Varieties of working memory as seen in biology and in connectionist/control architectures

WALTER SCHNEIDER
University of Pittsburgh, Pittsburgh, Pennsylvania

Biological and computational concepts that underlie the nature of working memory are briefly reviewed. The conceptualization of working memory has changed dramatically in the last 30 years. Current biological work has monitored several aspects of memory, including activation decay, sustained activation, long-term connection change, and differential structures for episodic (hippocampal formation) and procedural learning. Current connectionist modeling has identified factors including multiple-region-based processing, control processing as well as data storage, tradeoffs between fast- and slow-connection-change learning effects, and the speeding of acquisition via multiple levels of learning. The need to relate the biological, behavioral, and computational constraints into models of working memory is discussed. Finally, conceptualizations of working memory must acknowledge the need for human learning systems to be robust enough to operate in a dynamic world.

In the nearly 30 years since the seminal models of working memory were introduced (Atkinson & Shiffrin, 1968; Waugh & Norman, 1965), we have learned that the phenomena and mechanisms of working memory are far more complex than what the simple views of the 1960s suggested. In recent years, there has been important progress in the biological measurement of memory and in the computational modeling of memory effects. These new results have provided an important foundation for the study of short-term memory in its many facets. To deal with the complexity of human brain memory systems, we must incorporate a diversity of memories and utilize multiple behavioral and biological constraints to model the resulting systems.

In the 1960s, buffer models became the predominate structure for the understanding of memory. These portrayed a fairly simple view of memory as including several stages—sensory register, short-term store, and long-term store (Atkinson & Shiffrin, 1968). The models produced excellent fits to the results of laboratory tasks, as in the case of serial position curves obtained when subjects learned random word lists. However, the models missed critical qualitative features in the data (e.g., loading up short-term memory produced only minor deterioration in long-term learning; see Baddeley, 1986; Klapp, Marshburn, & Lester, 1983). Normal human performance was far more robust than that expected in the simple buffer models. (For example, how do humans recover from in-

terruptions that would clear small size buffers?—for an extended discussion, see Schneider & Detweiler, 1987.)

During the past 30 years, the behavioral data have become more complex, leading some researchers to question whether there even is an identifiable short-term memory (see Crowder, 1993). The behavioral data alone do not allow us to distinguish a single memory system with multiple time constants from several memory systems with different time constants. The researcher is faced with multiple models, all of which make equally good predictions. Appeals to parsimony have often encouraged researchers to utilize simple models, because the more complex models produce little improvement in the predictions and there are no sharp discontinuities in the behavioral data.

It is important to be very cautious about applying a parsimony criterion to the understanding of nervous systems. Neural systems have evolved through variations on a theme whereby new systems do not replace old systems but often arise through mutation in multiple ways to extend the range of cognitive function. When this mutation occurs, the multiple systems that coexist display a variety of qualitative differences, but they also overlap substantially during operation. In vision, for example, humans have both a rod system and a cone system, both of which respond to light, with the one more specialized for low-luminance monochrome, and the other, for high-luminance color. The visual system is also divided between the magno- and parvocellular pathways (see Van Essen & Anderson, 1990). Again, both respond to light, but one is more specialized for changing stimuli, whereas the other is more specialized for fine shape discrimination. It is likely that the biology of memory is analogously composed of multiple overlapping memory systems that are optimized for different classes of information storage. Indeed, this argument is becoming increasingly

prevalent in the memory literature (e.g., Squire, 1986; Tulving & Schacter, 1990).

In the 1990s we now have substantial biological data and simulation data illustrating a large variety of memory effects. Over the remainder of the decade, it will be a challenge to relate these data to mechanisms and from there to predictions of human learning in laboratory and nonlaboratory settings.

Biological Forms of Memory

The biology of memory works at substantially different time scales, employing quite different mechanisms. The neurobiological data illustrate activity decay, a maintainable short-term store, and a long-term store. Each of these operates by means of a different mechanism.

Activity decay is inherent in all biological signals. When a transmitter is released, there is a process of biological uptake at the receptive site, alteration of the membrane potential, passive decay of the membrane, and possibly active propagation of the potential. These operations occur at a variety of time scales, depending on the transmitters involved (see Hille, 1992; McCormick, 1990). It is important that some of these decays at the membrane

level occur over a range of hundreds of milliseconds. Slow inactivating Na^+ and K^+ channels can require decay over several seconds (see, e.g., Surmeier, Stefani, Foehring, & Kitai, 1991). Activity decay at multiple time constants from milliseconds to seconds is probably fairly generic throughout many areas of the cortex. The decay occurs over a time course that would be typical of sensory register type decay (see Long, 1980, for a review).

