SKILL LEARNING AND REPETITION PRIMING IN SYMMETRY DETECTION: PARALLEL STUDIES OF HUMAN SUBJECTS AND CONNECTIONIST MODELS[1]

Neal J. Cohen

Irene T. Abrams

Walter S. Harley

Lisa Tabor

Department of Psychology
The Johns Hopkins University

Terrence J. Sejnowski

Department of Biophysics
The Johns Hopkins University

ABSTRACT

The present paper is a preliminary report of our work exploring skill learning and repetition priming in parallel studies of mirror symmetry detection in humans and network models. The memory mechanisms supporting the acquisition of skill and repetition priming in humans have been the subject of much speculation. On one account, drawing on the distinction between procedural and declarative learning, these learning phenomena grow out of experience-based tuning and reorganization of processing modules engaged by performance in a given domain, in a manner that is intimately tied to the operation of those modules. Such learning appears similar to that suggested by the incremental learning algorithms currently being explored in massively-parallel connectionist models (e.g., the Boltzmann machine). In the present work, both learning phenomena were observed in the behavioral data from human subjects and the simulation data from the network models. The network models showed priming effects from the start of de novo learning despite being designed to handle generalization to new materials - the essence of skill learning - and without additional mechanisms designed to provide a temporary advantage for recently presented material. Priming occurred for the human subjects despite the use of novel materials for which pre-existing representations cannot already be present in memory. These findings support the notion that skill learning and repetition priming are linked to basic incremental learning mechanisms that serve to configure and reorganize processing modules engaged by experience.

---

[1] Supported by a Biomedical Research Support Grant from the Division of Research Resources, NIH and a grant from the Sloan Foundation to NJC, and by grants from the National Science Foundation, System Development Foundation, General Electric Corporation, Exxon Education Foundation, Allied Corporation Foundation, Westinghouse, and Smith, Kline & French Laboratories to TJS. We thank Caroline McKeldin for research assistance and Valerie Mehl for manuscript preparation. Reprint requests should be sent to Neal J. Cohen, Department of Psychology, The Johns Hopkins University, Baltimore, MD 21218.

INTRODUCTION

The memory mechanisms supporting the ability to acquire new skills and the additional facilitation of performance due to repetition priming effects (i.e., superiority in task performance for repeated materials) have been the subject of much recent work in cognitive science and cognitive neuropsychology. Skill learning and repetition priming are both exhibited as a consequence of experience in amnesic patients despite their impaired recall and recognition of the specific materials used to train and test skilled performance, poor recollection of the learning experiences during which skilled performance was developed, and poor insight into the nature of the knowledge underlying the increasingly skilled performance (see Cohen, 1985; Squire & Butters, 1984 for reviews). Skill learning and repetition priming have been dissociated from aspects of explicit remembering (recall and recognition) in normal subjects as well (see Kolers, 1979; Schacter, 1985; Tulving, 1985 for reviews). We have argued that these learning phenomena reflect the operation of a procedural memory system that influences the organization of perceptual and action systems in a way that does not depend upon explicit representation of particular learning experiences or of rules about the world (Cohen, 1985; Cohen, Eichenbaum, DeAcedo & Corkin, 1985; Squire & Cohen, 1984). On this view, skilled performance in a given domain grows out of the tuning and reorganization of processing and action modules engaged by performance in that domain; the learning that results consequent to experience in that domain is intimately tied to the operation of its processing components.

The characteristics of procedural learning and its relationship to the declarative system have been explored most extensively by John Anderson, who has modeled the procedural system in a production-system framework in the context of a broader conceptualization of memory and cognition (e.g., Anderson, 1982, 1983). Our view of the procedural system, however, and the explanation of skill learning and repetition priming that it suggests, seems to bear strong similarity to that of the incremental learning algorithms currently being explored in massively-parallel network models.

One example of network architectures is the Boltzmann machine, which has been applied successfully to such problems as figure-ground separation in visual perception (Sejnowski & Hinton, 1986). The processing units in this architecture are binary and connected symmetrically with coupling strengths or weights that can have positive or negative values. There exists a learning algorithm for the Boltzmann machine that allows the network to automatically find a set of weights for solving a problem given only examples of typical inputs and the required outputs (Hinton & Sejnowski, 1983; Ackley, Hinton & Sejnowski, 1985). The Boltzmann machine learning algorithm has recently been used for learning to detect the axis of mirror symmetry in checkerboard-based patterns (Sejnowski, Kienker & Hinton, 1986). This problem is a second-order predicate in the sense of Minsky and Papert (1969) and is beyond the capability of the perceptron learning algorithm (Rosenblatt, 1959) and Hopfield networks (Hopfield, 1982). The crucial difference that distinguishes Boltzmann machines from perceptrons and allows them to solve difficult problems is the presence of additional units between the input and output layers, called hidden units. These hidden units can be used as feature detectors for solving the problem; the learning algorithm discovers the optimal set of feature detectors by shaping the weights among units through incremental changes.

In nearly all previous work with connectionist models, the focus has been on the ability of a network to generalize from examples to new instances on which the network was not previously trained[2]. The work with human subjects suggests that skill learning, measured in terms of the improvement in performance for novel materials in the trained domain, is closely associated with repetition priming effects, the additional facilitation in performance specific to the materials actually presented during training. In order to examine whether these learning phenomena are linked in the network models as they are in humans, we have been exploring skill learning and repetition priming in both Boltzmann machine architectures and human subjects in studies of mirror symmetry detection for checkerboard-based patterns. It is important to note that this enterprise is not in any way intended to offer the Boltzmann machine implementation of symmetry-detection learning as a detailed model of how humans actually acquire skill in this domain. Rather it is intended to

---

[2] Although McClelland & Rumelhart (1986) and Carpenter and Grossberg et al. (1985) have considered repetition priming effects in models of word recognition.

explore whether simple, incremental learning algorithms can manifest both stimulus-generalizable skill learning and stimulus-specific repetition priming phenomena in the same networks.

Perception of mirror symmetries is superb in humans (e.g., Kohler, 1929; Garner, 1962; Bruce & Morgan, 1975; Barlow & Reeves, 1979), and, indeed, has been assumed to play an important role in the early stages of visual processing (e.g., Marr & Nishihara, 1978; Biederman, 1985). It is not surprising, then, that studies of mirror symmetry detection in humans have focused on its psychophysical characteristics rather than on its improvement with practice. The acquisition of visual symmetry recognition has been studied in pigeons (Delius & Nowak, 1982) and Boltzmann machines (Sejnowski, Kienker & Hinton, 1986), but this study is, to our knowledge, the first to examine the acquisition of mirror symmetry detection in humans. We assume that despite their already developed skill in symmetry detection, humans will show considerable practice effects, learning how best to apply symmetry-detection routines to this particular class of visual materials, e.g., learning about the critical visual features that are diagnostic of mirror symmetry for different axes in these particular checkerboard-based patterns.

An important question that has been raised in previous work on repetition priming in humans, and especially in work with human amnesic patients, is the extent to which priming depends upon a structured, pre-existing knowledge base in which particular facts can be temporarily activated by recent experience, rather than upon acquisition of new knowledge (e.g., see Fowler, Napps & Feldman, 1985; Schacter, 1985; Gordon, 1986). The present work addresses this issue in two ways: First, the checkerboard-based patterns used as stimuli in the present studies are novel materials, and are not a part of subjects' pre-existing knowledge structures. That is, any priming effects obtained here could not be accounted for by postulating the temporary activation of already stored logogen- or pictogen-like representations. Second, the computer simulation work explores the possibility of priming effects in networks learning de novo, where no pre-existing knowledge is provided about topography, symmetry, or checkerboard-based patterns. The present paper is a preliminary report of findings from our parallel studies of human subjects and computer simulations relevant to understanding the mechanisms of repetition priming and its relationship to skill learning.

METHOD

Subjects

Twenty-two students at The Johns Hopkins University volunteered to be subjects in the present studies. Of these, 12 served in the priming experiment and 10 served in the axis generalization experiment.

