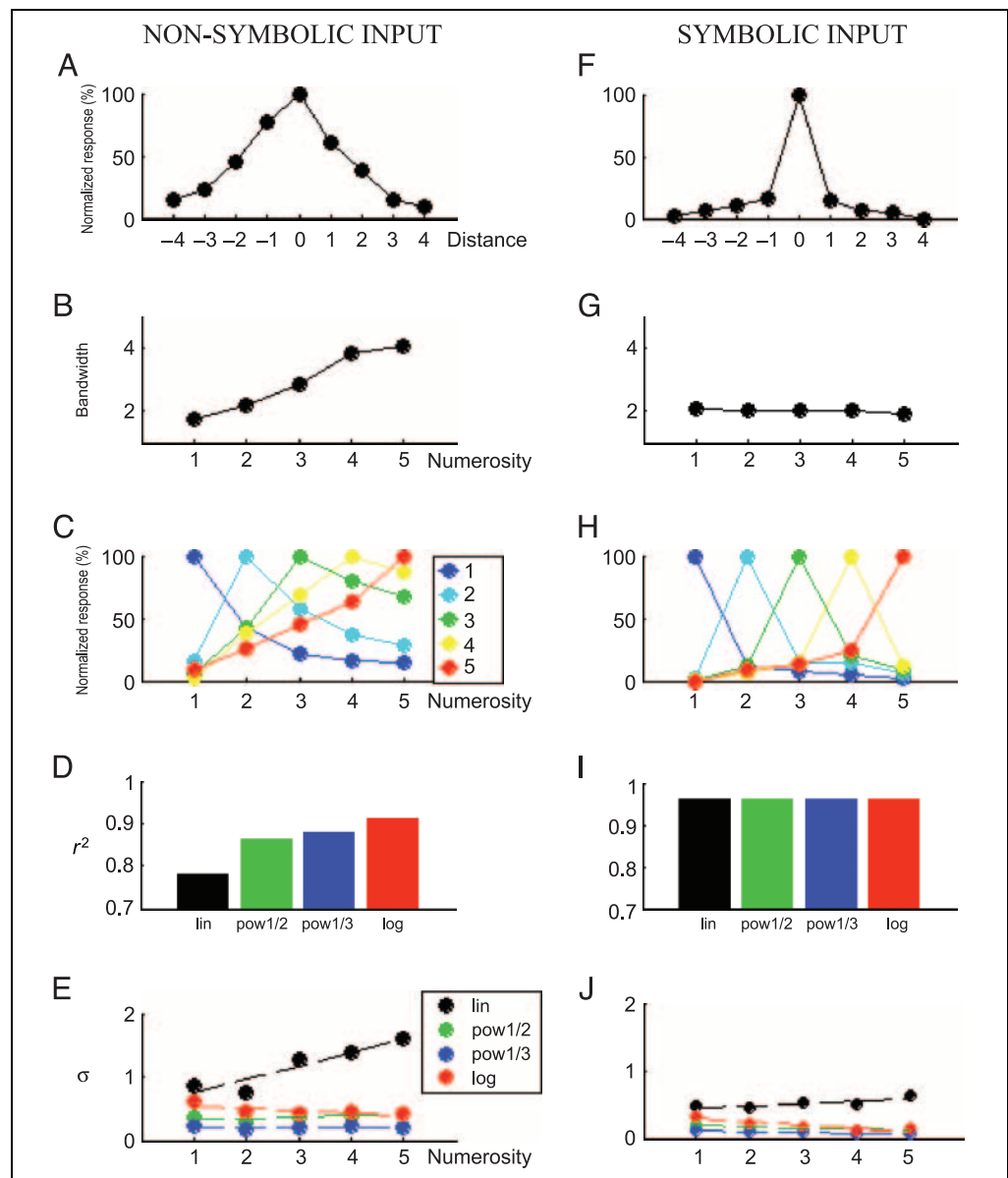
the input vector and the vector of weights feeding into a number field unit matters less for larger numerosities. Therefore, a response to near numbers will be relatively larger for larger numbers. This property accounts for the size effect, because larger numbers have larger bandwidths and will be more difficult to discriminate than smaller numbers.