Sustained recurrent activation is a form of short-term memory observed in multiple memory areas that is consistent with an actively maintained short-term store. Funahashi, Bruce, and Goldman-Rakic (1989) have presented primate monkeys with an oculomotor delayed response task (see Figures 1A and 1B). The monkey was required to fixate a central stimulus, one of eight possible peripheral cues was then blinked for 40 msec, the fixation stimulus then disappeared after a delay period (1.5–6 sec after the peripheral blink), and the monkey moved fixation to the location of the peripheral cue for a reward. The monkey had to maintain in memory the location of the cue during the fixation time. Figure 2 shows the basic results as a function of delay periods. Of the 288 principal sulcus neurons, 59% had task-relevant discharge and 30% had
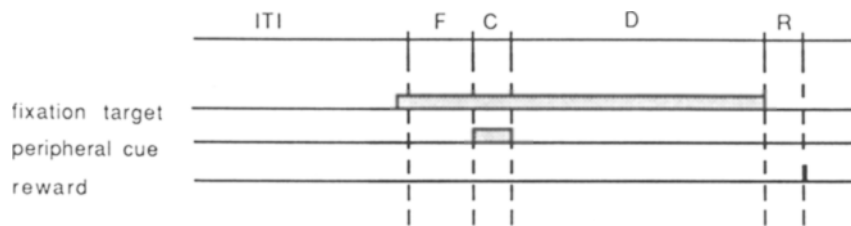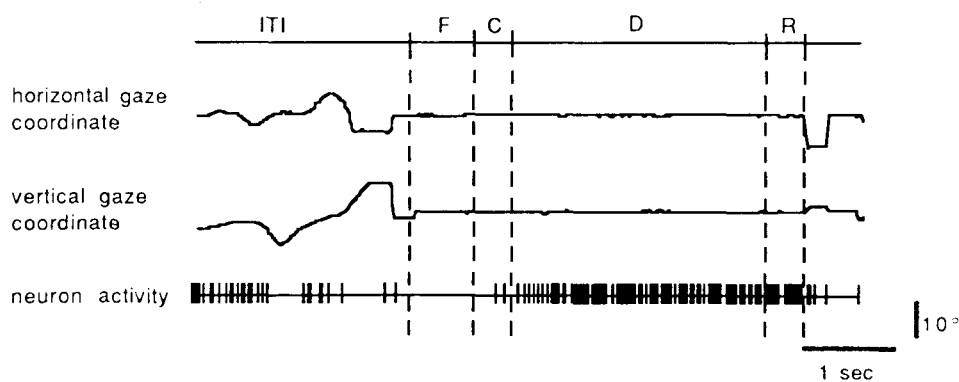

Figure 1. (A) Temporal sequence of events in the oculomotor delayed-response task. (B) Eye movements and single neuron activity during an oculomotor delayed-response trial. ITI, intertrial interval; F, fixation period (0.75 sec); C, cue period (0.5 sec); D, delay period (3 sec); R, response period (0.5 sec). From "Mnemonic Coding of Visual Space in the Monkey's Dorsolateral Prefrontal Cortex," by S. Funahashi, C. J. Bruce, and P. S. Goldman-Rakic, 1989, *Journal of Neurophysiology*, 61, pp. 331-349 (Figure 1). Copyright 1989 by the American Physiological Society. Reprinted by permission.
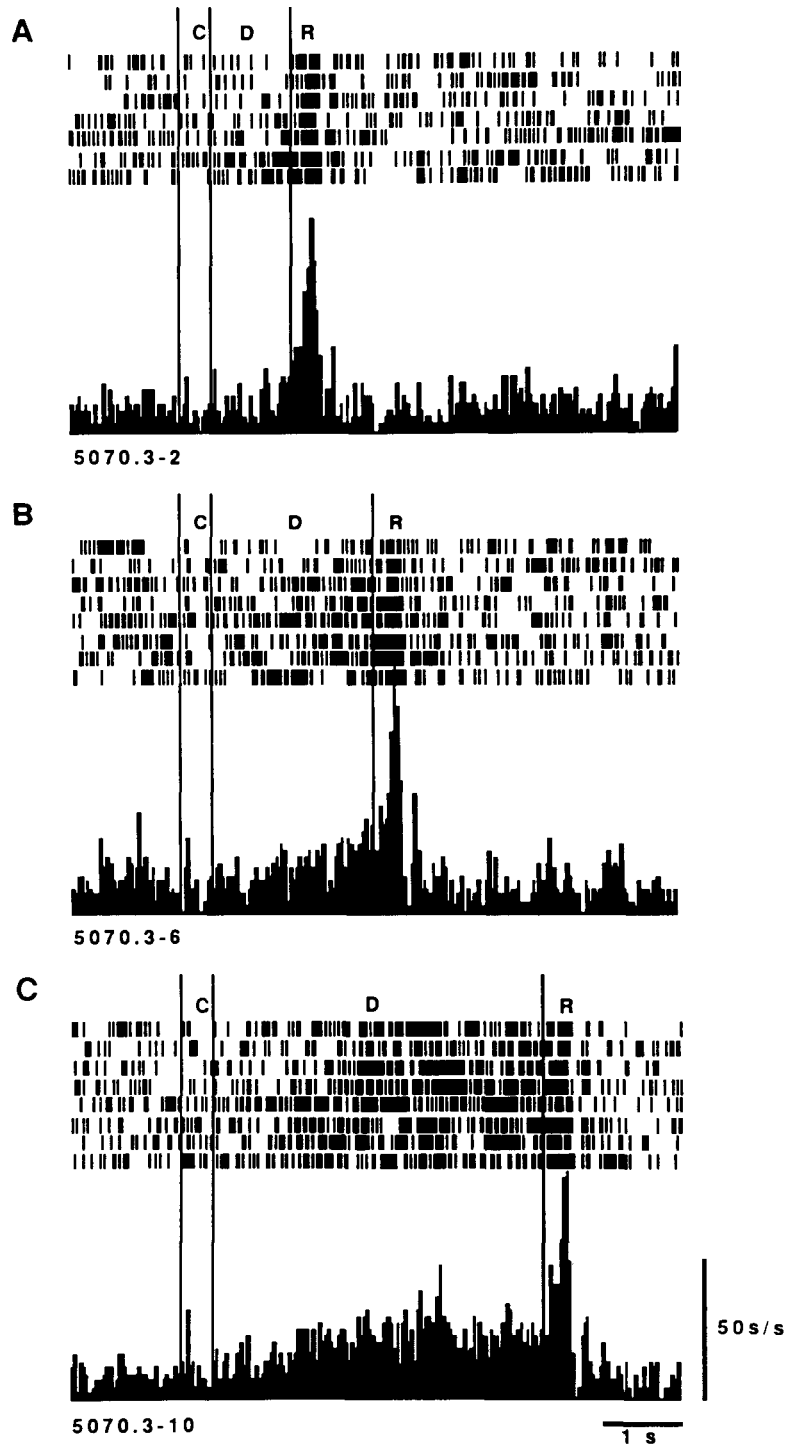
Figure 2. Directional delay period activity of a principal sulcus neuron (5070, right hemisphere) for three different delay durations (1.5-sec delay in A, 3-sec delay in B, and 6-sec delay in C). The data are from an experiment with 12 different types of trials: the three delay durations crossed with 4 visual cue locations (0° = right, 90° = up, 180° = left, and 270° = down). All cue eccentricities were 13°. Only the three histograms from the preferred direction (180° location) are shown. From "Mnemonic Coding of Visual Space in the Monkey's Dorsolateral Prefrontal Cortex," by S. Funahashi, C. J. Bruce, and P. S. Goldman-Rakic, 1989, *Journal of Neurophysiology*, 61, pp. 331-349 (Figure 12). Copyright 1989 by the American Physiological Society. Reprinted by permission.