Materials

Stimuli were computer-generated, mirror-symmetric visual patterns constructed by designating as purple approximately 40% (range = 31-49%) of 81 cells within a 9x9 blue display. Each pattern was symmetrical about one and only one of four axes: horizontal, vertical, left diagonal, or right diagonal. The precise distribution of purple and blue cells within a pattern, given a particular axis of symmetry, was determined randomly by an IBM PC. The patterns were approximately 13 cm on each side and subtended approximately 8 deg of visual angle.

Apparatus

Patterns were presented on an IBM PC color display controlled by an IBM PC. Reaction times and response axes were recorded via a Summagraphics optical mouse.

General Procedure

Subjects were seated individually in front of an IBM PC color display controlled by an IBM PC, and held a Summagraphics optical mouse in their dominant hand. They were presented with a series of 400 patterns (divided evenly among the 4 axes) organized into blocks of 60-80. Within each block there was an equal number of patterns representing each axis, with no axis occurring more than 3 times in succession. Subjects initiated each trial by pressing the left button of the mouse. Each trial consisted of a visual pattern presented for 83 msec followed, after a 17 msec unfilled delay, by a visual mask. The mask consisted of a regular, alternating-color checkerboard whose cells were identical in size and color to the test stimuli. It remained on the screen until a response was recorded, or until 8 sec had elapsed, ending the trial. Subjects were instructed to indicate the axis of symmetry by moving the mouse on the table-mounted pad in the direction analogous to the axis perceived on the display. They were instructed to respond as quickly and as accurately as possible, guessing whenever in doubt. After each response, feedback was given in the form of a re-presentation of the pattern with the addition of a line indicating the correct axis. The feedback pattern remained on the screen for 2 sec.

Prior to the start of the test series, subjects received a practice session consisting of 20 stimuli (divided equally among the axes presented in the subsequent experimental condition), each presented with a feedback line indicating the correct axis.

Exp. 1: Priming experiment

Subjects were trained on a combination of repeating (primed) and nonrepeating patterns. Of the 400 stimuli presented to each subject, there were 240 nonrepeating patterns (60 of each axis) common to all subjects, with each subject receiving a different, pseudo-randomized order. The remaining 160 stimuli were the primes, consisting of a set of patterns that repeated at one of three different rates: In the 1/10 condition, 4 patterns (1 of each axis) repeated in every 10+2 stimuli; in the 1/20 condition, 8 patterns (2 of each axis) repeated in every 20+6 stimuli; in the 1/40 condition, 16 patterns (4 of each axis) repeated in every 40+10 stimuli. Note that the overall percentage of prime stimuli in each series was maintained at 40% across the different conditions by doubling the number of prime patterns for each halving of the prime repetition rate. The prime patterns were yoked across the three priming-rate conditions such that a given prime pattern appeared in each condition. Stimuli were presented in 5 blocks of 80.

Exp. 2: Axis generalization experiment

This experiment consisted of a training phase, in which patterns were symmetric about one of only two axes, and a testing phase, in which patterns were symmetric about one of four axes, as in the priming experiment. The training phase consisted of 100 stimuli symmetric about either the left and right diagonal axes, for one group of subjects, or the horizontal-vertical axes, for another group. These stimuli, presented in 2 blocks of 50, were different for each subject. The testing phase consisted of 300 stimuli (75 of each axis) common to all subjects, with each subject receiving a different, pseudo-random order. The test stimuli were presented in 5 blocks of 60.

Computer Simulations

Networks

Two networks were studied. Both had 81 input units, representing the 81 cells of the 9x9 array, and 4 output units, representing the 4 possible axes of mirror symmetry. They differed with respect to the number and connectivity of the hidden units. One network had 12 hidden units, each fully and symmetrically connected to all of the input units and output units (hereafter called the global-unit network). The other network had 72 hidden units, each connected to the 4 output units and to one of 9 3x3 sections of the 9x9 stimulus array (hereafter called the local-unit network). For both networks, the hidden units were not interconnected.

General procedure

The learning algorithm and implementation details are the same as those described in Sejnowski, Kienker and Hinton (1986). All networks started with all weights set to zero except for the axis generalization experiment, in which the training phase ran until performance reached asymptote and then switched to the testing phase until performance again reached asymptote. Approximately 40,000-100,000 patterns were presented in each simulation.

Materials

Patterns were randomly generated following the same constraints as those used in the human experiments. Patterns were composed of approximately 40% (range = 31-49%) of the 81 cells in a 9x9 array being "on" in such a way as to be symmetric about one and only one of the four axes.

Exp. 1: Priming experiment

The networks were trained on a combination of repeating (primed) and nonrepeating patterns. In each simulation, two priming rates were included, with one being a multiple of 2 or 3 that of the other. In each case (with the exception of the condition with the highest priming rate; see below), there were at least 20 different patterns that served as prime stimuli for each priming rate, divided equally among the different axes. There were six different conditions, involving priming rates of 1/50 and 1/150, 1/100 and 1/200, 1/200 and 1/400, 1/400 and 1/800, 1/800 and 1/1600, and 1/833 and 1/2500.

Five simulation conditions were run with the global-unit network with several replications. For three of these, the overall percentage of prime stimuli was held constant at 30% (20 different patterns each at 1/100 and 1/200; 80 different patterns each at 1/400 and 1/800; and 160 different patterns each at 1/800 and 1/1600). For the other two simulations, the overall percentage of prime stimuli was permitted to vary, with 9 and 20 prime patterns respectively at 1/50 and 1/150 (31% prime stimuli), and 20 prime patterns each at 1/833 and 1/2500 (3.2% prime stimuli).

Five simulation conditions were run with the local-unit network with several replications. For two of these, the overall percentage of prime stimuli was held constant at 30% (40 different patterns each at 1/200 and 1/400; and 80 different patterns each at 1/400 and 1/800). For the other three simulations, the overall percentage of prime stimuli was permitted to vary, with 20 different patterns each at 1/100 and 1/200 (30% prime stimuli), at 1/200 and 1/400 (15% prime stimuli), and 1/400 and 1/800 (7.5% stimuli).

Exp. 2: Axis generalization experiment

In the training phase, the networks were presented with patterns symmetric about one of only two axes, either the left and right diagonal axes, for one set of simulations, or the horizontal-vertical axes, for another set of simulations. In the testing phase, the networks were presented with patterns symmetric about one of the four possible axes, as in the priming experiment.

RESULTS

Human Experiments

Exp. 1: Priming experiment

Subjects began training with performance well above chance levels and showed steady skill learning across the 5 blocks of the experiment. With practice, responses to the nonrepeated patterns became increasingly accurate

