Vendrell, P., Junque, C., Pujol, J., Jurado, M. A., Molet, J., & Grafman, J. (1995). The role of prefrontal regions in the Stroop task. *Neuropsychologia, 33*, 341–362.
Verin, M., Partiot, A., Pillon, B., Malapani, C., Agid, Y., & Dubois, B. (1993). Delayed response tasks and prefrontal lesions in man – Evidence for self generated patterns of behaviour with poor environmental modulation. *Neuropsychologia, 31*, 1379–1396.
Warach, S., Gur, R. C., Gur, R. E., Skolnick, B. E., Obrist, W. D., & Reivich, M. (1992). Decreases in frontal and parietal lobe regional cerebral blood flow related to habituation. *Journal of Cerebral Blood Flow and Metabolism, 12*, 546–553.
Watanabe, T., & Niki, H. (1985). Hippocampal unit activity and delayed response in the monkey. *Brain Research, 325*, 241–254.
Wilkins, A. J., Shallice, T., & McCarthy, R. (1987). Frontal lesions and sustained attention. *Neuropsychologia, 25*, 359–365.
Wilson, F. A. W., O'Scalaidhe, S. P., & Goldman-Rakic, P. S. (1993). Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science, 260*, 1955–1958.
Witkin, H. A., & Goodenough, D. R. (1981). *Cognitive styles: Essence and origins*. New York: International Universities Press.
Woltz, D. J. (1988). An investigation of the role of working memory in procedural skill acquisition. *Journal of Experimental Psychology: General, 117*, 319–331.
Yntema, D. B. (1963). Keeping track of several things at once. *Human Factors, 5*, 7–17.

Lovett, M., Reder, L. M., Lebiere, C. (1999). Modeling Working Memory in a Unified Architecture: An ACT-R Perspective. In Miyake, A. and Shah, P. (Eds). *Models of Working Memory*. Oxford University Press, pp. 135-182.

5 Modeling Working Memory in a Unified Architecture

An ACT-R Perspective

MARSHA C. LOVETT, LYNNE M. REDER, AND CHRISTIAN LEBIERE

FIVE CENTRAL FEATURES OF THE MODEL

We describe a model of working memory that is developed within the ACT-R cognitive architecture. Some of its main features are derived from the basic features of ACT-R:

(1) Processing depends on the current goal of the system.

(2) The accessibility of declarative and procedural knowledge varies with experience.

In addition, the following features are important to working memory in particular:

(3) There is a limited attentional resource, focused on the current goal, that increases the accessibility of goal-relevant knowledge relative to other knowledge.

(4) In more complex and memory-demanding tasks, this limited resource is spread more thinly thus impairing retrieval of goal-relevant items.

(5) The "capacity" of this attentional resource may vary from person to person, influencing the ability to access goal-relevant information across domains.

In performing almost any cognitive task, one must engage *working memory* to maintain and retrieve information during processing. For example, in mental arithmetic (e.g., multiplying large numbers without pencil and paper), one must hold intermediate results in memory while solving the problem. Similarly, in sentence processing, one must maintain various syntactic and

We would like to thank John Anderson, Mike Byrne, Andrew Conway, Sott Filipino, and all of the symposium participants for helpful comments. The work presented here was funded by Grant F49620-97-1-0054 from the Air Force Office of Scientific Research, and also supported in part by Grant N00014-95-1-0223 from the Office of Naval Research and Grant 1R01 MH52808-01 from the National Institutes of Mental Health.

semantic structures until subsequent processing reveals their roles. Because working memory is involved in so many tasks, studying its characteristics and its impact on cognitive processes is critical to gaining a deeper understanding of how people perform cognitive tasks in general.

Past research highlights two important results, each of which demonstrates that working memory modulates task performance. First, when the working memory demands of a task increase (either by increasing the number of items in a "pure" memory task or by increasing the difficulty of concurrent processing in a dual-task situation), errors and latencies tend to increase (e.g., Anderson, Reder, & Lebiere, 1996; Baddeley, 1986; Caplan, Rochon, & Waters, 1992). Second, groups of subjects who have been separately identified as having "low" or "high" working memory capacity exhibit different degrees of sensitivity to increases in working memory load (e.g., Engle, 1994; Just & Carpenter, 1992). Together, these results suggest a view of working memory as a cognitive resource that (a) can be allocated to enable the maintenance and processing of information, (b) is inherently limited, and (c) differs in supply across individuals.

Within this view, however, there are still multiple ways for working memory to be implemented. Computational modeling has therefore contributed greatly to this field because it requires that the working memory mechanisms posited by a given theory be specified in a rigorous (programmable) way. The resulting model provides a detailed account of how such mechanisms interact and yields quantitative predictions that can be compared with observed data. This makes possible the systematic comparison of models representing different theories. Computational models are also particularly appropriate for studying individual differences in working memory because they enable researchers to maintain the basic structure of their theory while perturbing a particular "individual difference" component. The different patterns of results that such a variablized model exhibits can be compared with the different patterns of results displayed across subjects performing the same task. Thus, computational models can be evaluated not only according to how well they fit aggregate data but according to how well they account for the observed differences among people.

Our approach to the study of working memory involves developing computational models of a few different working memory tasks and comparing model predictions to performance data – at both the aggregate and individual levels. We see several advantages to our approach. First, we obtain predictions from a running computer simulation. This offers quantitative predictions along several dimensions (e.g., latencies, percent correct, patterns of errors). Second, we evaluate each model by comparing its predictions to aggregate and individual subject data. The latter is particularly important because of the systematic subject-to-subject variation in working memory capacity that has been observed (e.g., Cantor & Engle, 1993; Daneman & Carpenter, 1980; Waters & Caplan, 1996). Third, we model several tasks, all using the same theory of working memory developed within the ACT-R framework (Anderson, 1993). This is important for testing whether a single theory can cover the variety of tasks in which working memory plays a role. This final point reflects the fact that our work is embedded within a unified cognitive architecture (cf. Newell, 1982), as are some other working memory theories (Kieras, Meyer, Mueller, & Seymour, Chapter 6, this volume; Young & Lewis, Chapter 7, this volume).

In this chapter, we describe the basic features of the ACT-R theory and relate them to several working memory issues (the designated questions). We present an ACT-R model that predicts working memory results at the aggregate level and then go on to report our work in progress – developing ACT-R models of *individual's* working memory performance. We summarize some of the encouraging results we have obtained thus far and discuss their implications for future work.

The ACT-R Theory

The ACT-R theory of cognition specifies a fixed computational architecture that applies to all cognitive tasks. Within this architecture one can develop ACT-R models for different tasks. The main difference among ACT-R models is not in their way of processing information but in the initial knowledge with which they are endowed. This initial knowledge includes the facts and skills that are relevant to the task being modeled and that are presumed to be known by the subject population being studied. For example, a model of elementary school students solving arithmetic problems would only represent certain arithmetic facts (e.g., 3 + 4 = 7) and skills (e.g., how to carry a digit). Regardless of the content of this initial knowledge, ACT-R assumes the same performance and learning mechanisms for all tasks. Specifically, knowledge is always learned, deployed, interfered with, and decayed in the same way. These mechanisms are implemented in a simulation program that can be used to generate a set of theoretical predictions for a given task model. The interested reader is invited to visit the ACT-R home page (at http://act.psy.cmu.edu/) to learn more about developing ACT-R models. Before describing ACT-R's mechanisms, we provide an overview of how knowledge is represented in the system.