significant delay-period activity. The responding neurons responded for longer periods of time when the delay was extended. If the activity of these units decayed during the delay interval, the animal was more likely to make an error (see Figure 3). If a second peripheral cue occurred during the fixation period, and the monkey had to change the content of memory, some of the units altered their response. These data indicate that a form of short-term memory has been directly monitored. It is of relatively short duration (less than 10 sec), under direct control of the subject, and it occurs in several areas of the cortex

(e.g., frontal eye fields, Bruce & Goldberg, 1985; temporal cortex, Miyashita & Chang, 1988; inferomedial temporal cortex, Brown, Wilson, & Riches, 1987; parietal cortex, Gnadt & Andersen, 1988; and prefrontal cortex, Funahashi et al., 1989). New results in this area will produce detailed specification of the duration over which such information is maintained.

This type of memory has been modeled by Zipser (1991) in a connectionist network in which the information is stored in the dynamic interaction within the neural population. The units excite one another, thereby main-

A. Correct trials

s f 0 3 - 3

40 s / s

1 s

B. Error trials

s f 0 3 - 1

40 s / s
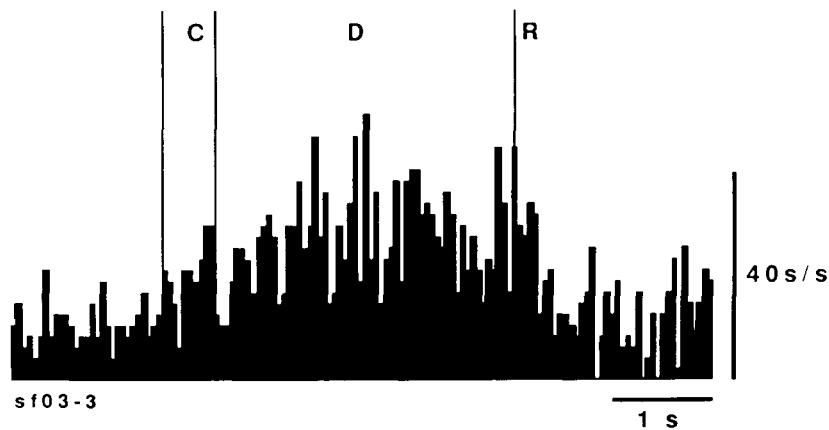
1 s
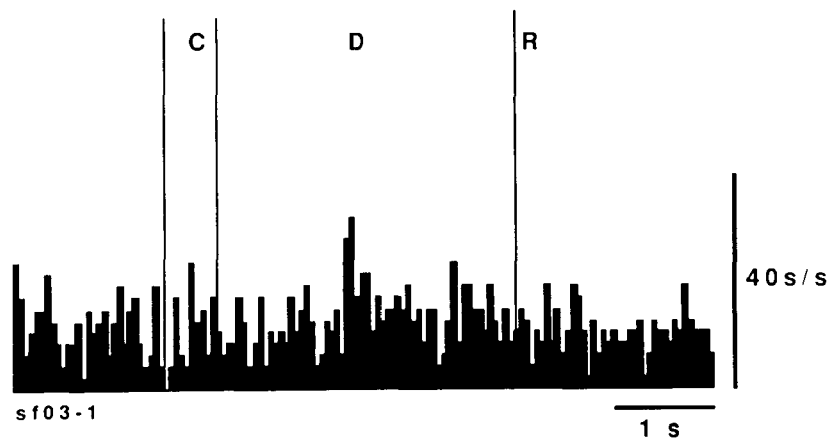
Figure 3. Comparison of delay period activity for correct trials with activity for error trials. The first correct trial and the first error trial were taken from each of the nine principal sulcus neurons that had excitatory directional delay period activity and for which the monkey made at least one error on a trial in the neuron's preferred direction. The top histogram sums the nine correct trials; the bottom histogram sums the nine error trials. From "Mnemonic Coding of Visual Space in the Monkey's Dorsolateral Prefrontal Cortex," by S. Funahashi, C. J. Bruce, and P. S. Goldman-Rakic, 1989, *Journal of Neurophysiology*, 61, pp. 331-349 (Figure 14). Copyright 1989 by the American Physiological Society. Reprinted by permission.