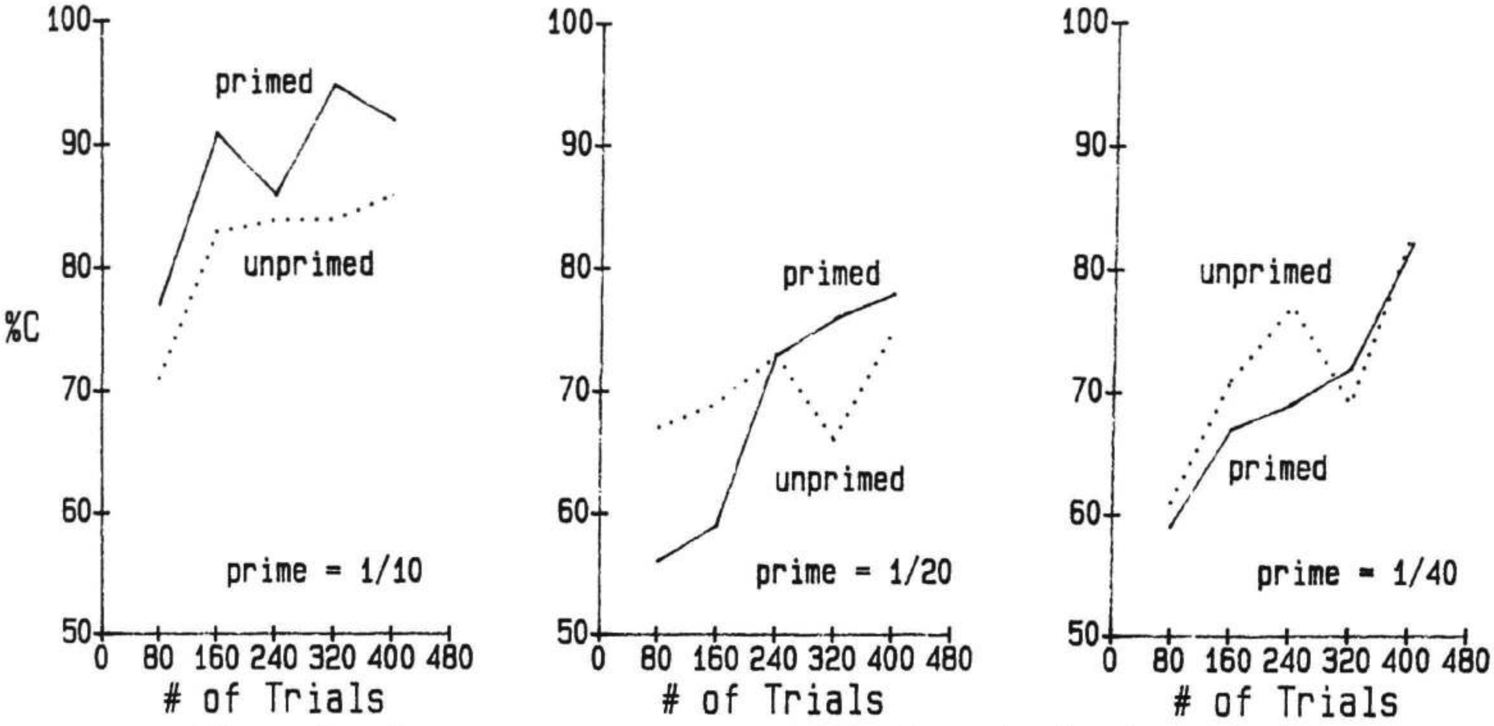

Figure 1. Percent correct across trials for primed and unprimed patterns. Across priming conditions (1/10, 1/20, and 1/40), the overall percentage of prime stimuli was held constant.

(see Figure 1) and had shorter latencies (see Figure 2). Subjects also showed clear repetition priming effects, measured both in terms of percent correct (Figure 1) and reaction time (Figure 2). Repetition priming was evident for all blocks for the 1/10 repetition rate and seemed to emerge by the last two blocks for the 1/20 repetition rate; no priming was observed for the 1/40 repetition rate.

One interesting result concerns performance for the different axes when analyzed separately. Performance was superior for the horizontal and vertical axes, for which subjects presumably have had a good deal more real-world experience prior to the present studies, than for the diagonal axes, which seem to play less of a role in our perceptual experience. The observed difference in performance was maintained across all 5 blocks of the experiment (see Figure 3).

Exp. 2: Axis generalization experiment

Results for the 2-axis training and 4-axis testing phases of this experiment are presented in Figure 4 (only the first 3 of the 5 testing blocks are shown). Data are presented separately for training on horizontal-vertical

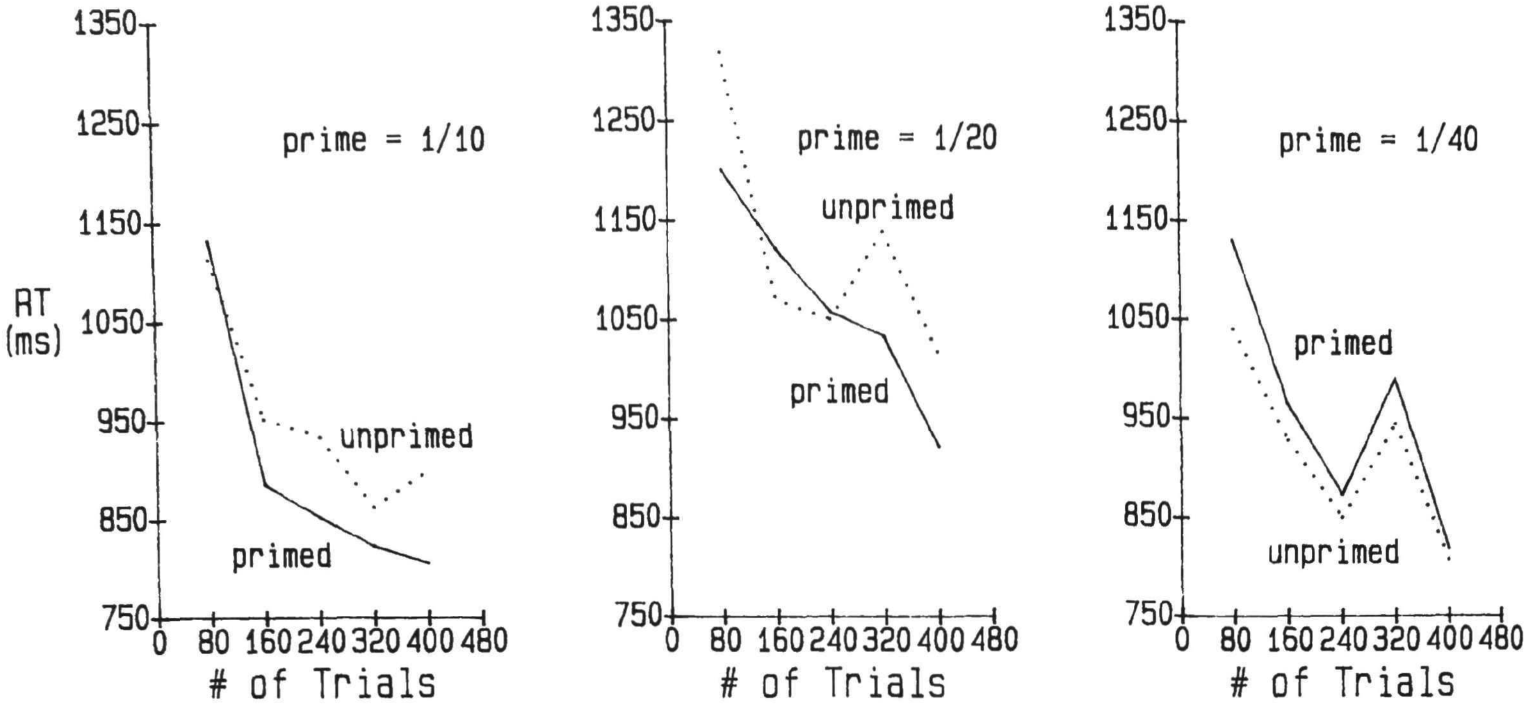

Figure 2. Reaction times (in msec) across trials for primed and unprimed patterns. Across priming conditions (1/10, 1/20, and 1/40), the overall percentage of prime stimuli was held constant.

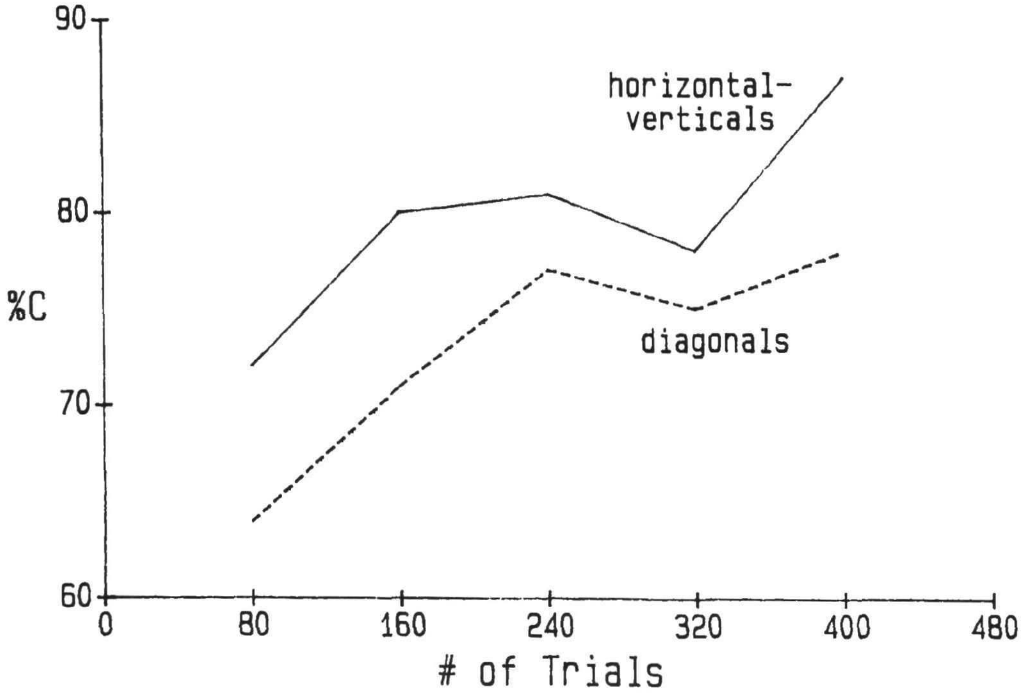

Figure 3. Percent correct across trials for horizontal-vertical or diagonal unprimed boards. Data are averaged across the different priming conditions.