Positive Skewing

Figure 3C also shows that positive skewing holds in our model nodes: For nodes that prefer quantity x, the response to quantity $x + 1$ was always stronger than the response to quantity $x - 1$ (and similarly for quantities $x + 2, x - 2$). The test for positive skewing yields $p < .001$.

Why is this property observed in the model? Let us assume that an input vector and number field unit weight vector mismatch in exactly two places. Then, the response of the number field unit will be stronger if the input has two nonzero values more than the unit than if the input has two nonzero values less. The result is that filtering curves will be positively skewed.

Linear or Logarithmic Scaling?

In their follow-up article, Nieder and Miller (2003) specifically contrasted the linear and logarithmic scaling hypotheses. We will now apply these authors' analyses to our own simulated data.

Gaussian functions were fitted on the tuning curves for each set of units that prefers a given numerosity separately. Four transformations of the numbers 1 to 5 were used: (1) linear coding, which amounts to no transformation; (2) transformation by raising each number by a power 1/2; (3) by a power 1/3; and (4) logarithmic transformation. Goodness-of-fit values (r^2) for the four models are shown in Figure 3D. Linear coding resulted in the worst fit, followed by the two power functions, and logarithmic compression yields the best fit, consistent with the findings on the neuronal data (see Nieder & Miller, 2003, their Figures 5 and 6). The three nonlinear coding schemes (power and logarithmic) fit reliably better than the linear one (Wilcoxon's signed rank test, $p < .001$).

The reason for the superiority of the nonlinear transformations is that they symmetrize positively skewed distributions, such as the distributions shown in Figure 3C. As a result, Gaussian functions, which are inherently symmetric, give a better fit to the data for the nonlinearly transformed models than the linear one, because the latter cannot fit the asymmetries in the filtering curves.

The estimated standard deviations of the Gaussian functions for the four models and the five numerosities are shown in Figure 3E. As can be seen, the standard deviations for linear coding tend to increase, but for the three other coding systems, the standard deviations are more or less constant, in line with the findings of Nieder

and Miller (2003). Besides increasing bandwidth, the fact that the standard deviations are increasing for the linear model ($p < .001$) is yet another indication that the filter functions show increasing variability with increasing numerosity. This is clearly predicted by Gallistel and Gelman's (1992) scalar variability assumption, but also by Dehaene's (2003) compressed coding assumption. However, the superior fit of the nonlinear coding schemes over the linear one favors compressed coding over scalar variability because scalar variability predicts symmetric filtering curves when no transformation occurs.

At this point, the model bears many resemblances to that developed by Dehaene and Changeux (1993). Like that model, the present model has a summation-coding scheme between visual input and numerosity detectors. Also like that model, numerosity detectors exhibit compressive scaling. Unlike their model, however, we have given an explicit rationale for summation coding (see backpropagation study) and trained the mappings from summation coding to numerosity rather than setting them by hand.

In sum, the properties of our number field nodes showed a striking similarity to the properties of the neurons described by Nieder and Miller (2003) and Nieder et al. (2002). Hence, we can conclude that unsupervised learning leads to the development of what has been called an internal "mental number line," which subjectively scales nonsymbolic numerical input in a logarithmic rather than a linear way. However, as noted by Dehaene (2001), the notion of a mental number line should not be taken too literally. Indeed, Nieder et al. observed that there was anatomically no topographic ordering among the number-sensitive neurons. Neither was such a topographic ordering observed in the model data, because we used a learning rule in which the distance between number field nodes is irrelevant.