taining an active pattern. Shedden and Schneider (1990) have utilized a positive feedback circuit to produce this type of memory buffer.

Connection change provides a long-term memory. Since the time of Hebb (1949), psychologists have generally assumed that long-term memory involves the change in connection weights between neurons. In the last decade, neuroscientists have tracked long-term changes in connection strength that last for durations of weeks in hippocampal neurons. The paradigm is referred to as long-term potentiation, or LTP (see Gustafsson & Wigstrom, 1988). It typically involves stimulating both a weak and a strong input to a pyramidal cell, and then showing that the evoked potential of the weak cell increases only if it is paired with the strong input, the one that is assumed to fire the cell. One type of LTP occurs after a modest number of pairings, occurring in a time scale of 20 sec and lasting for days and perhaps weeks, though not necessarily permanently. In the hippocampus, this change can produce dramatic changes in synaptic efficiency (e.g., doubling the excitatory effect of the connections).

One type of connection change appears to be modulated via the NMDA receptors. These receptors vary in concentration across cortical and subcortical areas, suggesting that some areas may have faster learning rates than others. Hippocampal areas showing high concentrations perhaps encode episodic information requiring fast learning, whereas cortical neurons having lower concentrations perhaps encode more procedural information.

Hippocampally based and cortically based memory effects show qualitative differences and differential locations of effects. Squire (1992) has provided an extensive review of the hippocampal literature. The hippocampus and perirhinal and parahippocampal cortices mediate a declarative memory system encoding episodic and semantic fact events, maintaining information for modest periods of time (from greater than 10 sec to minutes to perhaps days). This system is particularly important for episodic memory of recent events. Damage to the hippocampal formation causes severe amnesic responses in recognition tasks but still spares nondeclarative learning, such as that of skills, of priming conditions, and in habituation.

Physiological structures in different locations are involved in learning alternative information. Paller (1990) showed with evoked response potentials (ERPs) that a task involving completing partial word stems with words from a previously learned list produced anterior positive ERPs, whereas simple stem completion without such recollection produced less positive ERP responses. Simple priming did not produce the large anterior positivity, but did produce more effects on posterior electrodes. Recent PET blood flow studies (Squire et al., in press) have shown increased activity in the right hippocampal region for recall-based stem completion in contrast with priming tasks not requiring recall. A particularly exciting new development in the domain of functional MRI-based blood flow methodology is likely to allow detailed within-subject

mapping of the locations of different types of memory processing in the future (McCarthy, personal communication).

The variety of identified biological memories described above provides psychology theorists with a range of mechanisms for working memory. Memory is not one mechanism in one place. Activation decay and long-term change probably occur in all areas, since activation and connection-based transmission occur in all areas. In contrast, sustained activation and storage of declarative information may occur in more limited areas. These effects have been observed in only a few areas of the brain.

Connectionist-Based Conceptualizations of Working Memory

Recent changes in simulation trends have widened the range of conceptualizations of working memory. In the 1960s, analogies to computers produced the basic view of a short-term memory (analogous to the computer core memory) through which all information passed, a portion of which altered long-term memory (analogous to computer disk memory) as in Atkinson and Shiffrin (1968) and Waugh and Norman (1965). This view has been elaborated by Baddeley and his colleagues (Baddeley, 1983, 1986; Baddeley & Hitch, 1974) to include separable subsystems consisting of an articulatory loop, a visuospatial scratchpad, and a central executive. In the late 1980s, interest in connectionist modeling grew greatly (see Rumelhart & McClelland, 1986; Schneider, 1987). This provided a novel class of information storage analogies that can more directly implement the parallel distributed nature of storage that is expected to occur in biological systems. Schneider and Detweiler (1987) have reviewed the traditional models of short-term memory and discussed alternatives suggested from connectionist processing. Connectionist modeling suggests a number of themes that are biologically inspired and that can be implemented in connectionist simulations.