axes versus diagonal axes (middle and bottom panels), and are also averaged together as trained versus untrained axes (top panel). The skill learning exhibited by subjects for new patterns did not generalize across axes. After two blocks of training with two axes, performance on the untrained axes at the start of the 4-axis testing phase was no better than - indeed, on average was poorer than - performance at the beginning of the training phase. Note, however, that these data do not take into account differences in number of response alternatives between the training and testing phases. To the extent that the observed performance includes some amount of guessing distributed

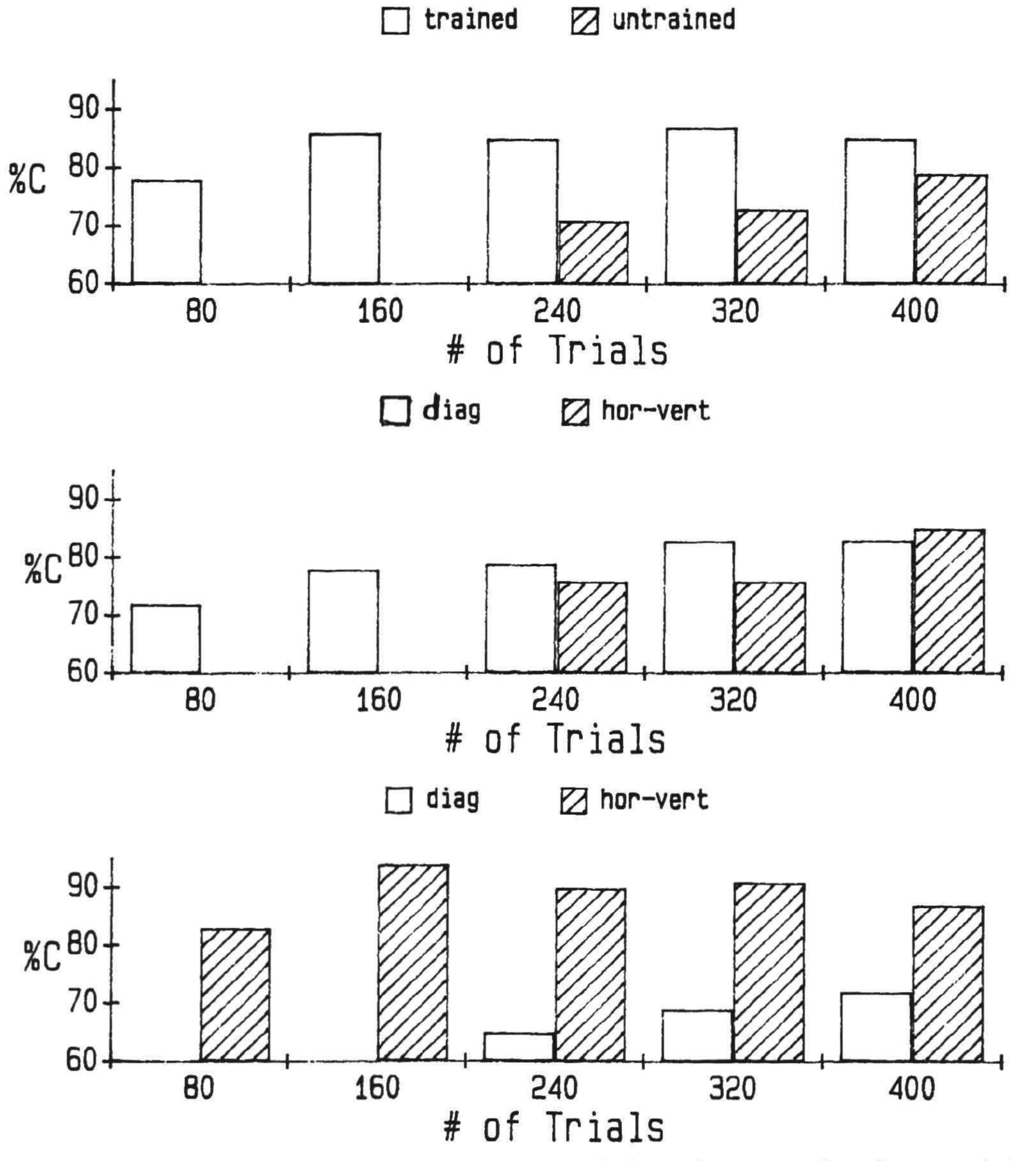

Figure 4. Percent correct across trials when trained on either diagonal (middle) or horizontal-vertical axes (bottom) and tested on all four axes. Overall effect of training is shown at top.

among the set of response alternatives, the contribution of guessing to performance would be different for training than for testing, artificially lowering the scores on the testing phase. Our inspection of the errors and observations of subjects taking the test suggests that rather little contribution is made to the observed performance by distributing guesses among the set of response alternatives, thereby lessening its presumed impact.

The clearest finding of this experiment was the difference between performance for the diagonal axes and performance for the horizontal-vertical axes, amplifying the effect seen in Exp. 1. First, looking only at the training results, subjects' initial performance was better for the horizontal-vertical axes (bottom panel) than for the diagonal axes (middle panel), and there was more improvement over the two training blocks for the horizontal-vertical axes than for the diagonal axes. Moreover, the effects of 2-axis training on 4-axis test performance was dramatically different for the two sets of axes: Training on the horizontal-vertical axes produced a huge advantage in performance for these axes over the diagonal axes when tested in the 4-axis condition (bottom panel), whereas training on the diagonal axes served only to boost performance on these axes to the level of the horizontal-vertical axes when tested in the 4-axis condition (middle panel).

Computer simulations

Exp. 1: Priming experiment

For each network, simulations showed both skill learning for nonrepeated items and repetition priming for repeated items, even over very long lags (see Figures 5-8). The finding of priming in the global-unit network for the 1/1600 condition (see Figure 6) deserves particular emphasis. The priming effect was apparent with fewer than 12 presentations of a given prime pattern spread out over 20,000 stimuli. In fact, in a subsequent study (Cohen, Abrams, Harley, Tabor, Gordon, & Sejnowski, 1986), we have demonstrated priming with as few as 6 presentations of a given prime stimulus spread out among 20,000 patterns. In addition, the priming effect was remarkably reproducible across replications. That is, for those sets of simulations that included overlap among the priming rates (e.g., for the local-unit network, the three simulations had priming rates of 1/100 and 1/200, 1/200 and 1/400, and 1/400 and 1/800; look across panels of Figure 7), the performance of