Symbolic Components of ACT-R

ACT-R is a hybrid system with symbolic and subsymbolic aspects. Knowledge is represented symbolically, whereas the processes acting on knowledge occur at a subsymbolic level. In this section, we describe the two symbolic representations in ACT-R. Declarative knowledge (for facts) is represented as a network of interconnected *nodes,* and procedural knowledge (for skills) is represented as a set of *productions*. Figure 5.1a depicts a node representing the fact 3 + 4 = 7. This node is labeled "Addition-fact$_i$" and is linked to

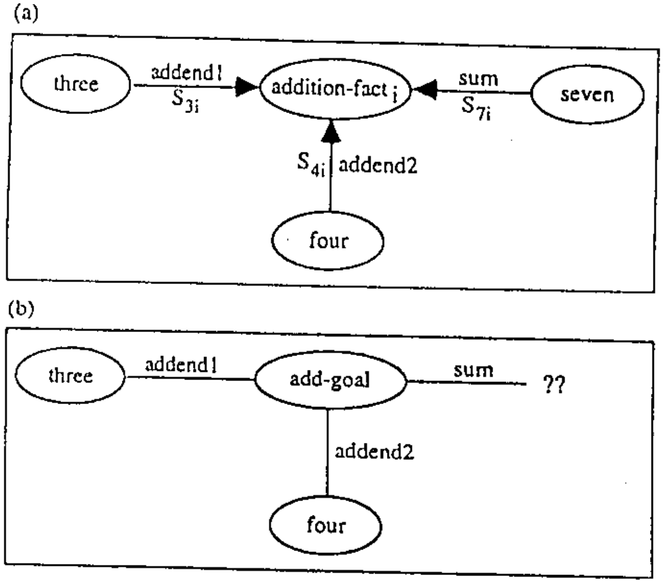

Figure 5.1. (a) A declarative node representing the fact "3 + 4 = 7." (b) A node representation of the goal to add 3 + 4.

"three," "four," and "seven," which comprise the fact's first addend, second addend, and sum, respectively. The links in Figure 5.1a vary in strength (S_{ji}), indicating the strength of the relationship between connected concepts – nodes that frequently occur together will have high associative strengths. Separate from declarative knowledge is the set of all productions held by the system. Each production represents a contingency for action. A production takes the form IF <condition>, THEN <action>. Here, <condition> specifies the circumstances under which the production is relevant, and <action> specifies a possible action to be taken. Table 5.1 presents a production for retrieving the sum in an arithmetic problem.

The processing of declarative and procedural knowledge in ACT-R is primarily driven by the *current goal* of the system (e.g., finding the sum of three plus four). The current goal contains the information in the focus of attention and uses a declarative node structure (Figure 5.1b). Its contents are either established by previous processing (e.g., when one part of a problem is solved, attention is switched to another part) or by stimuli in the environment (e.g., upon hearing a loud noise, the focus of attention may switch to process the sound). Figure 5.2 sketches the cycle of processing. First, the goal acts as a filter to select only those productions relevant to its current

Table 5.1. *Sample Production for Arithmetic Retrieval*

IF	the current goal is to add $a + b$ when no sum has been computed and an addition fact stating that the sum of $a + b$ is c can be retrieved
THEN	update current goal's sum to be c

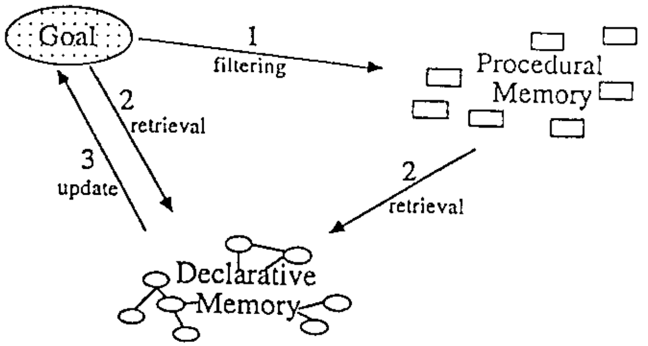

Figure 5.2. The role of the current goal in ACT-R's processing cycle: (1) It acts as a filter on procedural memory so that only goal-relevant productions are available. (2) Along with the selected production, the goal (via its links to declarative memory) makes some nodes more accessible than others. (3) Once a declarative node is retrieved, information from that node is used to update the current goal.

state. For example, the production in Table 5.1 is relevant to the goal in Figure 5.1b because both specify adding two numbers when no sum has been computed. Of the productions offering such a match to the current goal, the production with the highest expected utility (estimated from past use) is selected for continued processing. (See Lovett & Anderson, 1996, for more details on this selection mechanism and how it enables the successful modeling of various problem-solving data.) Second, the retrieval specified by the selected production is attempted. This retrieval is influenced by both the current goal (via its connections to declarative memory) and the selected production (via the retrieval pattern specified in its condition). In the addition example, an addition fact involving "three" and "four" is retrieved because the current goal includes "three" and "four" and the Retrieve-Sum production specifies retrieving an addition fact. Third, the contents of the retrieved node are used to update the current goal according to the production's action specification, for example, if "3 + 4 = 7" is retrieved, "seven" is added to the current goal's "sum" link. Then, the cycle of processing is reinitiated with the modified goal.

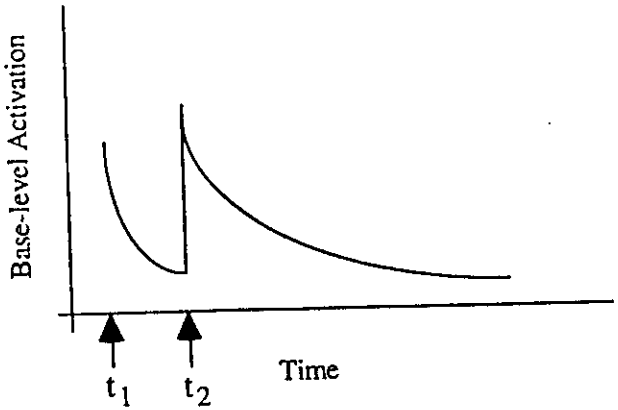

Figure 5.3. The time course of base-level activation for a node that is created at time t_1 and later accessed at time t_2 (see Equation 1).

Subsymbolic Processing in ACT-R

The above processing cycle describes performance at a symbolic level. Learning and other performance mechanisms, however, are defined in terms of processes at the subsymbolic level. Here, we focus on the subsymbolic mechanisms acting on declarative knowledge because they are most relevant to working memory.

For declarative knowledge, *activation* is the main unit of "currency" for learning and processing. Each node has a certain base-level activation that influences its accessibility. Specifically, when a fact is first learned (stored as a new node), it is endowed with an initial activation. Each time that fact is accessed (retrieved), it receives a boost to its base-level activation (learning). However, each of these "boosts" decreases as a power function of time (forgetting). Figure 5.3 shows how base-level activation changes with time for a node that was created at time t_1 and later accessed at time t_2. The ACT-R function for base-level activation of node *i* is

$$B_i = \log(\Sigma t_k^{-d}), \tag{1}$$

where t_k is the time lag since the kth access of node *i* and *d* is the decay rate. A node's base-level activation reflects that node's prior history of use. Because this prior history of use is a good index of future likelihood of use, the system is adaptive: Nodes that have high activation (high past use) are both more likely to be needed and more accessible.