To conclude, we have shown how, with a simple learning rule and relatively few trials, number-selective neurons with compressed scaling can emerge in an initially undedicated neural network. Nevertheless, as will be shown in the next section, one should be careful to extrapolate this scaling property to other input modalities as, for example, symbolically (e.g., verbally) presented numbers.

SIMULATION 2: LEARNING TO REPRESENT THE MEANING OF SYMBOLS

Whereas animals and young infants typically experience numbers only in the form of numerosities, children also learn to attach different symbols to different numerosities. For example, when four objects are presented, a child may learn to associate the auditory label "four" to this particular numerosity, and later also the written labels "four" and "4." Now that we have a model that captures the details of how neurons can represent numerical information when trained with nonsymbolic

input, this model provides an ideal opportunity to investigate how the exclusively human ability to represent numerical information by means of symbols may originate from these early available representations.

To simulate the developmental stage where children learn to associate symbols with numerosities, nonsymbolic and symbolic input were presented together to the model, starting from its state after Simulation 1. For example, on a particular trial, "four" could be presented both as a numerosity and as a symbol. Numerosities and symbols are represented in the summation field and symbolic field, respectively (see Figure 2). During symbol learning, the summation to number field connections obtained from Simulation 1 can either be preserved or allowed to change. In the absence of any strong a priori developmental arguments, we simulated both situations. Because the results were very similar, we only report simulation results with changeable summation to number field mappings.

Fifteen replications were performed. The initial weights of each of the replications were the final values obtained after the 15 replications of Simulation 1 (i.e., replication 1 in Simulation 2 started with the final values of replication 1 in Simulation 1, and so on). On each simulation, numerosities 1–5 were chosen at random for 1000 trials. The sensitivity parameter β was set to a lower value (=2) than in Simulation 1, because the extra symbolic input in Simulation 2 results in more differentiation between number field nodes. After learning, the model was again tested by presenting each numerosity and symbol, and the activation values of number field units were recorded.

Results and Discussion

Numerosity Input

After training, the number field units behaved like in Simulation 1 if, again, only numerosity input was presented (e.g., filtering, increasing bandwidth, positive skewing). Nevertheless, individual weights did undergo substantial change: The mean correlation (mean taken over the 15 replications) between preferred numerosity over number units before and after Simulation 2 was .124 ($p < .001$). The fact that this value is significantly different from 0 indicates that the weights from Simulation 1 set a bias for the weights from the summation field in Simulation 2. The fact that the correlation is relatively low, however, indicates that there was reorganization due to the symbolic input.

Symbolic Input

After learning, the number field nodes that responded maximally to a particular numerosity also preferred the corresponding symbolic code. In particular, the mean correlation between preferred numerosity and number

calculated over all number field nodes was equal to .986 ($p < .001$). Again, no topographic organization as a function of number preference was observed. Regarding the properties of number-selectivity, the following picture emerged. When stimulated with symbolic input only, the number field nodes still exhibited the filter property ($p < .001$), as can be seen in Figure 3F. However, the bandwidths are smaller with symbolic than with nonsymbolic material (see Figure 3G, mean bandwidth = 2.00; difference with nonsymbolic bandwidth, $p < .001$). This makes the tuning curves more "peaked" for symbolic input. This difference in peakedness is not due to the different values of the sensitivity parameter β that were used in Simulations 1 and 2. If the same parameter value had been used in both simulations, the difference would have been even more pronounced.