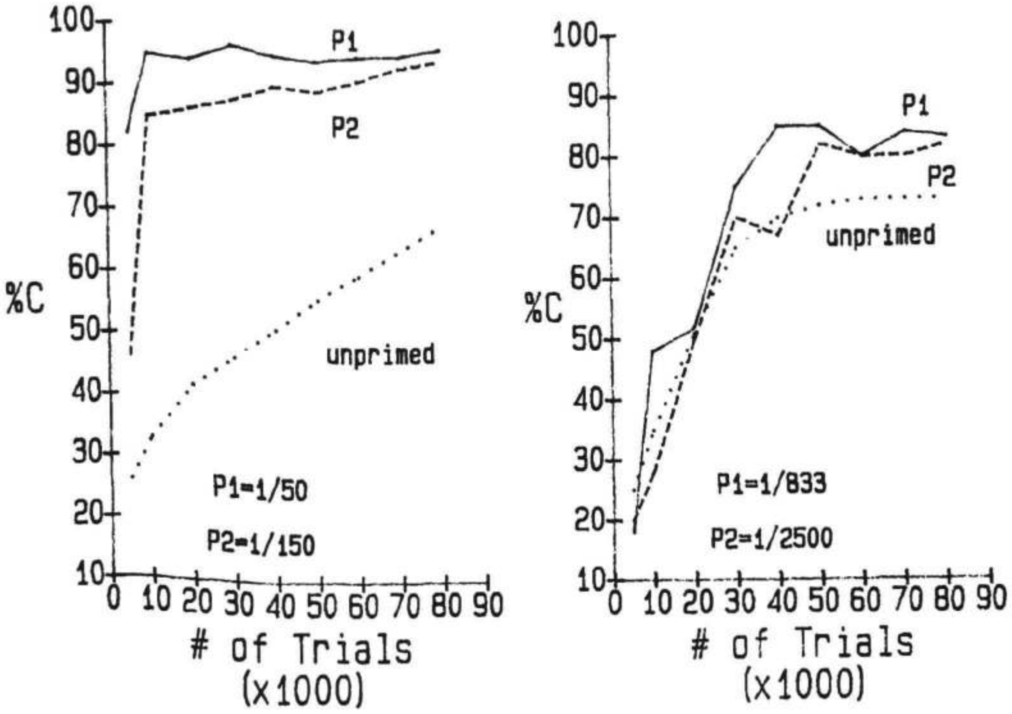

Figure 5. Percent correct for global-unit network for primed and unprimed stimuli. Each of two simulations tested performance on two different priming rates. The overall percentage of prime stimuli varied across simulations.

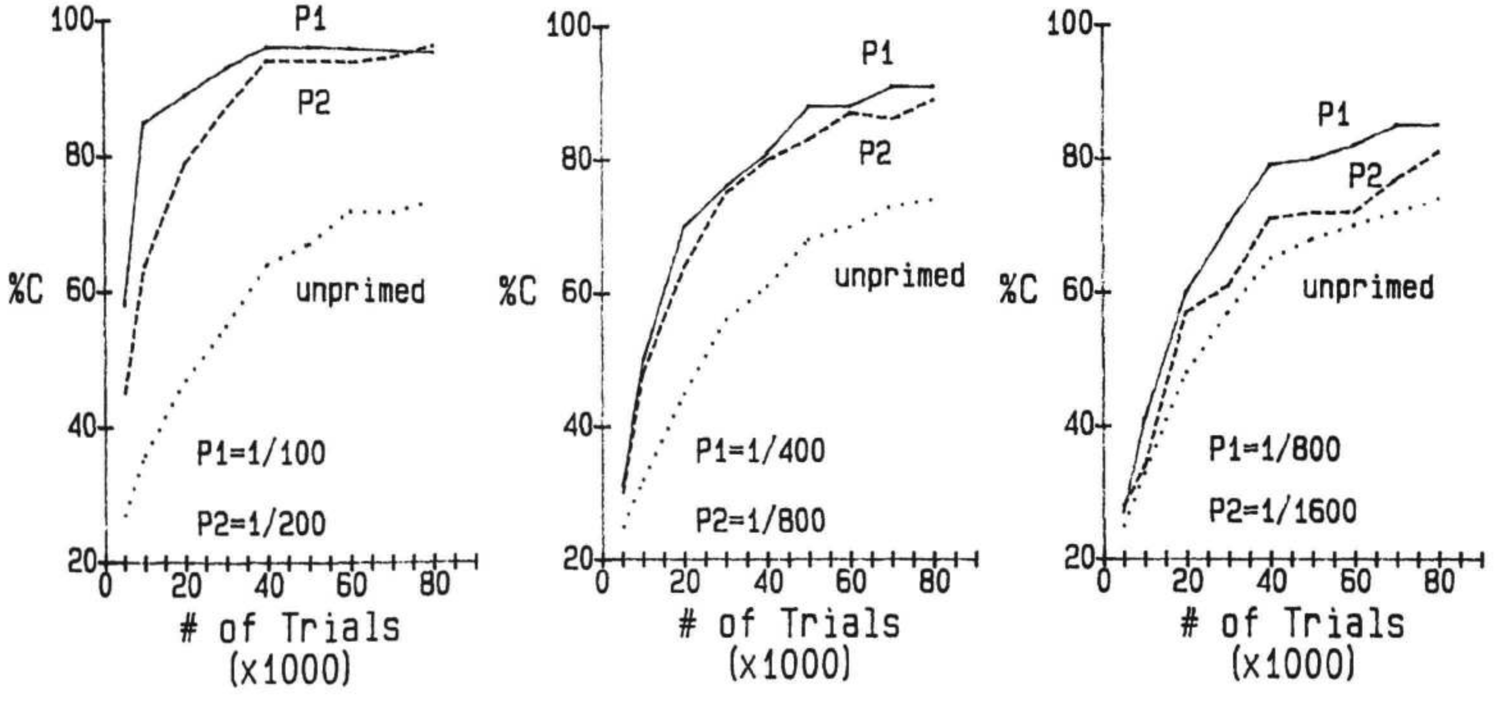

Figure 6. Percent correct for global-unit network for primed and unprimed stimuli. Each of three simulations tested performance on two different priming rates. The overall percentage of prime stimuli was maintained across simulations.