The first theme is that processing occurs in multiple regions that specialize in different classes of processing (see Figure 4) such as visual, motor, and lexical processing. Some modules specialize in context-based storage of declarative information. In addition, there is a control structure that moderates transmissions in the system. These modules correspond to basic biological levels of processing (see Mishkin & Appenzeller, 1987). All these models have multiple memories, such as activity decay and long-term connection change. Some connection change occurs at fast rates for declarative storage but at slow rates for skill encoding.

The second theme is that there is a need for a control structure as well as for a data storage structure. This need was well recognized in the buffer models such as that of Atkinson and Shiffrin (1968). In a distributed system, each module can store information, but the transfer of information must be coordinated. The controller is assumed to have memory concerning which modules have transmitted and the activity and priority of messages needing

A

AUDITORY

| PHONEMES |
| SYLLABLES |
| WORDS |

VISUAL

SPEECH

CONTEXT

LEXICAL

MOOD

SEMANTIC

| TASKS |
| SEQUENCES |
| MOVEMENTS |

MOTOR

B

VISUAL   AUDITORY   SPEECH

CENTRAL CONTROL

━━ MESSAGE VECTOR

CONTROL SIGNALS:

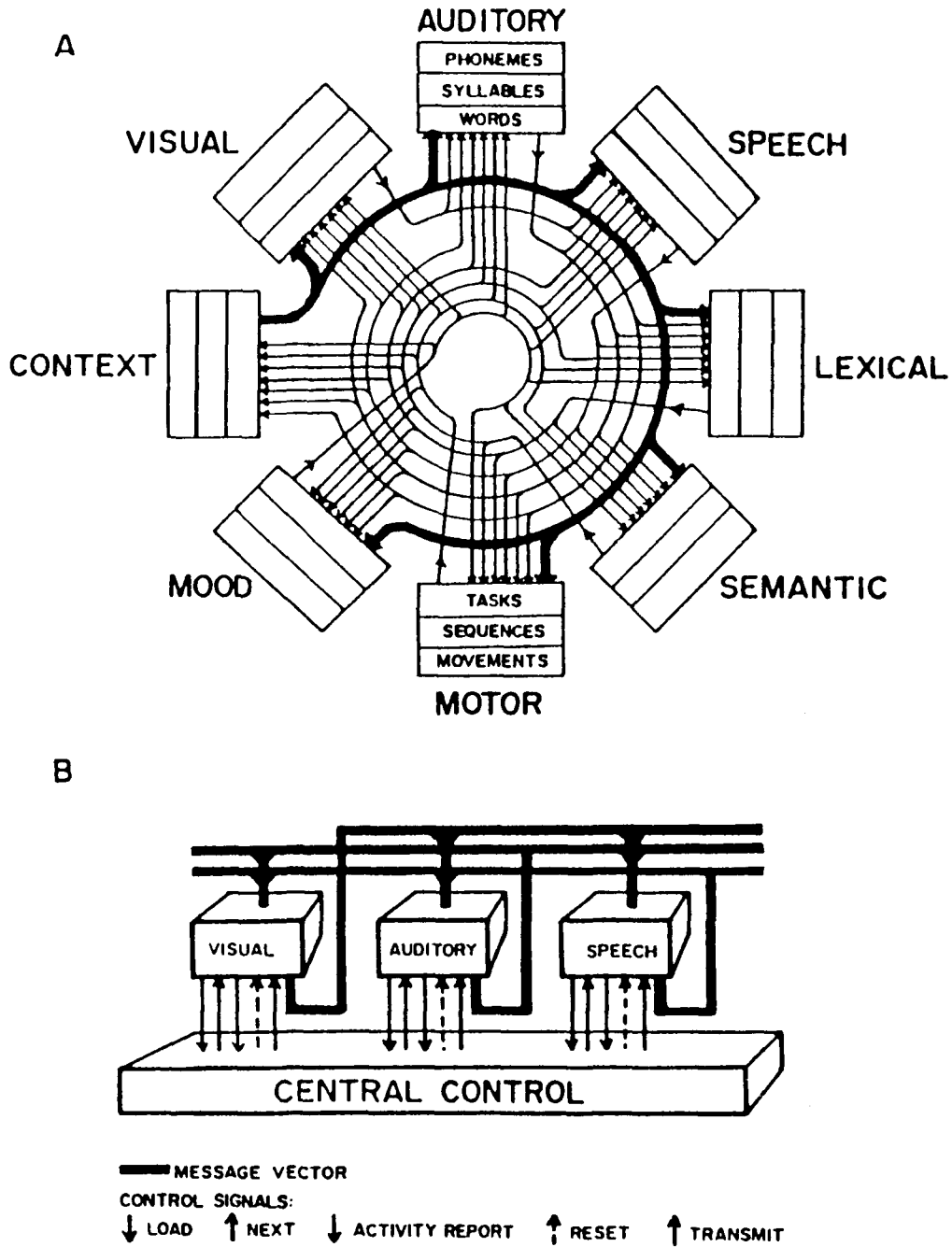↓ LOAD   ↑ NEXT   ↓ ACTIVITY REPORT   ↑ RESET   ↑ TRANSMIT

Figure 4. Diagram of a system-level description of the Schneider and Detweiler (1987) connectionist control architecture, illustrating the regions of processing. Each region represents a series of connectionist processing levels. The first and last level of each region (last level for input regions and first for output regions) is assumed to input to the inner loop of connections between regions. The modules on the inner loop have separate message vectors to each of the other modules to which they connect. The context module output message vector is thickened to highlight one set of outputs. The lower figure represents a side view of the system-level diagram. All the regions of the inner loop connect to a central control system that routes control signals between regions of the inner loop. The system manages message traffic and context storage to maintain reliable communication and robust processing across regions. From "A Connectionist/Control Architecture for Working Memory," by W. Schneider and M. Detweiler, 1987, in G. H. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. 21, pp. 53-119, Figure 3). Copyright 1987 by Academic Press. Reprinted by permission.

to be transmitted. Retrieval failure can occur through either the loss of the data from a regional module or the failure of the control system to transmit information to or from the regional module.

A third theme is that short-term information can be stored in many ways. It can be stored as decay of a single trace (Ratcliff, 1978). It can be stored by positive feedback within a module (Shedden & Schneider, 1990), maintaining information via recurrent connections within the module. It can be stored via dividing a region into multiple buffers to hold a small number of elements (Schneider & Detweiler, 1987). Or it may be stored in a recurrent network. Recurrent networks (see Elman, 1990) are a new branch of connectionist models in which the output from the intermediate level of units connect back to intermediate units. This looping provides these networks with the ability to compact a sequence of inputs into one code (e.g., given a vector representing CAT, to output the sequence of letters C, A, and T that make up that input).

A fourth theme consists of a tradeoff between learning speed and retroactive and proactive interference. In connectionist models, one sets a parameter to determine how quickly a connection changes, on the basis of the difference between the desired and the actual output of the network. For example, consider a connectionist model that must learn a list of 20 paired associates. As each input is presented, the network predicts the output, and then the difference between the network-produced output and the desired output results in an error signal. The connections are changed to reduce that error. If the learning rate is set to one, all the error can be eliminated in one trial so that the network can exactly produce the input–output pair of the last trial. However, if the learning rate is high, retroactive interference is complete, causing complete forgetting of all previously learned information (see Schneider & Detweiler, 1987, Figure 6) . The faster the learning rate, the less the proactive interference and the greater the retroactive interference problem. Most connectionist learning programs use low learning rates (e.g., values below .1), but this produces learning that requires hundreds and sometimes thousands of learning trials to acquire input learning sets.