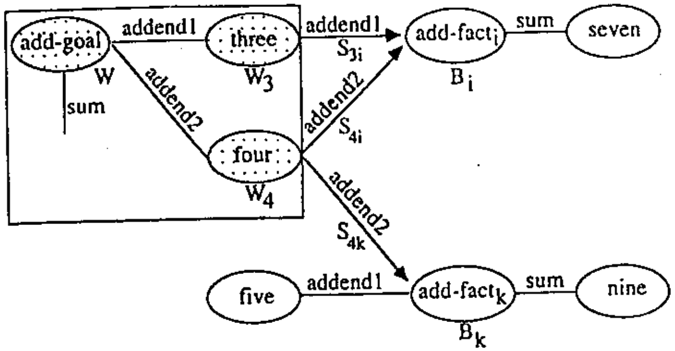

Figure 5.4. The goal (dotted) and declarative nodes for an arithmetic task (adding three plus four). The goal's source activation *W* is divided between the goal nodes "three" and "four" and then is spread along the arrows to addition facts in declarative memory.

Base-level activation represents a node's overall accessibility but does not account for any effects of the current context. In ACT-R, the current goal represents a person's focus of attention and thus drives context effects. The goal propagates "attentional" activation to declarative memory, raising the accessibility of some nodes relative to others. For example, Figure 5.4 shows the current goal (dotted) propagating activation along the arrow-links to related nodes in declarative memory. Specifically, the goal's attentional activation, or *source activation W*, is divided among the goal nodes "three" and "four," and then these shares (W_3 and W_4) are each spread among neighboring nodes in declarative memory. The amount of source activation received by neighbor node *i* from goal node *j* is the product of W_j (the source activation from goal node *j*) and S_{ji} (the strength of the link between goal node *j* and neighbor node *i*). A node that is connected to the goal by more than one link receives source activation along all those links. For instance, in Figure 5.4, "add-fact$_i$" receives $W_3 \cdot S_{3i} + W_4 \cdot S_{4i}$ units of source activation. The *total activation* of declarative node *i*, then, is the sum of its base-level activation and its received source activation:

$$A_i = B_i + \Sigma_j W_j \cdot S_{ji}. \tag{2}$$

The main implication of this equation is that both past use of a fact and its relevance to the current goal jointly determine that fact's accessibility. Familiar facts will be more accessible than unfamiliar facts because of their difference in use, and contextually relevant facts will be more accessible than irrelevant facts because of their difference in association to the current goal. These activation mechanisms have been used to model many memory experiments successfully (Anderson & Matessa, 1997).

In summary, ACT-R claims that when people focus their attention on the current goal, they are directing additional activation to declarative memory elements that are related to that goal. Source activation spreads from each goal node in proportion to the strength of association between that node and various declarative nodes. Declarative nodes also differ in terms of base-level activation, which is determined by their past history of use. One way of conceiving ACT-R's distinction between base-level activation and source activation is that the former represents the activation of each node without any context effects and the latter represents a kind of attentional activation that gets dynamically applied to particular nodes based on the current context.

Performance Functions for Declarative Retrieval

The latency T_i for retrieving node i is a function of total activation A_i. Specifically,

$$T_i = Fe^{-A_i}, \tag{3}$$

where F is a time-scaling factor. An important implication of Equation 3 is that there is a cost for accessing a declarative node; the lower the node's activation, the greater the time taken to retrieve it. To generate ACT-R's prediction for the time to take a certain observable step, we add all the *retrieval times* (Equation 3) and *action times* for the productions leading up to that step. (The default action time is 50 ms; productions that involve motor actions have longer action times.) Thus, ACT-R predicts longer latencies on average for steps requiring longer production sequences and/or the retrieval of less activated declarative nodes.

To handle nodes with very low activation, ACT-R posits a retrieval threshold below which nodes will not be retrieved. (One can think of this threshold as a fixed waiting time; if a node is not retrieved within that time, the retrieval attempt is aborted.) Of the nodes with above-threshold activation, it would be natural for the maximally activated node to be retrieved. However, ACT-R recognizes that there may be noise in the system, so noise is added to each node's total activation, and then the node with the highest "noisy" activation is selected. The following equation describes the behavior of a system with this noisy selection process:

$$P(\text{retrieve}_i) = e^{A_i/s}/\textstyle\sum_j e^{A_j/s} \tag{4}$$

where the denominator sums over all nodes competing to be retrieved and s represents the level of activation noise.[1] This equation states that the lower a node's total activation relative to competing nodes, the lower its chance of being retrieved.

[1] This noise is approximately Gaussian, centered at 0, and with spread parameter s. The retrieval threshold is included as one of the "competing nodes" in Equation 4 to reflect the fact that a node's "noisy" activation must be greater than its competitors' and the retrieval threshold in order to be retrieved.

The probability of declarative node i *not* being retrieved – an error of omission – is just the complement of Equation 4. Errors of commission (when an incorrect node is retrieved) are produced by a mechanism known as partial matching. This mechanism allows a node to be retrieved even if it does not exactly match the production's retrieval condition. Such a node competes based on a reduced total activation, A_i':

$$A_i' = A_i - M, \tag{5}$$

where M represents a mismatch penalty. This penalty is a measure of the "psychological distance" between node i and the node described in the production's retrieval condition.[2] It produces a bias against retrieving nontarget nodes. Nevertheless, with activation noise, nodes that are slightly dissimilar to the production's specification may be retrieved, especially if their total activation is high enough to compensate for the mismatch penalty.

In summary, both latency and probability of retrieval are nonlinear functions of a node's total activation (after noise and partial matching are taken into account). As we will show, this nonlinearity has an impact on how individual differences manifest themselves in our model predictions.

Working Memory in ACT-R

Basic Mechanisms and Representations in Working Memory

Before describing the mechanisms and representations of working memory in ACT-R, it is important to define working memory in terms of this theory. There are two ways to do so. One is to equate working memory with the *content* that is being maintained during processing (e.g., the elements representing the memory items in a working memory task). This content-oriented definition identifies working memory as a subset of the entire declarative memory. That is, working memory is not a special repository of information but just those declarative nodes that are highly activated because they have been stimulated from the environment and/or are strongly linked to the current goal (source activation). According to this definition, working memory mechanisms are just the mechanisms acting on declarative memory: learning, decay, and attentional activation.

The second way to define working memory emphasizes the *process* that enables memory elements to be concurrently maintained. This definition takes working memory as the propagation of source activation from the current goal. Defining ACT-R's working memory in these terms emphasizes the attentional activation mechanism that differentially activates items relevant to the current

[2] If declarative nodes were represented as feature vectors, the psychological distance between them would be a function of the overlap in their representations. Since declarative nodes are represented symbolically in ACT-R, this distance function is specified in the model description.

context. This process-oriented definition of working memory is complementary with the content-oriented one above: Nodes in the highly activated subset of declarative memory receive an important part of their activation from the process that spreads source activation. Note that other working memory researchers have similarly considered working memory as the union of content and process (e.g., Byrne, 1998; Cowan, Chapter 3, this volume).