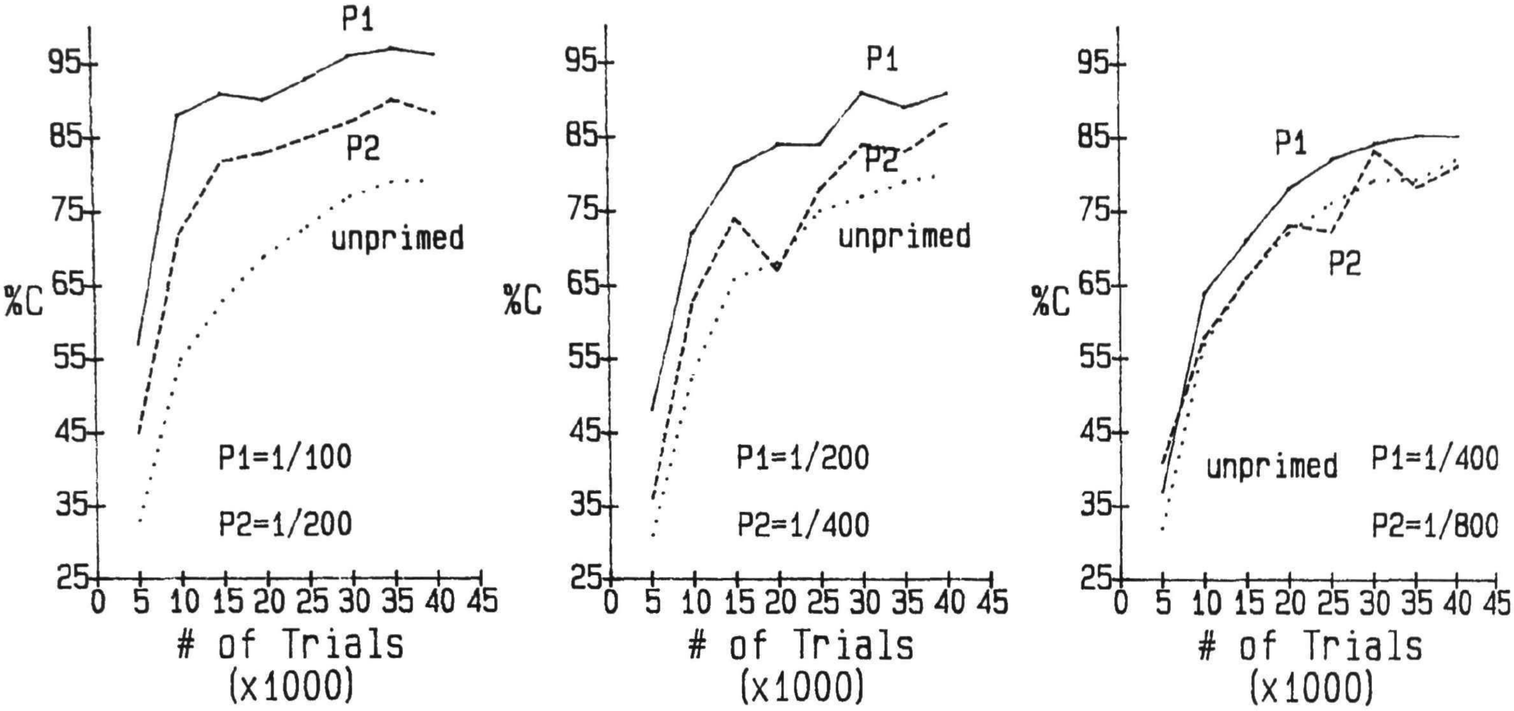

Figure 7. **Percent correct for local-unit network for primed and unprimed stimuli. Each of three simulations tested performance on two different priming rates. The overall percentage of prime stimuli was permitted to vary conditions.**

network for the same priming rate on different occasions was virtually identical. The performance of both networks for the different axes was nearly invariant, unlike the results for human subjects.

The two networks differed in a number of interesting ways. First, though both networks started learning de novo, the local-unit network learned much more rapidly, reaching a higher level of performance for nonrepeated stimuli in 40,000 trials (see Figures 7 & 8) than did the global-unit network in 80,000 trials (see Figures 5 & 6). Second, despite the superior skill learning of the local-unit network, the global-unit network was the more sensitive one to priming, showing a priming effect even at a priming rate of 1/1600 (see Fig. 6) and at 1/3200 (Cohen et al., 1986), whereas the local-unit network showed only marginal priming at a priming rate of 1/800 (see Figures 7 & 8). Third, the global-unit network but not the local-unit network showed a trade-off between skill learning and repetition priming as a function of priming rate when the overall percentage of prime stimuli was permitted to vary. Thus, performance for the prime stimuli was better in the 1/50 and 1/150 condition than in the 1/833 and 1/2500 condition, whereas performance for the nonrepeated stimuli was poorer in the 1/50 and 1/150 condition than in

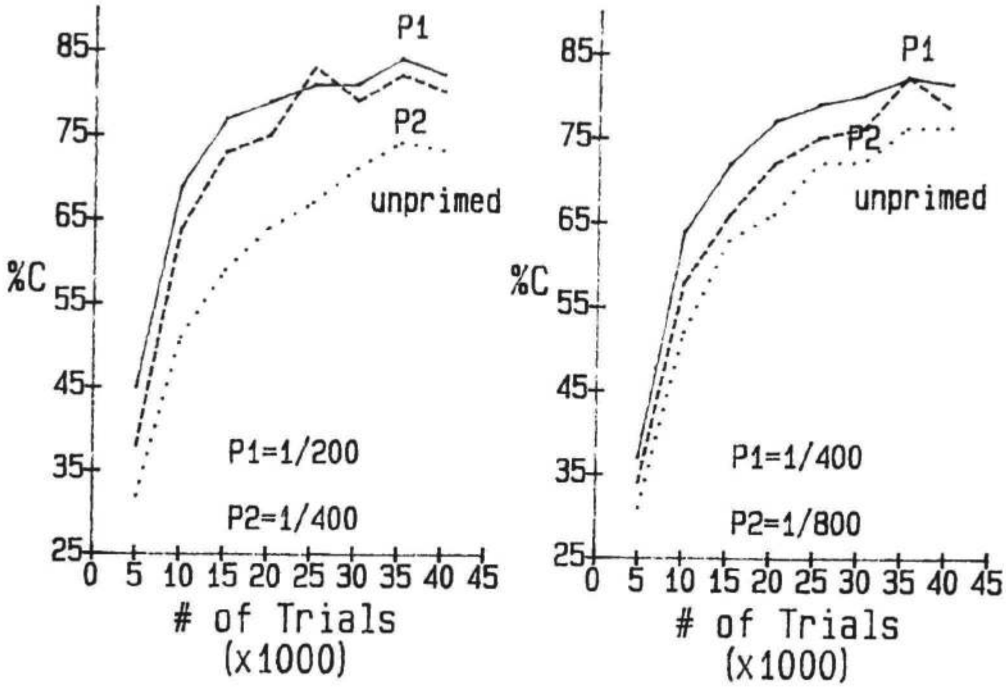

Figure 8. Percent correct for local-unit network for primed and unprimed stimuli. Each of two simulations tested performance on two different priming rates. The overall percentage of prime stimuli was maintained across simulations.

the 1/833 and 1/2500 condition (see Figure 5). The trade-off disappeared, however, when the overall percentage of prime stimuli was held constant across the different priming rate conditions (see Figure 6). For the local-unit network, there was no trade-off regardless of whether the overall percentage of prime stimuli varied or was constant (see Figures 7 & 8).

Inspection of the weights for the hidden units of the two networks was illuminating. The hidden units in the global-unit network behaved in the manner reported by Sejnowski et al. (1986). The weights were frequently antisymmetric about one or more axes, and often were also symmetric about one or more other axes. The amount of antisymmetry was striking. The spatial distribution of the weights to input units corresponding to different portions of the receptive field varied considerably among hidden units and represented a number of different types of geometric features as well as some isolated cells in the array. Many of the geometrical features, such as linear stripes or angles, were quite global, spanning the entire width or length of the array. Frequently, such features were represented in one hidden unit along with their complement in a different unit. Finally, the weights to some