The fact that the filter property is preserved with symbolic input cannot be due to similarity in the input vectors because symbols were coded in an arbitrary manner, and so the input vectors in this field, coding for, say, number 1 and 2, have no greater similarity than the input vectors for numbers 1 and 4. The reason this property emerged nevertheless is the following. Suppose numerosity 3 (in the summation field) and the symbol for 3 (in the symbol field) are presented together. The activation of number field nodes coding for three will become activated, and hence, the connections between the active symbolic field node and the number field nodes will be adapted. However, due to the filter property in the mapping from the summation field to the number field, number field units coding for 2 and 4 will also be (weakly) activated. For this reason, the connection between the symbol field node coding number three, and number field nodes that code for two and four, will be adapted to some degree. Hence, the filter property emerges in the symbol to number field mapping as well. However, because the filtering property appears only indirectly with symbolic input, namely, as the result of a coupling with nonsymbolic numerosity input, the bandwidths are smaller. The discrepancy in bandwidths means that the system originally devoted to process nonsymbolic numerosity can represent the same kind of information with more precision when symbols are used. From this, the prediction follows that the distance effect observed in number comparison should be smaller for symbolic stimuli than with nonsymbolic stimuli, which is confirmed experimentally by Buckley and Gillman (1974). Notwithstanding this increased efficiency, symbolic representations still bear traces of the way that neural systems were originally shaped by nonsymbolic experience.

Contrary to the filter property, increasing bandwidth was not transferred from nonsymbolic to symbolic number processing (see Figure 3G). As can be seen from Figure 3H, the degree of positive skewing was very small compared with the skewing obtained with

nonsymbolic stimuli (difference $p < .01$) but it was systematic ($p < .001$). Possibly due to the low bandwidth values, it was not possible to empirically differentiate the different transformation models (linear, power, logarithmic), and all had high r^2 values (about .96, see Figure 3I). In addition, all standard deviations were about equal over different numerosities for the four models (see Figure 3J).

Because we observed no increasing bandwidth in Simulation 2 with symbolic input, it remains to be shown that a size effect is still possible with this model. Indeed, with symbolic input, a size effect is reliably obtained in numerical comparison, just as with numerosity input, although the size effect does appear to be weaker with symbolic input (Buckley & Gillman, 1974). We have demonstrated elsewhere (Verguts, Fias, & Stevens, in press) that the size effect observed with symbolic input in number comparison can also be localized in the mapping from the "mental number line" to output. In that article, a network was trained in a comparison task, and it presented symbolic numbers with frequencies proportional to their daily-life frequencies (Dehaene & Mehler, 1992). What we found was that the mapping from the mental number line to output nodes exhibited exactly the required nonlinear compressive pattern of connection weights that is necessary to generate a size effect. This is empirically supported by the fact that in priming experiments with tasks that do not rely on these mappings, but that do rely on the numerical representations (Reynvoet, Fias, & Brysbaert, 2002), distance-related priming effects were observed that were the same for primes smaller and larger than the target (consistent with absence of pronounced positive skewing). Moreover, the priming effects and absolute response times did not depend on number size (consistent with constant bandwidth over numbers; Figure 3G). Hence, the size effect may have different origins depending on the stimulus format: nonlinear mappings to output in case of symbols and number field coding in addition to nonlinear mappings in the case of nonsymbolic input.

GENERAL DISCUSSION

We have argued for the existence of a summation coding system between sensory input and numerosity detection. We described a model that uses such a system and showed how it can learn mappings from both summation and symbolic input fields to a common set of number detectors. The properties of this model were found to be consistent with both neurophysiological (Nieder & Miller, 2003; Nieder et al., 2002) and behavioral data (distance and size effect). We will now return to the three questions that were posed in the beginning of the article concerning the input to number detectors, the learnability of such detectors, and the relation between symbolic numerical

cognition and elementary numerical abilities that are shared with other animals.