Both definitions identify the basic mechanisms of working memory as the spreading of source activation (which primarily affects nodes strongly linked to the goal) and the general declarative mechanisms of base-level learning and decay (which affect all of declarative memory, of which working memory is a part). These mechanisms are always at work, influencing the accessibility of declarative knowledge and allowing for both context and learning/forgetting effects.[3]

With working memory defined, we can now address the issue of representation. Because working memory is a subset of all declarative memory, working memory representations are the same as those used for declarative nodes in general. Throughout declarative memory, however, different node structures (representations) are used for different kinds of information. For example, the arithmetic fact depicted in Figure 5.1a includes the features "addend1," "addend2," and "sum," whereas the memory-list item depicted later in Figure 5.8 incorporates the features "trial," "position," and "value." Each node incorporates its represented features in the linked structure. Similar items, such as addition facts "3 + 4 = 7" and "3 + 5 = 8," can thus be related in several ways: They can use the same node structure (both have links for addend1, addend2, and sum), they can have common elements (both include "three"), and they can have strong associative links ("3 + 4 = 7" and "3 + 5 = 8" may co-occur frequently). Note that it is also possible for different people (e.g., an expert vs. a novice) to represent the same object in different ways by encoding different features. Thus, a pair of nodes representing two items may be more or less similar for different people, both in terms of their representational structures and in terms of their strengths of association.

Such differences in representation influence processing in ACT-R. For example, the partial matching mechanism described above allows for interference among *similarly represented* items. Thus, confusions involving the retrieval of one node for another will tend to be limited to nodes of the same structure (e.g., misretrieving one arithmetic fact for another, or confusing one word for another). Moreover, the greater the feature overlap between two items, the more likely such confusions are to occur. That is, confusions between phonologically similar words are more likely than confusions between phonologically dissimilar words, and confusions between related arithmetic facts are more likely than confusions between unrelated facts. Thus, interference effects arise mainly because of representational similarity. This implies that in dual-task situations where the two tasks require processing of similarly represented items, interference effects can substantially impair performance. Such effects have been found in several studies (e.g., Baddeley & Lieberman, 1980; Logie, 1986; Shah & Miyake, 1996).

[3] The learning/forgetting of base-level activation (Equation 1) allows for a variety of memory effects that are not the focus of this chapter. For example, familiar (highly practiced) declarative nodes will show slower decay than new nodes. Such differences have been observed and modeled with separate decay-rate parameters by Healy, Fendrich, Cunningham, and Till (1987). In the system we are describing, such a familiarity effect arises naturally from a single decay parameter: The more practiced an item, the more activation boosts that are combined to produce the item's total activation. Adding together more power-decay functions (each with a common decay rate) produces a combined function with slower decay.

The Nature of Working Memory Limitations

In ACT-R, working memory limitations are invoked by the constraint that source activation is limited and must be divided among the goal nodes, that is,

$$\Sigma_j W_j = W. \tag{6}$$

Here, W represents the total amount of attention focused on the current goal and W_j the share of source activation propagating from goal node j. Equation 6 implies that, for a fixed W, less source activation will be spread from each goal node the more goal nodes there are. Hence, in complex tasks (more goal nodes) there will be a smaller modulating effect of source activation on the accessibility of related declarative nodes.[4]

The share of source activation that is propagated from each goal node is spread among the different links emanating from that goal node. For example, in Figure 5.4, the source activation W_4 is spread to both add-fact$_j$ and add-fact$_k$. (In a complete model, there would be many more links.) Note that the link strengths S_{ji} are learned values that reflect the co-occurrence between nodes. As a default, the learned link strength S_{ji} is approximated by the expression $C - \log(n_j)$, with n_j as the number of nodes linked to node j. Thus, the more links emanating from a given goal node, the less source activation that is propagated along each link. Hence, for more memory-intensive tasks (i.e., more elements connected to each goal node), a smaller amount of source activation reaches any one linked node.

The key implication of a *limit* to source activation (W) is that the size of its modulating effect will be reduced in cases when source activation is divided among more goal nodes and spread among more links to declarative memory. Since relative amounts of total activation determine the latency and probability of retrievals (Equations 3 and 4), this reduction in the effect of source activation leads to degraded memory performance. Therefore, retrievals in

[4] For simplicity, we take source activation as divided evenly among the n nodes in the current goal: $W_j = \frac{W}{n}$. Thus, in Figure 5.4, $W_3 = W_4 = \frac{W}{2}$. Of course, an unequal division of W is possible.

complex, memory-loaded tasks will be less accurate and take longer than will retrievals in simpler tasks. This effect holds regardless of whether the memory load is part of a memory span task or part of a dual-task situation. We will show examples of each of these below.

In summary, ACT-R posits a limit in attentional resources (or source activation W) that leads to degraded memory retrieval in complex and memory-loaded tasks. The attentional limit is *not* a cap on total activation in the system but a cap on a particular kind of dynamic, attentional activation spreading from the goal. This attentional limit produces a limit in the degree to which goal-relevant items can be *differentially* activated. Such goal-based modulation of processing is similar to that proposed by O'Reilly, Braver, and Cohen (Chapter 11, this volume). Moreover, because this limitation affects the degree to which memory items can be differentially activated, it posits that *relative* activation levels are more important than absolute activation levels; thus, it is compatible with theories that attribute working memory limitations to inhibitory processes (e.g., Conway & Engle, 1994; Kane, Hasher, Stoltzfus, Zacks, & Connelly, 1994; Stoltzfus, Hasher, Zacks, Ulivi, & Goldstein, 1993; Zacks & Hasher, 1994).

The Control and Regulation of Working Memory

The primary mechanisms controlling working memory have already been delineated: (a) spreading source activation and (b) learning and decaying base-level activations. These mechanisms combine to produce the total activation levels of all declarative nodes (including those highly activated nodes that comprise working memory). Total activation, in turn, affects the accessibility of declarative nodes (Equations 3 and 4).

The main regulatory processes among these are the limit to source activation (W) and the decay of base-level activation. The limit on W constrains the degree to which goal-relevant nodes can be differentially activated relative to other nodes. Because this limitation specifies a fixed amount of source activation that must be shared and spread, there is a gradual degradation in memory performance as task complexity and memory demand increase (i.e., more complex tasks lead to thinner spreading of source activation, reducing its benefits). This implies that only a limited number of declarative nodes can be effectively differentiated (based on added source activation) from the remaining elements in declarative memory.

The decay of base-level activation, on the other hand, does not reflect a limitation in the amount of activation but rather in the duration of activation. This mechanism specifies that declarative nodes' activation (and hence general accessibility) will decrease as a power function of the time since they have been used (Equation 1). Therefore, nodes will tend to remain above threshold for only a limited amount of time. Like the spreading of source activation, this decay mechanism helps to keep working memory to a relatively small size.

Given these regulatory processes, there are two ways a node can be maintained in a highly activated state: (a) It can be part of the current goal and thus be in the focus of attention (e.g., when the concept "three" is activated by seeing a "3" in the environment),[5] (b) it can be strongly connected to one or more goal nodes and thus receive a large amount of source activation (e.g., when the fact "3 + 4 = 7" in memory is highly relevant to the goal).