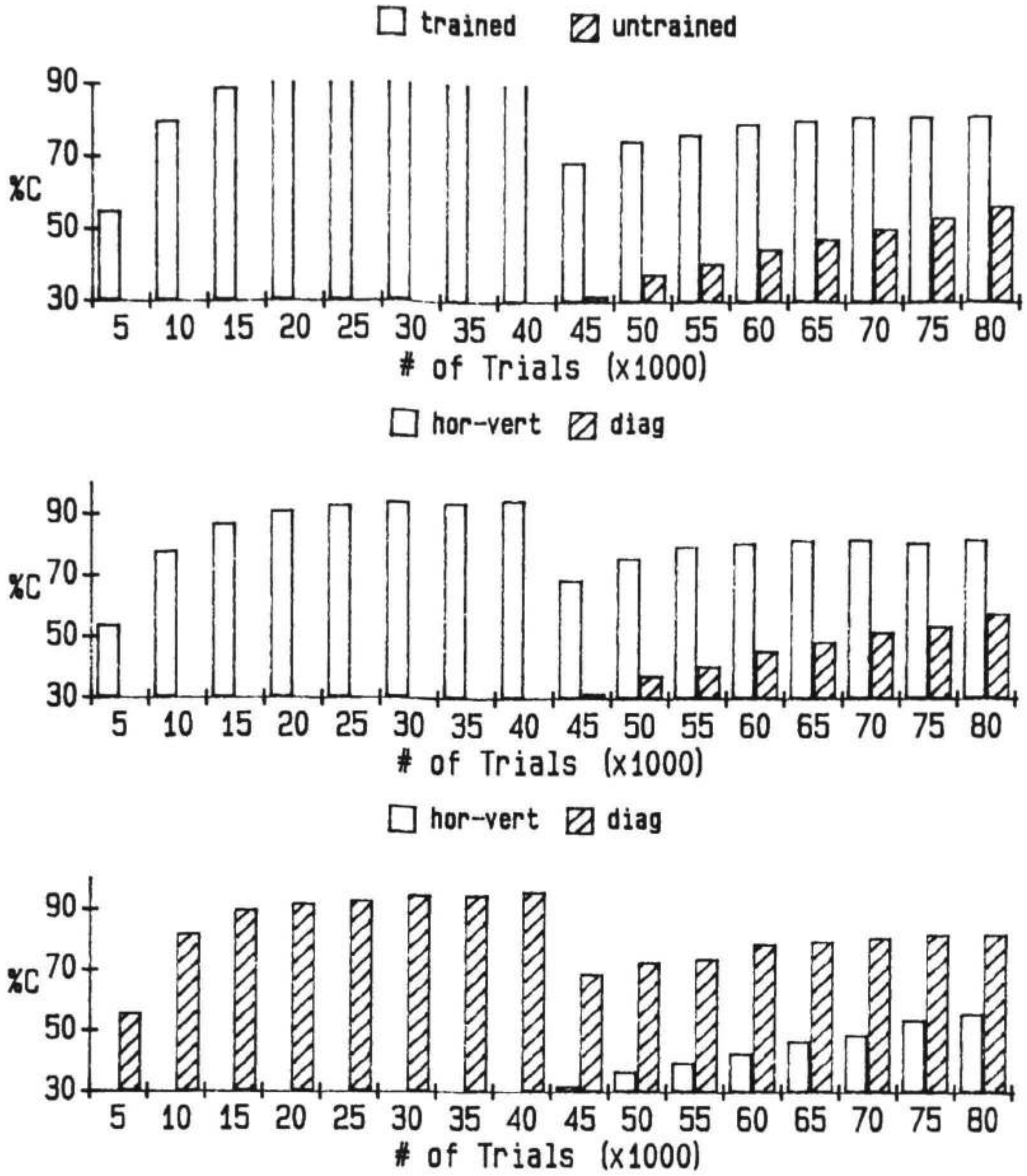

Figure 9. Percent correct for global-unit network when trained on either horizontal-vertical (middle) or diagonal axes (bottom) and tested on all four axes. Overall training effect is shown at top.

hidden units were very small, suggesting that these units played little or no role in the solution to the problem.

The hidden units in the local-unit network behaved in a similar fashion, although their receptive fields were quite local, being restricted to particular 3x3 sections of the 9x9 input array. Here, too, the number of units whose weights were set up to detect antisymmetry was striking. The presence of hidden units with weights representing geometrical features and their complement in a unit with a matching receptive field was noted here as well. An interesting aspect of the hidden units in this network was the apparent reliance of the network on units with receptive fields in the center 3x3 section of the input array. This section is relevant to detection of any of the four axes. Accordingly, the weights for such units were typically sensitive to multiple axes and, in particular, were antisymmetric for one or more axes while being symmetric for one or more others.

Exp. 2: Axis generalization experiment

Both networks trained on a particular pair of axes showed a decrement in overall performance when switched to testing with all 4 axes (see Figures 9 &

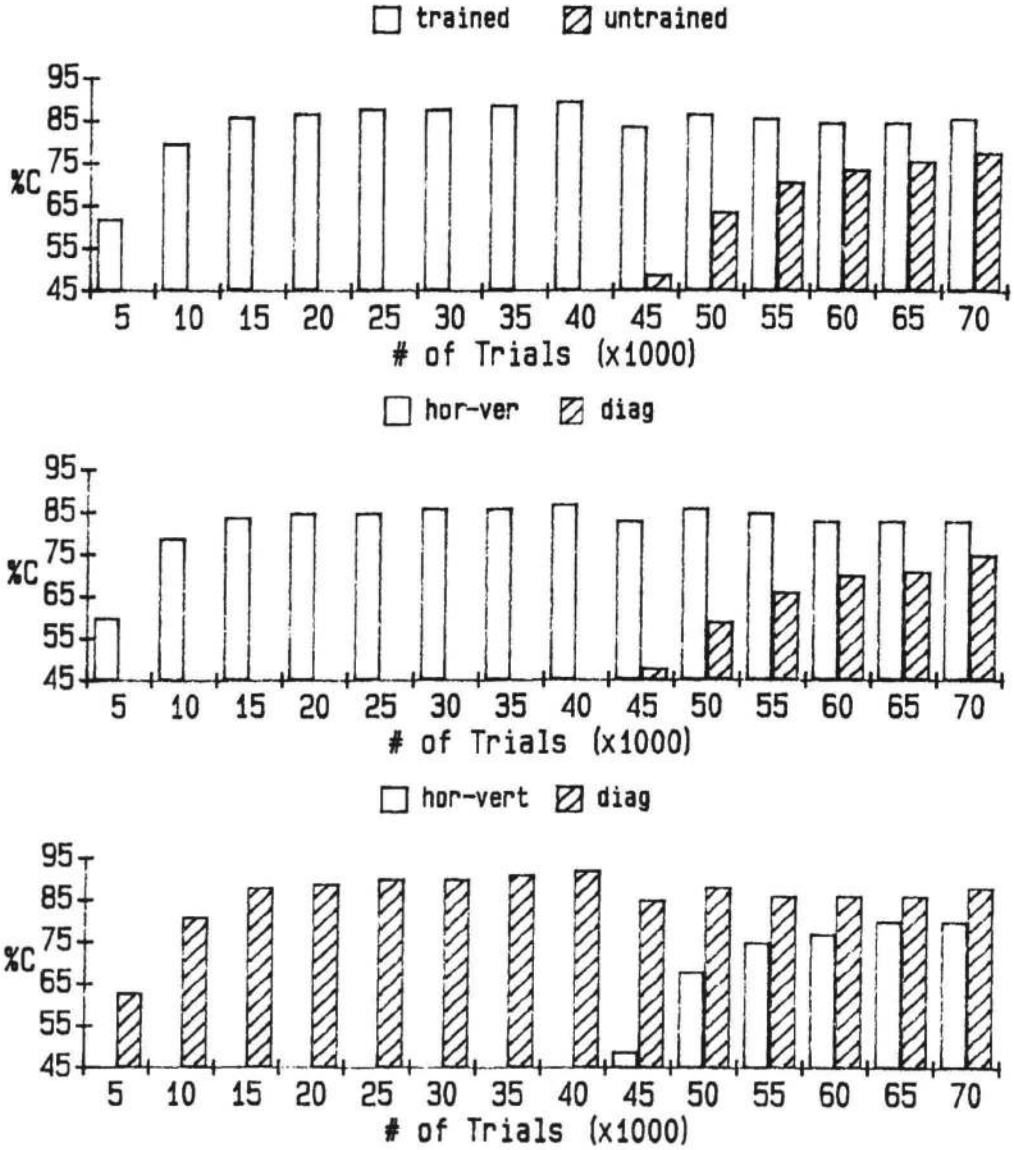

Figure 10. Percent correct for local-unit network when trained on either horizontal-vertical (middle) or diagonal axes (bottom) and tested on all four axes. Overall training effect is shown at top.

10). Not only did performance decline for the trained axes, especially in the global-unit network, but performance on the previously untrained axes was poorer at the start of the 4-axis testing than was performance for the other pair of axes at the start of the 2-axis training. Again, this was particularly evident for the global-unit network. Unlike the performance of human subjects, neither of the networks showed a preference for the horizontal-vertical axes over the diagonal axes, or vice versa. (Note that these data, like those presented for human subjects, do not take into account the difference between the training and testing phases in number of response alternatives. Preliminary data from simulations in which the networks are trained on 2 axes and then tested on only the other 2 axes suggests that the previously untrained axes neither benefit nor suffer from prior training on other axes.)

The weights for the hidden units in the two networks were similar to those discussed above. One finding of interest that comes from this particular task was the preponderance of hidden units in the local-unit network that were responsive to both of the two axes during the 2-axis training phase. Inspection of the hidden units with receptive fields in the center 3x3 section of the input array revealed an organization of weights that nearly always was symmetric for one axis and antisymmetric for the other. For the other units, in most of the cases in which the unit was responsive to only one axis, it was responsive to antisymmetry in that axis. Hidden units with receptive fields in portions of the input array not relevant to distinguishing between the axes being tested had small weights and were inactive. Upon switching to 4-axis testing, these inactive units rapidly took on weights appropriate to the new axes, making this network rather more responsive to the change in stimulus parameters than was the global-unit network. The global-unit network was forced to reconfigure its hidden-unit weights to conform to the new stimulus parameters.