Input to Number Detectors

Given that summation neurons act as a hidden layer, the question is then how such a system can be implemented. A set of neurons with the following three properties would be able to solve the task: (1) The neurons have large receptive fields. (2) They (spatially) summate responses of other (earlier) neurons with small receptive fields. The latter identify the presence of objects in specific locations (spatial coding). (3) The summation neurons have a differential sensitivity to input signals. To show that these three properties are sufficient, let us suppose we have a set of neurons that code for the presence of an object in a particular location. We will call them location neurons (see also Dehaene & Changeux, 1993). Location neurons may also be sensitive to a number of other dimensions, but for definiteness, we assume that they only respond to spatial location. Spatial coding can be done in various reference frames (e.g., eye, head, body; Cohen & Andersen, 2002). From Trick and Pylyshyn's (1994) theory that the ability to quickly enumerate a small number of elements in a visual display is attributable to the ability to preattentively tag a limited number of visuospatial positions for further attentive processing, we tentatively assume that location neurons code objects positions in an eye-centered reference frame. Location neurons signal the presence of an object in a small area only (Property 2), and summation neurons add the contributions of these earlier neurons (Properties 1 and 2). Therefore, each summation neuron works as an accumulator that has monotonously stronger or weaker input when a larger number of objects is presented. Due to Property 3, the numerosity in the display can be inferred from the population of summation neurons.

Although a system based on such properties has not yet been described neurophysiologically in the context of number detection, these properties have been reported to hold for a number of neurons, most notably, in parietal cortex. In particular, a subset of the neurons in the lateral intraparietal area (LIP) and parietal area 7a are plausible candidates for location and summation neurons, respectively. In the macaque monkey, neurons in area 7a have large receptive fields (up to 60°; Siegel & Read, 1997; Property 1). They receive input from area LIP (Siegel et al., 1997), which contains neurons with much smaller receptive field (Ben Hamed, Duhamel, Bremmer, & Graf, 2001, our purported location neurons; Property 2). Finally, neurons that are monotonously sensitive to an increase of a particular feature dimension have been reported in a number of cortical areas (e.g., in auditory association cortex, Edwards, Alder, & Rose, 2002; in areas V4 and MT, Cheng, Hasegawa, Saleem, & Tanaka, 1994; in somatosensory,

premotor, and prefrontal cortex, Romo, & Salinas, 2003), which verifies Property 3 at least in nonparietal areas. It remains an issue of investigation whether neurons with all the required properties can be located in parietal (or other) cortical areas. If verified, this would clarify why numerical operations have so often been associated with parietal cortex in imaging studies (e.g., Fias, Lammertyn, Reynvoet, Dupont, & Orban, 2003; Pinel, Dehaene, Riviere, & Lebihan, 2001). To sum up, then, summation coding provides both a natural and plausible intermediate stage for numerosity detection.

Learnability of Number Detectors

Dehaene (2002) has argued for the innateness of a number detection system on the grounds that it appears very early in life (e.g., number-selective neurons in a kitten of 8 days old, Thomson, Mayers, Robertson, and Patterson, 1970; number discrimination ability in infants, Wynn, 1992) and across a wide range of species. However, the fact that number-sensitive neurons can develop in a simple network architecture, and after only few learning trials, as demonstrated in this article, shows that the innateness assumption is premature. Each neuron whose functional properties are not fixed and that is connected to summation neurons can, in principle, develop number-selective properties. Note that summation neurons themselves can also be easily learned, as demonstrated in the backpropagation study. On the other hand, the high consistency in cerebral localization suggests that it is predetermined which parts of the brain become involved in quantity representation. Hence, although representational constraints are not necessarily required, there may be innate architectural constraints on which neurons come to function as number detectors (Elman, Bates, Johnson, Parisi, & Plunkett, 1999).

Relation with Symbolic Cognition

The use of language and other symbol systems is one of the distinguishing features between human and animal cognition. An open question is how symbol-processing systems are related to more primitive cognitive systems. One possibility is that representations and processes underlying symbol use are distinct from primitive mechanisms. Alternatively, symbolic cognition may be integrated in more primitive systems. In other words, uniquely human cognitive abilities may be rooted in rudimentary functions common to many species because of a general evolutionary advantage (Hauser, 2000). The current study focused on this issue from the perspective of number processing.

Dehaene (2001) has proposed the hypothesis that the use of numerical symbols is rooted in brain circuits devoted to presymbolic processing of numerosities by linking symbolic representational systems to a logarithmically scaled representation, which would then account for the size and distance effect in symbolic number comparison.