The tradeoff of fast and slow learning implies that if a system learns fast it will forget rapidly. Faced with such a tradeoff, it would be advantageous to evolve both fast and slow learning rate systems. Slow learning systems would learn skills when repeated experience is provided. Fast learning systems would rapidly learn specific information, but the availability of that information would be transient because of the severe retroactive interference effects. This dual system view seems to parallel the physiological data regarding fast episodic and slow skill learning effects (see Mishkin & Petri, 1984; Squire, 1992).

The fifth theme is the need for multiple levels of connectionist learning, perhaps to quickly acquire a task and guide the development of skill acquisition. Standard connectionist learning often "learns" at rates far too slow for human performance. For example, to learn digital

logic truth tables with six inputs required 10,835 trials for a connectionist network and only 216 trials for humans (Schneider & Oliver, 1991). Human learning contrasts sharply with typical connectionist learning in that humans can benefit greatly from instruction, acquire complex behavior in a few trials, and decompose complex tasks into smaller subtasks (see Schneider & Oliver, 1991). These characteristics result in large-scale improvements in learning effectiveness. Given sufficient trials, a three-layer network can learn complex relationships. However, the learning time may require more trials than could occur in the lifetime of a human, and hence the learning system may be of limited survival value (e.g., requiring 2 billion stimulus presentations to learn a seven-state finite state grammar; see Servan-Schreiber, Cleeremans, & McClelland, 1988).

Having multiple levels of learning can greatly speed skill acquisition and facilitate permanent storage and fast responding. To enable learning in a human time scale and to allow instruction to benefit connectionist learning, Schneider and Oliver (1991) developed a two-level connectionist control architecture. Modules in a data network performed normal connectionist input mapping in three-layer nets in a data network. In addition, there was a recurrent connectionist network that controlled parameters of the data modules. The control network was a fast-learning recurrent net that could learn declarative information to perform the task. For example, to perform an AND task, it would learn the sequence of: "attend to the visual inputs, compare to 1 state, if all 1s attend to the input state model, set to all 1 state." This declarative representation could be learned in 20 trials per rule, which is faster than humans learned the task (36 trials per rule). This declarative knowledge allowed the system to interpretively execute control steps to perform the task in a serial effortful manner. Having obtained controlled execution of the task, the connectionist control architecture could continue to practice without a teacher and acquire the knowledge in the connectionist data network in a modest number of trials (932 trials to learn six input gates relative to a human learning time of about 2,000 trials for developing automatic gate knowledge; see Schneider & Oliver, 1991). The dual system of fast declarative learning and slower skill learning allows the system to acquire the task in small numbers of trials within the control network. The control network can, after a small number of trials, exercise the data network, slowly modifying long-term connections. After many trials the trained data network will then execute low-effort high-speed automatic behaviors. The automatic behaviors are also less susceptible to the effects of retroactive interference.