The Relationship of Working Memory to Attention and Consciousness

Attention and our notion of working memory are closely related. As we have described above, working memory processing is heavily influenced by source activation, the "attentional energy" that is directed at the current goal and spread from the goal nodes. We view the limitations of working memory as stemming from limited attentional resources (Equation 6). Several of the tasks that we discuss and model in this chapter involve dual-task procedures in which attention must be divided among concurrent goals (e.g., reading and remembering). In our models of these dual-task situations, source activation is spread more thinly among a greater number of goal nodes – essentially dividing the attentional resource W between the two goals – leading to worse performance than under single-task conditions.

In lay terms, "attention" often refers to a dimension of motivation or alertness, suggesting that people can direct more or less of their attention to the current task. We have not modeled this kind of attention directing, but it could be modeled by varying the total amount of source activation or how it is shared among different components of the current goal. This is an interesting issue to explore in connection with how people strategically (vs. automatically) allocate attentional resources.

Relating working memory to consciousness is difficult. The most natural link under our framework is to consider those nodes with above-threshold activation as accessible to conscious awareness. Declarative nodes below this threshold still vary in total activation; this affects the processing required to bring them into awareness (cf. Reder & Gordon, 1997). For example, one's own name is so highly practiced, very little environmental stimulation can bring it into awareness (e.g., cocktail party effect). It is also possible to refer to our second definition of working memory – the spreading of source activation from the goal – and link consciousness to the ability to maintain a focus of attention that influences the processing of information.

[5] Note that in the current system, the current goal does not undergo decay. Thus, it is difficult to explain forgetting of the current goal in ACT-R. See the section on Biological Implementation for a generalization of ACT-R that handles this problem and makes connections to neural activation data.

The Unitary Versus Non-Unitary Nature of Working Memory

As implied above, the value of *W* impacts processing (i.e., retrieval probabilities and latencies) of all kinds of information. It represents a general resource that is used in any task that involves retrieval from declarative memory. According to ACT-R, then, higher *W* will lead to better performance in any task, all else being equal. This leads to one of the hypotheses explored in our research: that the performance of an individual on one task is predictive of his or her performance on another task because both tasks tap into the same attentional resource *W*. This unitary hypothesis is similar to Engle, Kane, and Tuholski's (see Chapter 4, this volume); indeed, our model offers an information-processing account consistent with their analysis that supports a latent variable underlying working memory performance.

It is important to note, however, that all else is not always equal. For example, different situations may involve processing different kinds of declarative nodes (e.g., those representing spatial versus verbal information) and the interference among nodes across these situations may differ (see Basic Mechanisms section). Also, various tasks and subject populations may differ in terms of the base-level activation of the nodes involved, which can make *W*'s effect on performance seem variable across tasks. For example, the latency of retrieval (Equation 3) is a nonlinear function of the sum of base-level and source activation: When base-level activations are high, the modulating effect of *W* is small and when base-level activations are low, *W*'s effect is large. Finally, in many situations, participants' strategies for approaching a task may differ. These differences can affect the number and timing of retrievals and mask a common effect of *W*, especially if people choose strategies that compensate for their working memory capacity. This discussion highlights several of the sources of variability we try to reduce to study true differences in the limit to source activation. However, in most tasks, it is likely that both kinds of variability (domain-general and -specific) influence working memory processes (cf. O'Reilly et al., Chapter 11, this volume).

It is interesting to contrast this view with other views that posit separate working memory capacities for different types of processes. For example, Shah and Miyake (1996) argue for separate verbal and spatial working memory capacities. Although it would be difficult to discriminate between models that specify unitary versus separate working memory pools, we prefer to posit a unitary working memory and ascribe individual differences in performance patterns across tasks to (a) different patterns of experience and (b) differences in buffer use. As discussed earlier, differences in the amount of practice at various types of tasks (strength of productions), different strategies (sets of productions), and differences in knowledge representations (declarative nodes) may result in performance differences that vary with the type of task. These differences could produce performance differences across tasks even under a single working memory capacity.

Others have argued that more interference occurs within the same modality for dual tasks (Baddeley & Lieberman, 1980; Brooks, 1968; Logie, 1995) because of sharing of an input or processing buffer. Differences in modality-specific buffers may also account for some of the observed difference among individuals that has been taken as evidence for separate spatial and verbal working memory pools.

The Relationship of Working Memory to Long-Term Memory and Knowledge

As we noted before, working memory is not a special repository of information but rather a subset of declarative memory that is distinguished by higher activation levels. Otherwise, working memory elements are processed (e.g., they undergo learning and forgetting) just like any other node.

It is worth relating our view of working memory to the approach of Ericsson and Kintsch (1995) and Ericsson and Delaney (Chapter 8, this volume). We believe that their findings and our theoretical approach are compatible because learning processes play a large role in both. Their subject SF was trained to recall digit strings of over 80 digits. This superior digit span was achieved with a lot of training on recalling digits, practice at converting digits into running times (e.g., "826 is 8.26 seconds, an excellent 2 mile time"), and knowing ahead of time the length of the digit string so that the appropriate tree structure could be "loaded." Each of these processes was presumably learned and refined over SF's long training. For example, SF clearly had specialized, highly practiced procedures for recoding digits into running times (facilitated by the fact that he was an accomplished long-distance runner), special tree structures on which to hang these running times, and so on (see Chase & Ericsson, 1981, 1982 for more discussion). In contrast, SF was not able to recall letter strings that were at all exceptional in length. This speaks to the specificity of what SF learned and argues that his *W* was not altered, only his procedures and declarative knowledge.

The Biological Implementation of Working Memory

The question of how working memory is implemented in the brain is a difficult one: It requires bridging the gap between the abstract theoretical construct of working memory and the complex biological processes of the brain – neither of which is completely understood. Researchers have used three different approaches to attack this problem: (a) consider their working memory model analogous to the human brain and compare performance of both systems after specific impairments, (b) analyze the predictions of their model and look for these features in brain-imaging data, and (c) extend their model to refer to biological mechanisms and explore the model's performance. The first two approaches apply to the theory described in this chapter, so we will discuss them below. The third approach, exemplified by the work of O'Reilly et al. (Chapter 11, this volume), does not apply directly because ACT-R's processes have traditionally been described at a functional level rather than a biological one. (However, see Lebiere & Anderson, 1993, for a connectionist implementation of ACT-R.)

The first approach – comparing model performance to human performance when both have undergone some impairment – is relatively common in computational modeling (e.g., Plaut, 1996). This approach was applied to four ACT-R models developed by Kimberg and Farah (1993). The four models simulated the Stroop task, the Wisconsin Card Sorting Test, a context memory task, and a motor-sequencing task. These tasks were chosen for their variety and for the patterns of errors they evoke among frontal-lobe-damaged patients. Each model included the productions and interconnected declarative nodes required for performing the task. In its unimpaired form, each model performed with a high level of accuracy. However, when the strengths of the links between declarative nodes were reduced (representing relatively dispersed neurological damage), each model showed a level and kind of impairment akin to that exhibited among frontal-lobe-damaged patients. For example, the impaired motor-sequencing model performed inappropriate actions on each device (e.g., pulling instead of twisting a twister), and the impaired Stroop model exhibited greater interference from unattended stimulus features (it was slower to name the ink color of color-name words).