DISCUSSION

The present work documents skill learning and repetition priming in symmetry detection for both human subjects and the Boltzmann machine. The behavioral data from humans and the simulation data from the network models share certain qualitative similarities, but are different in some significant quantitative and qualitative respects. The difference quantitatively between human and model performance is striking: The number of trials required for the networks to learn the task is some two orders of magnitude greater than that required by humans. Moreover, the performance of the networks for unprimed stimuli seems to plateau at approximately 70-80% correct, depending upon the condition, whereas the performance of our human subjects continues to improve with practice, and will approach 100% correct with enough trials. Note, however, that humans come to this task with considerable topographic knowledge and familiarity with symmetry detection, and their initial performance on the task is well above chance. By contrast, the Boltzmann machine starts its learning de novo; it must learn simultaneously about both symmetry and the critical aspects of symmetry detection for these particular materials. Its performance on the present task starts at chance. In addition to this difference in what must be learned, the discrepancy between the relatively small number of units used in the network models compared with the presumably huge number of neurons in the human visual system sensitive to symmetry is likely to be an important factor. Finally, note that it is extraordinarily difficult to know how to scale the performance of the network models vis-a-vis human performance in the absence of specific claims about the psychological relevance of "trials" or "machine operations" for these models. Until such claims are made, conclusions based on quantitative comparisons can be offered only tentatively.

In terms of qualitative comparisons of the experimental data and simulation data, it is noteworthy that humans showed a strong superiority for detecting vertical symmetries. The superiority of performance in both experiments for horizontal-vertical axes over diagonal axes was due overwhelmingly to superior performance for the vertical axis. Across the 5 blocks of Exp. 1, performance for the vertical axis (84% correct) showed nearly as much advantage over performance for the horizontal axis (77% correct) as for the diagonal axes (74% correct). This was as true in the initial block (% correct: v = 75, h = 66, d = 63) as in the final block (% correct: v = 91, h = 84, d = 81). This result is consistent with previous psychophysical work (Barlow & Reeves, 1979), but at odds with the performance of the networks. The superiority for vertical symmetries is thought to be due at least in part to the sensitivity of cells near the vertical midline to local features around the midline. Other features relevant to processing mirror symmetries must be processed more globally. Our modeling attempts included both a local-unit and global-unit network, a difference in connectivity of the hidden units that had important implications for the way in which the model was able to internally represent mirror symmetries. The global-unit network, with less pre-wired structure than the local-unit network, was more influenced by the structure imposed by the stimulus array: It remained responsive to primes over long lags and was more sensitive to the change from 2 to 4 axes. The limited receptive fields of processing elements in the local-unit network would seem to provide a closer model of the type of processing used by biological visual systems, although the connectivity among the local processing units actually established by learning in our networks may be quite different from that attained by biological systems. Evaluation of which of these networks provides a closer match to human performance awaits further testing.

The most important finding of the present studies is the emergence of priming from the network model without any additional mechanisms, such as short-term changes to the weights. These networks were designed to handle generalization, which is the essence of skill learning; the fact that they prove to be sensitive to individual items even over enormous lags is a striking finding, one that further links skill learning and repetition priming. This is the basic qualitative similarity between the simulation and human performance data.

Two conclusions can be drawn from the study of symmetry detection in humans and massively-parallel network models. First, priming for humans in this task cannot depend upon the activation of some pre-existing representation of the primed materials, since the stimuli used in these studies were novel. Second, priming is a latent property of incremental learning in the parallel network model and occurs from the start of de novo learning. Thus, both the behavioral data from humans and simulation data from the model suggest that repetition priming is not a separate mechanism and does not depend on an already structured system. Priming may instead be an integral feature of the basic learning process that configures and reorganizes processing modules. Parallel work with humans and network models, such as that reported in the present paper, can provide important insights about possible mechanisms of learning and memory.

REFERENCES

Ackley, D.H., Hinton, G.E., & Sejnowski, T.J. (1985). A learning algorithm for Boltzmann machines. Cognitive Science, 9, 147-169.

Anderson, J.R. (1982). Acquisition of cognitive skill. Psychological Review, 89, 369-406.

Anderson, J.R. (1983). The Architecture of Cognition. Cambridge: Harvard University Press.

Barlow, H.B. & Reeves, B.C. (1979). The versatility and absolute efficiency of detecting mirror symmetry in random dot displays. Vision Research, 19, 783-793.

Bruce, V.G. & Morgan, M.J. (1975). Violations of symmetry and repetition in visual patterns. Perception, 4, 239-249.

Biederman, I. (1985). Human image understanding: Recent research and a theory. Computer Vision, Graphics, and Image Processing, 32, 29-73.

Carpenter, G.A. & Grossberg, S. (1986). Neural dynamics of category learning and recognition: Attention, memory consolidation, and amnesia. In J. Davis, R. Newburgh, & E. Wegman (Eds.), Brain Structure, Learning, and Memory. AAAS Symposium Series.

Cohen, N.J. (1984). Preserved learning capacity in amnesia: Evidence for multiple memory systems. In L.R. Squire & N. Butters (Eds.), Neuropsychology of Memory. New York: Guilford Press.

Cohen, N.J., Eichenbaum, H., DeAcedo, B.S., & Corkin, S. (1985). Different memory systems underlying acquisition of procedural and declarative knowledge. Annals of the New York Academy of Sciences, 444, 54-71.

Delius, J.D. & Nowak, B. (1982). Visual symmetry recognition by pigeons. Psychological Research, 44, 199-212.

Fowler, C.A., Napps, S.E., & Feldman, L. (1985). Relations among regular and irregular morphologically related words in the lexicon as revealed by repetition priming. Memory & Cognition, 13, 241-255.

Garner, W.R. (1962). Uncertainty and Structure as Psychological Concepts. New York: Wiley.

Gordon, B. (1986). Preserved learning of novel information in amnesia: Evidence for multiple memory systems. (submitted)

Hinton, G.E. & Sejnowski, T.J. (1983). Optimal perceptual inference. Proceedings of the IEEE Computer Society Conference on Computer Vision & Pattern Recognition, Washington, DC, 448-453.

Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational abilities, Proceedings of the National Academy of Sciences (USA), 79, 2554-2558.

Kohler, W. (1929). Gestalt Psychology. NY: Harcourt Brace.

Kohlers, P.A. (1975). A pattern-analyzing basis of recognition. In L.S. Cermak & F.I.M. Craik (Eds.), Levels of Processing in Human Memory. Hillsdale, NJ: L. Erlbaum.

Marr, D. & Nishihara, H.K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. Perceedings of the Royal Society (London; series B), 200, 269-294.

Minsky, M. & Papert, S. (1969). Perceptrons. Cambridge: MIT Press.

Rosenblatt, F. (1959). Principles of Neurodynamics. NY: Spartan.

Rumelhart, D.E. & McClelland, J.L. (1986). Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Cambridge: MIT Press.

Schacter, D.L. (1985). Priming of old and new knowledge in amnesic patients and normal subjects. Annals of the New York Academy of Sciences, 444, 41-53.

Sejnowski, T.J. & Hinton, G.E. (1986). Separating figure from ground with a Boltzmann Machine. In M.A. Arbib & A.R. Hanson (Eds.), Vision, Brain & Cooperative Computation. Cambridge: MIT Press.

Sejnowski, T.J., Kienker, P.K., & Hinton, G.E. (in press). Learning symmetry groups with hidden units: Beyond the perception. Physica D.

Squire, L.R. & Butters, N. (1984). Neuropsychology of Memory. NY: Guilford Press.

Squire, L.R. & Cohen, N.J. (1984). Human memory and amnesia. In J.L. McGaugh, N.W. Weinberger, & G. Lynch (Eds.), Neurobiology of Learning and Memory. NY: Guilford Press.

Tulving, E. (1985). How many memory systems are there? American Psychologist.