The sixth theme from the modeling and review of the memory literature (Schneider & Detweiler, 1987) is the need to have robust memory systems. The buffer-oriented systems do not provide mechanisms that allow information to be recovered after an interruption that flushes the buffers. Buffer models tend to create an amnesia-prone operation similar to that observed in the patient H.M.

(Milner, 1966). If distracted from the task at hand, the patient H.M. must be reinstructed from the beginning to restart the task. If long-term storage is slow and the short-term buffers are flushed through an interruption, a buffer model must be reloaded. Normal humans are quite adept at dealing with interruptions and at restarting with little lost effort (e.g., phone interruptions do not cause one to lose the full train of thought developed in the hour before). This is a serious problem for buffer interpretations of memory. To duplicate the robustness of human processing, it may be beneficial to develop overlapping memory systems that can reload the context that existed before the interruption.

A fast-learning connectionist context learning mechanism (see Schneider & Detweiler, 1987) illustrates such robust processing. If a context module had fast-learning connections into many modules, it could associate the current active state in each of the modules to a slow, varying context vector. If an interruption occurred, the current activation of the modules might be cleared to process the interruption. By transmitting the context vector state before the interruption, the pre-interruption states could be re-evoked. However, note that to allow fast learning of the context, the system would show severe retroactive interference. Hence, although the fast context learning systems allow dealing with interrupts, it is not a substitute for the slower connectionist learning that supports long-term retrieval with much less severe retroactive interference.

Dealing With the Complexity of Human Memory and the Models

In the preceding description, I have detailed a set of components and concepts that need to be integrated into models of working memory and memory in general. Memory is not just a buffer, association net, set of slave processors, or control structure but more likely is made up of all of these things and more. It is the result of multiple biological mechanisms with different time courses, including activity decay (subsecond), sustained activation (seconds), connection change (minutes to weeks), and hippocampally and cortically based storage. The phenomena of working memory are many (see reviews by Baddeley, 1976, 1986; and by Schneider & Detweiler, 1987), and detailed models must account for effects such as the limited storage of three to five items; acoustic confusions; sequential processing; problems with digit span and reverse digit span; difficulty in maintaining order information; the nature of rehearsal, episodic, and semantic memory distinctions; retroactive and proactive interference effects; buildup of proactive interference; release from proactive interference; the benefit of elaborate rehearsal; and long-term memory recency effects (see Schneider & Detweiler, 1987, for details of these phenomena).

In evaluating our conceptions of human working memory, we must consider whether the resulting system is robust and can survive in the complex dynamic world in which humans have evolved. This need for robust processing may have resulted in the evolution of multiple memory systems with both overlapping and complementary functions. Humans need a very fast memory to maintain information when it is no longer sensorially available. Such an episodic storage system may be crucial for dealing with interruptions and maintaining information to train a slow-learning-skill and semantic fact system. The slow-learning system may in turn provide for fast automatic execution of well-practiced skills and very stable long-term retention (e.g., over decades) of well-practiced information.

The developing base of physiological, behavioral, and computational constraints provides a three-legged foundation from which to build a detailed model of human memory. Currently, there is a sufficient range of results in each of these areas to provide a basis for such modeling (see, e.g., Schneider & Detweiler, 1987; Schneider & Oliver, 1991). Although it is true that such models are much more complex than the original buffer models (e.g., that of Atkinson & Shiffrin, 1968), they incorporate many new and important contributions into our understanding of the varieties of working memory. This multiple-constraint approach has the prospect of accounting for the many facets and mechanisms of working memory.

REFERENCES

ATKINSON, R. C., & SHIFFRIN, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2, pp. 89-195). New York: Academic Press.

BADDELEY, A. D. (1976). *The psychology of memory*. New York: Basic Books.

BADDELEY, A. D. (1983). Working memory. *Philosophical Transactions of the Royal Society of London: Series B*, **302**, 311-324.

BADDELEY, A. D. (1986). *Working memory*. Oxford: Oxford University Press/Clarendon Press.

BADDELEY, A. D., & HITCH, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47-89). New York: Academic Press.

BROWN, M. W., WILSON, F. A. W., & RICHES, I. P. (1987). Neuronal evidence that inferomedial temporal cortex is more important than hippocampus in certain processes underlying recognition memory. *Brain Research*, **409**, 158-162.

BRUCE, C. J., & GOLDBERG, M. E. (1985). Primate frontal eye fields: I. Single neurons discharging before saccades. *Journal of Neurophysiology*, **53**, 603-635.

CROWDER, R. (1993). Short-term memory: Where do we stand? *Memory & Cognition*, **21**, 142-145.