The impairment procedure applied to these models (weakening strengths of association) is closely related to reducing W, the source activation from the goal. Both "impairments" reduce the amount of source activation arriving at to-be-retrieved declarative nodes, thereby affecting retrieval performance. Since impairing the spreading activation process in a computational model produces effects similar to frontal lobe damage, the logic of this approach suggests that spreading source activation may have its counterpart in the human brain. It also suggests that working memory processes like spreading source activation are associated with frontal lobe function. Although this model-to-brain mapping is at an abstract level, it is consistent with the role of prefrontal cortex posited in the O'Reilly et al. model (Chapter 11, this volume).

The second approach to understanding the biological implementation of working memory involves comparing the predictions of a working memory model with data recorded from the brain. This approach does not speak directly to the mechanisms underlying working memory but can shed light on whether a proposed mechanism is consistent with observed changes in neural activation. Relating such data to models of working memory is a more recent approach (Cohen et al., 1994; Just, Carpenter, Keller, Eddy, & Thulborn, 1996; Lü, Williamson, & Kaufman, 1992a). For example, Lü, Williamson, and Kaufman have studied the habituation responses of various cortical (and subcortical) regions of the brain using event-related potential

and magnetic source imaging techniques. One of their findings relevant to our work involves the response of two areas in the auditory cortex. In particular, they found that the event-related field response (100-ms component) for repeated tones shows an activation trace that decays with time after the presentation of the tone. They postulate that this trace (measured in terms of reduced responsiveness at increasingly short ISIs) reflects the neural availability of the tone for processing by working memory. Moreover, they have shown that the rate of decay varies considerably across individuals and relates to behavioral data. The shape of the decay they observed is well captured by an exponential function

$$e^{-(t-t_0)/\tau}, \tag{7}$$

where t is the current time, t_0 is the time of the tone's onset, and τ is the decay rate.

This neural activation function seems superficially similar to the ACT-R declarative learning mechanism, where nodes' base-level activations consist of accumulated power-decay functions. However, such a mapping suffers from the fact that the time course of decay studied by Lü et al. (1992a, b) was on a much shorter time scale (and fit by an exponential function). The more appropriate mapping we have found relates the amount of source activation W to the decay rate τ, assuming a slight extension to ACT-R's treatment of source activation. (See Appendix A.) The basic idea involves viewing the goal's source activation as a "leaky capacitor" (Sejnowsky, 1981) that needs to be continually pumped with source activation to maintain a fixed level W. That is, instead of W being automatically and instantaneously transferred to a new goal, source activation of the previous goal would gradually decline (à la Equation 7), and source activation of the new goal would gradually build to asymptote. In such a system, the total source activation across all goals (past and present) is a limited amount W that relates directly to the decay rate τ from Equation 7. It is noteworthy that Lü et al. (1992b) found individual differences in the decay rate τ for different subjects, just as we propose that differences in W will predict working memory differences among subjects.

This new view of W allocation helps explain why the brain might employ a limit to attentional resources such as attention or source activation. If attention could be allocated without limit (or, under our new generalization, if source activation did not decay), the system would not be able to focus differentially on the current goal. Not only would this make it difficult to distinguish the current goal from past goals, but goal-based modulation of declarative information would become ineffective with so many goals simultaneously activating so many declarative memory elements. This analysis suggests a certain computational efficiency to maintaining limited attentional resources and is similar to the analysis of Young and Lewis (Chapter 7, this volume) supporting functional limitations to working memory.

Table 5.2. *Critical Productions for Memory-Loaded Algebra Task*

Substitute for a:
IF the goal is to solve an equation with *a* in it
* and *f* is the first element of the memory list
THEN substitute *f* for *a*
Substitute for b:
IF the goal is to solve an equation with *b* in it
* and *s* is the second element of the memory list
THEN substitute *s* for *b*
Invert-transformation:
IF the goal is to solve an equation *term1 op1 constant1 = constant2*
** and *op2* inverts *op1*
THEN transform equation to the form *term1 = constant2 op2 constant1*
Collect-sum:
IF the goal is to solve an equation that contains *c + d*
** and *s* is the sum of *c + d*
THEN replace *c + d* by *s*

The Role of Working Memory in Complex Cognitive Activities

As mentioned above, ACT-R posits that memory performance is degraded in more complex tasks because the limited amount of source activation (*W*) must be shared among more elements. In this section, we exemplify this prediction in the context of a model of algebra problem solving. (See Anderson et al., 1996, for details.) This particular algebra task emphasizes working memory by incorporating a memory-load component. Specifically, in each trial the participant had to encode a digit list, hold the list in memory while solving an algebra problem, and then recall the digit list. The difficulty of both subtasks was manipulated over trials. For example, the digit span was either 2, 4, or 6 digits, and the algebra problems required either one or two transformations. Finally, the trials were divided into two types: those for which the algebra equation had all numeric constants (e.g., $3x-8=7$) and those for which the equation required substituting digits from the digit span (e.g., $ax-b=7$, where subjects had to substitute the first and second elements of the digit span for *a* and *b*, respectively, and then solve). The latter trial type was included to study working memory effects when the two subtasks required integrated processing.

Table 5.2 shows some production rules used in the model for this task. Notice that each asterisk (*) indicates retrieval from the memory list, and each double asterisk (**) indicates retrieval of an arithmetic/algebra fact. The probability and latency of these retrievals depends on the amount of source activation reaching the to-be-retrieved nodes. Figures 5.5a and 5.5b show that these amounts will differ between high-load and low-load trials. In Figure

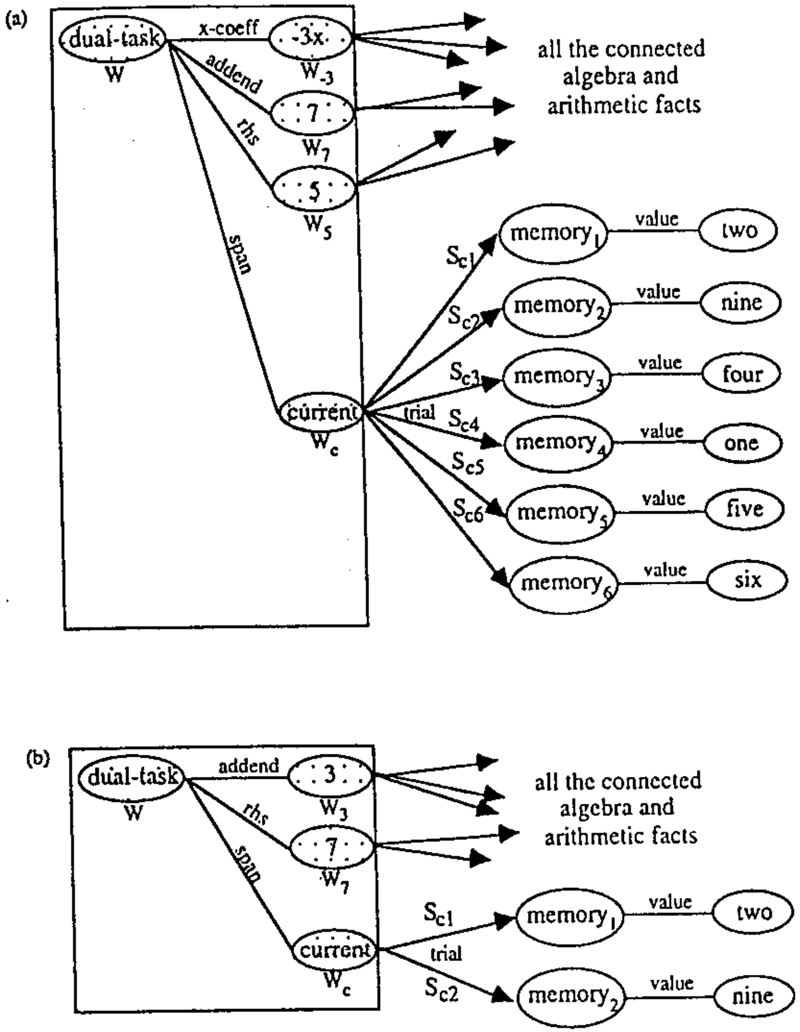

Figure 5.5. Goal and declarative memory for the algebra-memory dual task. (a) Representation of a trial with high-complexity algebra problem and high memory-load digit span. (b) Representation of a trial with low-complexity algebra problem and low memory-load digit span.