The present study agrees with this proposal in the sense that it shows that a neural system that was originally devoted to process nonsymbolically presented numerical information can learn to represent the meaning of arbitrary number symbols. In doing so, our model also shows how the representation of symbolic number meaning comes to inherit some of the properties of nonsymbolically construed representations. In particular, the key property that number-selective neurons act as a filter over numerosity (meaning that such a neuron responds maximally to its preferred numerosity and responds less and less to more distant numerosities) was transferred to symbol processing. This explains why human number processing is still subject to a distance effect, despite the fact that in principle the ability to use symbolic systems allows to uniquely define a given numerical value separate from other numerical values (4 is 4, not 5).

Furthermore, the present work argues that the acquisition of a symbolic system does not consist of merely linking symbols to the representations devoted to nonsymbolic processing, as demonstrated by the fact that the transfer of properties is not complete. Notwithstanding the fact that symbol use inherits the filter property, humans benefit from using symbols. We showed that the filter property is only partially inherited in the sense that the number-selective filters act in a more finely tuned manner when provided with symbolic input compared with nonsymbolic input. Consequently, the meaning of numerical symbols can be represented with more, but not absolute, precision. This is empirically supported by smaller distance effects when humans compare digits and number words as compared when they compare collections of dots (Buckley & Gilman, 1974).

In contrast to the filter property, our model exhibits no transfer of increasing bandwidth and only minimal transfer of positive skewing. Because these properties are a source of the size effect observed in number comparison, our model makes two predictions. First, no effects of number size are expected to emanate from the number representations themselves when tested with symbolic stimuli. The most clear testing conditions of this hypothesis are provided by number priming experiments (Koechlin, Naccache, Block, & Dehaene, 1999). A number of priming experiments shows distance-related priming, reflecting access to numerical representations. Importantly, the nature of this priming effect does not depend on number size (Verguts et al., in press; Reynvoet et al., 2002). Second, the size effect is expected to be smaller in symbolic than in nonsymbolic stimuli. This seems to be the case in the results reported by Buckley and Gilman (1974), but a systematic com-

parison of input formats is not yet available. In Verguts et al. (in press), we argue how the residual size effect in symbolic number comparison is attributable to mapping the numerical representations to a binary (smaller than, larger than) output mechanism for number comparison.

At the neural level, we predict that neurons of the type described by Nieder et al. (2002) exist in humans, that they are also used for symbolic input, but that the properties of these neurons will differ depending on the input format. Fias et al. (2003) already verified that in a comparison task, symbolic and nonsymbolic material may activate the same parietal regions. Recent studies provide evidence for a crucial involvement of parietal areas along the intraparietal sulcus in number processing (e.g., Pinel et al., 2001), with the fMRI blood oxygenation level-dependent (BOLD) signal exhibiting a distance effect. We hypothesize that these regions will also be activated by nonsymbolic material, and that the different nature of symbolic and nonsymbolic processing are expressed in differential BOLD signal signatures.

To conclude, this modeling study shows how symbolic cognition can be grounded in neural systems devoted to deriving abstract information from perceptual input. The same neurons represent numerical information with both symbolic and nonsymbolic stimulus formats, but differently for the two formats. This result sheds a new and specific light on our understanding of the systems involved in numerical processing.

METHODS

Statistical Tests

Here we describe the different tests that are used on the model data. For each test, the relevant reference distribution was generated by simulation, and the observed test statistic was compared with this distribution. In this way, the correct p value could be obtained.