ELMAN, J. L. (1990). Finding structure in time. *Cognitive Science*, **14**, 179-211.

FUNAHASHI, S., BRUCE, C. J., & GOLDMAN-RAKIC, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology*, **61**, 331-349.

GNADT, J. W., & ANDERSEN, R. A. (1988). Memory related motor planning activity in posterior parietal cortex of macaque. *Experimental Brain Research*, **70**, 216-220.

GUSTAFSSON, B., & WIGSTROM, H. (1988). Physiological mechanisms underlying long-term potentiation. *Trends in Neurosciences*, **11**, 156-162.

HEBB, D. O. (1949). *The organization of behavior*. New York: John Wiley.

HILLE, B. (1992). *Ionic channels of excitable membranes* (2nd ed.). Sunderland, MA: Sinauer Associates.

KLAPP, S. T., MARSHBURN, E. A., & LESTER, P. T. (1983). Short-term

memory does not involve the "working memory" of information processing: The demise of a common assumption. *Journal of Experimental Psychology: General*, **112**, 240-264.

LONG, G. M. (1980). Iconic memory: A review and critique of the study of short-term visual storage. *Psychological Bulletin*, **88**, 785-820.

McCORMICK, D. A. (1990). Membrane properties and neurotransmitter actions. In G. M. Shepherd (Ed.), *The synaptic organization of the brain* (3rd ed., pp. 32-66). Oxford: Oxford University Press.

MILNER, B. (1966). Amnesia following operation on the temporal lobe. In C. W. M. Whitty & O. L. Zangwill (Eds.), *Amnesia* (pp. 109-133). London: Butterworth.

MISHKIN, M., & APPENZELLER, T. (1987, June). The anatomy of memory. *Scientific American*, **256**, 80-89.

MISHKIN, M., & PETRI, H. L. (1984). Memories and habits: Some implications for the analysis of learning and retention. In L. R. Squire & N. Butters (Eds.), *Neuropsychology of memory* (pp. 287-296). New York: Guilford.

MIYASHITA, Y., & CHANG, H. S. (1988). Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature*, **331**, 68-70.

PALLER, K. A. (1990). Recall and stem-completion priming have different electrophysiological correlates and are modified differentially by directed forgetting. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 1021-1032.

RATCLIFF, R. (1978). A theory of memory retrieval. *Psychological Review*, **85**, 59-108.

RUMELHART, D. E., & McCLELLAND, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 1. Foundations.* Cambridge, MA: MIT Press/Bransford Books.

SCHNEIDER, W. (1987). Connectionism: Is it a paradigm shift for psychology? *Behavioral Research Methods, Instruments & Computers*, **19**, 73-83.

SCHNEIDER, W., & DETWEILER, M. (1987). A connectionist/control architecture for working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 21, pp. 54-119). New York: Academic Press.

SCHNEIDER, W., & OLIVER, W. L. (1991). An instructable connectionist/control architecture: Using rule-based instructions to accomplish connectionist learning in a human time scale. In K. Van Lehn (Ed.), *Architectures for intelligence: The 22nd Carnegie Mellon symposium on intelligence* (pp. 113-145). Hillsdale, NJ: Erlbaum.

SERVAN-SCHREIBER D., CLEEREMANS, A., & McCLELLAND, J. L. (1988). *Encoding sequential structure in simple recurrent networks* (Tech. Rep. CMU-CS-88-183). Pittsburgh, PA: Carnegie Mellon University, Department of Computer Science.

SHEDDEN, J. M., & SCHNEIDER, W. (1990). A connectionist model of attentional enhancement and signal buffering. In *The Twelfth Annual Conference of the Cognitive Science Society* (pp. 566-573). Hillsdale, NJ: Erlbaum.

SQUIRE, L. R. (1986). Mechanisms of memory. *Science*, **232**, 1612-1619.

SQUIRE, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, **99**, 195-231.

SQUIRE, L. R., OJEMANN, J., MIEZIN, F., PETERSEN, S., VIDEEN, T., & RAICHLE, M. (in press). Activation of the hippocampus in normal humans: A functional anatomical study of memory. In *Proceedings of the National Academy of Sciences.* New York: National Academy of Sciences.

SURMEIER, D. J., STEFANI, A., FOEHRING, R. C., & KITAI, S. T. (1991). Developmental regulation of a slowly-inactivating potassium conductance in rat neostriatal neurons. *Neuroscience Letters*, **122**, 41-46.

TULVING, E., & SCHACTER, D. L. (1990). Priming and human memory systems. *Science*, **247**, 301-306.

VAN ESSEN, D. C., & ANDERSON, C. H. (1990). Information processing strategies and pathways in the primate retina and visual cortex. In D. C. Van Essen & C. H. Anderson (Eds.), *An introduction to neural and electronic networks* (pp. 43-72). New York: Academic Press.

WAUGH, N. C., & NORMAN, D. A. (1965). Primary memory. *Psychological Review*, **72**, 89-104.

ZIPSER, D. (1991). Recurrent network model of the neural mechanism of short-term active memory. *Neural Computation*, **331**, 679-684.