5.5a, very little source activation is reaching the relevant nodes in declarative memory. This is because there are many goal nodes (less source activation per goal node) and there are many links from these goal nodes to declarative memory (thinner spreading of each share of source activation). In contrast, in Figure 5.5b, fewer goal nodes are needed to represent the simpler algebra problem (more source activation per goal node) and there are fewer links from these nodes (better spreading of each share of source activation).

Because of these differences in source activation, the to-be-retrieved nodes will have less total activation on high-load trials than on low-load trials, which should lead to differences in probability and latency of recall. Notice that this prediction implies that manipulating the difficulty of one task should affect performance on that task as well as on the other task. For example, making the algebra equation more complex leads to smaller shares of source activation per goal node, thus leading to less source activation spreading to arithmetic facts and to memory items. Less source activation will make both arithmetic retrievals and memory span retrievals more error prone and slower. Moreover, making the algebra equation more complex increases the number of steps required to reach a solution. With more steps and slower retrievals, there will be a longer delay on high-load trials before the memory span can be recalled. That extra delay incurs greater decay for the memory digits, making their retrieval even more difficult. This shows that, given limited source activation, making one task more difficult can impair performance on both tasks.

Figure 5.6 shows some of the quantitative predictions of the model along with the corresponding observed results. Notice that both the model and participants take longer to solve the algebra problems when they are more complex, when they require substitution, and when the memory list is longer. In terms of string recall, accuracy suffers the longer the list and when the algebra problems are complex, but there is no reliable effect of substitution on memory.[6]

To generate predictions from this model, five free parameters were estimated. For the experiment shown in Figure 5.6, there were 48 data points in total to be fit, and the obtained goodness-of-fit statistic was $\chi^2(df = 43) = 79.3$. This value indicates a good fit (as can be seen), but there is significant residual variance not predicted. One possible source of that residual variance is unaccounted-for variation among participants in their working memory capacities. In the next section, we show how individual differences can be incorporated into the modeling framework described thus far and how doing so can improve overall model fits.

[6] The lack of an effect of substitution on memory span is probably due to the fact that the longer delay from making the substitution counteracts the extra practice of the first two digits when they are accessed for substitution.

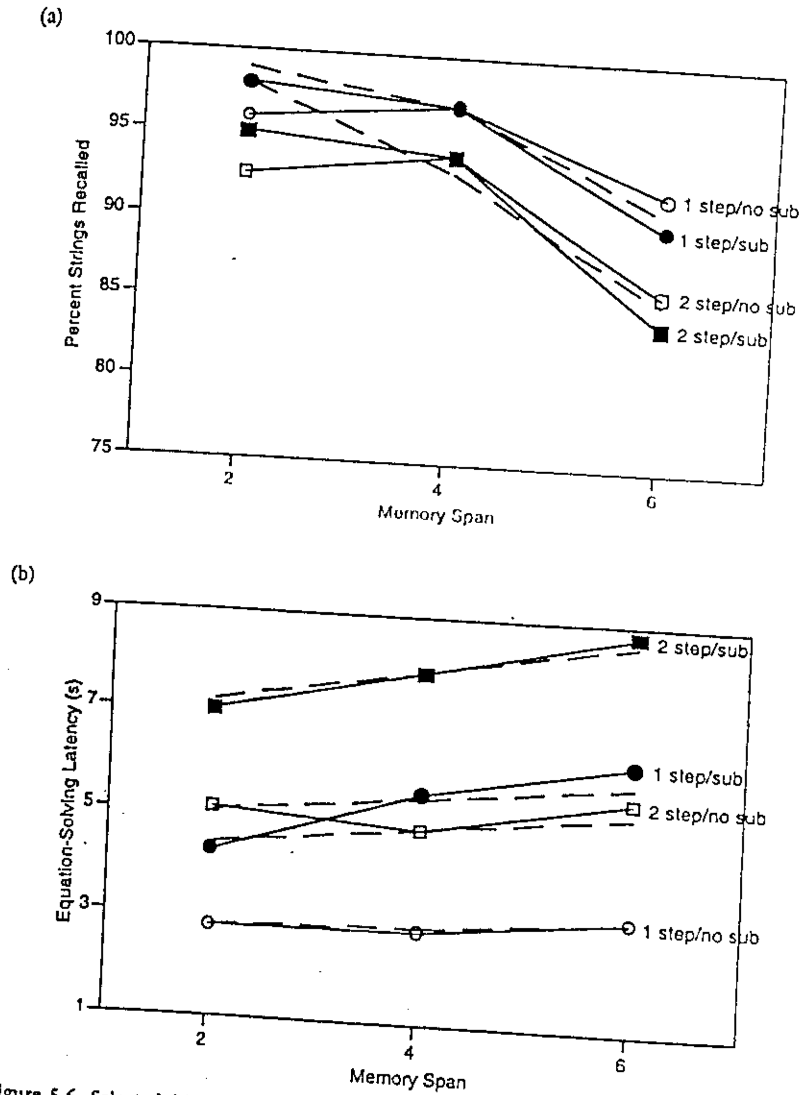

Figure 5.6. Selected (a) accuracy and (b) latency results, adapted with permission of Academic Press from Anderson, Reder, and Lebiere (1996). Solid lines are behavioral data, dashed lines are model predictions.

Modeling Working Memory Effects at the Individual Level

In this section, we describe some of the issues involved in modeling working memory effects at the individual subject level. We take *W* (the amount of source activation propagating from the goal) as the parameter reflecting individual differences in working memory. That is, in contrast to previous work (e.g., Anderson et al., 1996), which took *W* as a fixed limitation across subjects ($W \equiv 1$), we take *W* as a limitation that varies across the population ($W \sim \text{Normal}(1,\sigma^2)$). The results we present here are preliminary in nature but are encouraging with respect to our main hypothesis that varying the *W* parameter can capture individual differences in working memory performance. Our secondary hypothesis, based on the unitary nature of *W* in ACT-R, is that individual differences reflected in the *W* parameter will be similar across tasks and situations within the same individual. That is, we take an individual's *W* parameter to be relatively stable across time and tasks.[7] The topics discussed subsequently cover several issues associated with these hypotheses: theoretical issues (What are ACT-R's predictions under different values of *W*?), empirical issues (How can current research paradigms be adjusted to focus on individual differences caused by different working memory capacities and not by different strategies?), and modeling issues (How does variability in *W* impact model predictions?).