Filtering Test

For each of the 15 replicated data sets in either Simulation 1 or 2, curves as in Figure 3C or 3H were generated. The number of violations of single-peakedness in the tuning curves was counted. For example, with the tuning curve for neurons preferring numerosity 3, if the response to numerosity 1 was larger than that to numerosity 2, this was a violation; if the response to 5 was larger than to 4, this was also a violation. The number of such violations over the 5 numerosities and the 15 data sets served as the test statistic. For each data set, the maximal number of violations is 12, so the test statistic could, in principle, range from 0 (*no violations*) to $12 \times 15 = 180$. The reference distribution was obtained by generating random numbers and checking

the number of violations from monotonicity in the same way as for the data.

Increasing Bandwidth

In the nonsymbolic case, bandwidths for each of the 15 data sets (as plotted in Figure 3B) were calculated, and the number of violations of monotonous increase over numerosity was counted. Numerosities 1 and 5 were not included in these calculations because the bandwidths for these numerosities were partly based on a linear extrapolation from the filter averages, in line with the procedure of Nieder et al. (2002). The reference distribution was obtained in the way as described in the previous paragraph. In the symbolic case, no statistics could be meaningfully calculated because the bandwidth was almost always equal to 2 (see Figure 3G).

Positive Skewing

Again, the curves in Figure 3C and H were used as a starting point, and the number of times that the response to numerosity $x - 1$ was smaller than to $x + 1$ was counted (x being the preferred numerosity). Hence, tuning curves for 2, 3, and 4 only could be used, and a sign test was applied (Siegel & Castellan, 1989).

Model Fits

The fit of the different models was compared by Wilcoxon's signed rank test (cf. Nieder & Miller, 2003).

Standard Deviations

To check the increase in standard deviations of the linear model, the procedure for increasing bandwidths described above was used. In this case, however, the standard deviations for numbers 1 and 5 were not partly arbitrarily defined, hence, these numbers were used here along with 2, 3, and 4.

Differences between Input Formats

Similar statistics were used to test the difference between the two input formats (symbolic and nonsymbolic), except that this time, the difference between the basic test statistics for the two formats was calculated and used as the relevant statistic.

REFERENCES

Ahmad, K., Casey, M., & Bale, T. (2002). Connectionist simulation of quantification skills. *Connection Science, 14,* 165–201.

Ben Hamed, S., Duhamel, J. R., Bremmer, F., & Graf, W. (2001). Representation of the visual field in the lateral intraparietal area of macaque monkeys: A quantitative receptive field analysis. *Experimental Brain Research, 140,* 127–144.

Buckley, P. B., & Gillman, C. B. (1974). Comparisons of digits and dot patterns. *Journal of Experimental Psychology, 103,* 1131–1136.

Cheng, K., Hasegawa, T., Saleem, K. S., & Tanaka, K. (1994). Comparison of neuronal selectivity for stimulus speed, length, and contrast in the prestriate visual cortical areas V4 and MT of the macaque monkey. *Journal of Neurophysiology, 71,* 2269–2280.

Chey, J., Grossberg, S., & Mingolla, E. (1998). Neural dynamics of motion processing and speed discrimination. *Vision Research, 38,* 2769–2786.

Cohen, Y. E., & Andersen, R. A. (2002). A common reference frame for movement plans in the posterior parietal cortex. *Nature Reviews Neuroscience, 3,* 553–562.

Dehaene, S. (2001). Is the number sense a patchwork? *Memory & Language, 16,* 89–100.

Dehaene, S. (2002). Single-neuron arithmetic. *Science, 297,* 1652–1653.

Dehaene, S. (2003). The neural basis of the Weber–Fechner law: A logarithmic mental number line. *Trends in Cognitive Sciences, 7,* 145–147.

Dehaene, S., & Changeux, J. P. (1993). Development of elementary numerical abilities—A neuronal model. *Journal of Cognitive Neuroscience, 5,* 390–407.

Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representation of numbers in the animal and human brain. *Trends in Neurosciences, 21,* 509.

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition, 43,* 1–29.

Edwards, C. J., Alder, T. B., & Rose, G. J. (2002). Auditory midbrain neurons that count. *Nature Neuroscience, 5,* 934–936.

Elman, J. L., Bates, E. A., Johnson, M. H., Parisi, D., & Plunkett, K. (1999). *Rethinking innateness: A connectionist perspective on development.* Cambridge: MIT Press.

Fias, W., Lammertyn, J., Reynvoet, B., Dupont, P., & Orban, G. A. (2003). Parietal representation of symbolic and nonsymbolic magnitude. *Journal of Cognitive Neuroscience, 15,* 47–56.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition, 44,* 43–74.

Hauser, M. D. (2000). *Wild minds: What animals really think.* New York: Henry Holt.

Koechlin, E., Naccache, L., Block, E., & Dehaene, S. (1999). Primed numbers: Exploring the modularity of numerical representations with masked and unmasked semantic priming. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1882–1905.

Kohonen, T. (1995). *Self-organizing maps.* Berlin: Springer-Verlag.

Martinetz, T. M., & Schulten, K. J. (1991). A ''neural gas'' algorithm learns topologies. In T. Kohonen, K. Mäkisara, O. Simula, & J. Kangas (Eds.), *Artificial neural networks* (pp. 397–402). Amsterdam: North-Holland.

Maunsell, J. H. R., & Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. *Journal of Neurophysiology, 49,* 1127–1147.

McNellis, M. G., & Blumstein, S. E. (2001). Self-organizing dynamics of lexical access in normals and aphasics. *Journal of Cognitive Neuroscience, 13,* 151–170.

Moyer, R. S., & Landauer, T. K. (1967, September 30). Time required for judgments of numerical inequality. *Nature, 215,* 1519–1520.

Nieder, A., Freedman, D. J., & Miller, E. K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. *Science, 297,* 1708–1711.

Nieder, A., & Miller, E. K. (2003). Coding of cognitive magnitude: Compressed scaling of numerical information in the primate prefrontal cortex. *Neuron, 37,* 149–157.

Orban, G. A., Kennedy, H., & Bullier, J. (1986). Velocity sensitivity and direction selectivity of neurons in areas V1 and V2 of the monkey: Influence of eccentricity. *Journal of Neurophysiology, 56,* 462–480.

Pinel, P., Dehaene, S., Riviere, D., & Lebihan, D. (2001). Modulation of parietal activation by semantic distance in a number comparison task. *Neuroimage, 14,* 1013–1026.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56–115.

Reynvoet, B., Brysbaert, M., & Fias, W. (2002). Semantic priming in number naming. *Quarterly Journal of Experimental Psychology, 55,* 1127–1139.

Romo, R., & Salinas, E. (2003). Flutter discrimination: Neural codes, perception, memory, and decision making. *Nature Reviews Neuroscience, 4,* 203–218.

Siegel, S., & Castellan, N. J. (1989). *Nonparametric Statistics for the Behavioral Sciences.* Singapore: McGraw-Hill.

Siegel, R. M., & Read, H. L. (1997). Construction and representation of visual space in the inferior parietal lobule. In J. Kaas, K. Rockland, & A. Peters (Eds.), *Cerebral cortex* (pp. 499–525). New York: Plenum.

Spelke, E. S. (2002). Developing knowledge of space: Core systems and new combinations. In A. M. Galaburda, S. M. Kosslyn, & Y. Christen (Eds.), *The languages of the brain* (pp. 239–258). Cambridge: Harvard University Press.

Thomson, R. F., Mayers, K. S., Robertson, R. T., & Patterson, C. J. (1970). Number coding in association cortex of the cat. *Science, 168,* 271–273.

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review, 101,* 80–102.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review, 108,* 550–592.

Verguts, T., Fias, W., & Stevens, M. (in press). A model of exact small number representations. *Psychonomic Bulletin & Review.*

Wallis, G., & Rolls, E. T. (1997). Invariant face and object recognition in the visual system. *Progress in Neurobiology, 51,* 167–194.