Theoretical Issues: Capturing Individual Differences in Working Memory

Just as varying the memory load of a task distributes source activation more or less thinly and leads ACT-R to predict differences in performance across task versions, varying the limit on source activation across individuals (for a fixed task) leads ACT-R to predict differences in performance across individuals. As described above, the more complex the task, the less source activation added to goal-relevant nodes, making them less likely to be retrieved. The same effect is obtained by keeping task complexity constant and decreasing the total amount of source activation (*W*) propagating from the goal. Thus, ACT-R predicts that people with lower attentional capacity *W* will have lower probability of retrieving goal-relevant nodes, all else equal. The value of *W* also has an effect on retrievals through the latency function (Equation 3). Here, when *W* is lower, goal-relevant nodes have less total activation, making their retrieval latencies longer. Longer latencies can have a secondary effect on other retrievals – they incur delayed processing of other information or less time for rehearsals, so other nodes' base-level activations will tend to be lower when their retrieval is attempted. These effects can accumulate across a task to produce markedly different behavior among subjects, as we will show.

It is noteworthy that our explanation of individual differences in working memory, though fundamentally based on the *W* parameter, also relies critically on the timing differences caused by differences in *W*. Thus, one view of our approach is that it provides a computational account of the processing-speed theory of age-related working memory differences (Salthouse, 1996). If the amount of attentional energy, *W*, decreases with age, then all retrieval latencies will be slower (not to mention somewhat more error prone) in older populations, making overall processing times slower. A related computational account of processing-speed theory by Byrne (1998) directly manipulates a rate parameter to produce differences between young and old populations. Our model, on the other hand, manipulates the parameter *W* thus indirectly affecting processing rate.

Other research suggests that decrements in performance among elderly subpopulations may be due to decrements in working memory capacity that are linked to inhibitory processes (Connelly & Hasher, 1993; Hasher & Zacks, 1988; Kane et al., 1994; Stoltzfus et al., 1993). Many of the effects seen in the elderly, especially under dual-task conditions, are reminiscent of problems associated with frontal-lobe-damaged patients. Postulating a lower value of *W* for individuals in either population might explain how these clinical behaviors arise. For example, smaller *W* can produce distractibility or inability to stay on task because less source activation propagates to goal-relevant items, making it easier for goal-irrelevant stimuli to capture attention.

Empirical Issues: Studying Individual Differences in Working Memory

The first experiment we report was designed to explore whether we could observe and model individual differences – subject-by-subject differences – in a dual-task memory experiment. We were specifically interested in capturing individual differences in working memory capacity. This meant that we had to be careful to design our experiment so as to minimize the introduction of other sources of individual differences. For instance, we sought to reduce between-subjects variability in motivation and on-taskedness by having an experimenter in the room to monitor each participant's progress. The main type of variability we sought to reduce, however, was *strategic variability*. We focused on designing the experiment so that subjects would tend to use the same strategy to perform the task. Building a model that matched this strategy would make the model more accurate and hence its parameter values (for both global and individual subject parameters) more interpretable.

The task we devised for this purpose is a variant of the digit working memory task developed by Oakhill and her colleagues (e.g., Yuill, Oakhill, & Parkin, 1989). It employs a dual-task procedure in that subjects had to read a sequence of digits aloud while maintaining in memory a selected subset of

[7] There may be moment-to-moment fluctuations of *W* within an individual (see section on Attention and Consciousness), but we are hypothesizing that the differences in *W* between individuals are potentially greater.

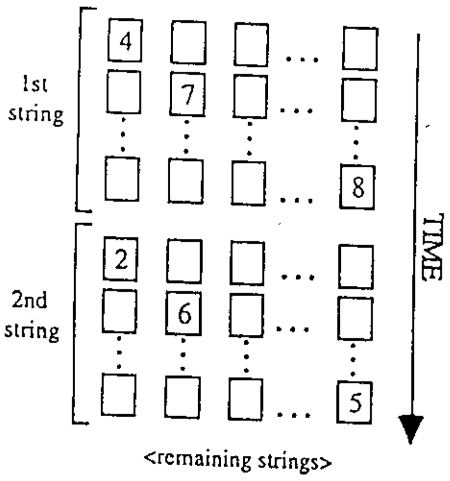

Figure 5.7. Time-based display of a trial in the digit working memory task.

those digits. Figure 5.7 shows the time-stepped presentation of a single trial. Digits were presented individually, appearing in one box and then disappearing before the next digit appeared in the next box. Subjects had to keep pace with the presentation rate by reading the digits aloud. (We used two presentation rates, 0.5 s and 0.7 s.) The rightmost digit in each string (8 and 5 in the figure) was to be remembered for later recall. These to-be-remembered digits were presented for double the presentation rate to allow for extra "memorizing" time; thus, the slower presentation rate offered more end-of-string time. After all the strings for a given trial were presented, subjects were prompted to recall the rightmost digits in the order they were presented. Recalling these string-final digits is analogous to recalling the sentence-final words in the Reading Span Task (Daneman & Carpenter, 1980).

This digit working memory task is distinguished from related working memory tasks in several ways. First, we maintained a precise digit-presentation rate via computer presentation. This reduced the variability from subjects choosing their own rates. Second, because our presentation rates were quite fast, variability owing to different rehearsal strategies was reduced – there was little time for any kind of rehearsal. Third, we varied the presentation rate to study its impact on memory performance. Note that a slower presentation rate increases the difficulty of the memory task by elongating the delay between storing and recalling the memory digits, but it leaves more time for additional processing of the memory digits. Fourth, we presented the different trial types in a random order. This eases the assumption that subjects come to each trial with an equal allocation of resources and eliminates potentially confounding time-based effects (e.g., learning, strategy change, boredom, fatigue). Finally, we included strict recall instructions for our task: The goal was to recall both the identity and position of each memory digit. Specifically, subjects' recall had to proceed once through the memory list without corrections or backtracking but with the possibility of skipping unknown digits. This procedure eliminated variability in recall order and reduced potential variability in recall strategies.

This task offers several options for manipulating task difficulty. A few that we have explored are (a) number of strings per trial (i.e., number of to-be-recalled digits), (b) number of digits to be read per string, and (c) interdigit presentation rate. Finally, to verify the similarity of different subjects' strategies in approaching this task, at the end of the experiment, we asked subjects to describe their approach to the task. A commonly reported strategy involved rehearsing memory digits at the end of each string after encoding the current to-be-remembered digit. We have incorporated this information into our model. More importantly for our study of individual differences, the frequency of this response suggests that subjects did not differ greatly in the strategies they applied to the task. Also, because of the simple nature of the items to be remembered and the fact that chunking strategies (i.e., encoding the digits into related groups) were so rare, it is reasonable to assume that undergraduate students would not differ greatly in their representations or their relevant knowledge about numbers. These arguments suggest that any memory differences to be found among subjects in this task would largely be a result of "architectural" differences in working memory capacity. At the very least, our procedure reduced other sources of variability more than is typically done.

Modeling Issues I: The Basic Model and Its Aggregate Predictions

The processes required to perform this task involve reading, storing, and recalling digits. Based on subjects' reports, it also appeared that people rehearsed digits (as time allowed) at the end of each string. We designed our model of the task to reflect all of these processes. Our model represents the two main goals of this task separately: reading digits and recalling digits. This corresponds to people switching between the goals as they process a single trial. Figure 5.8a depicts a goal to recall the digit in the first position of the current trial. Note that the representation of the corresponding memory digit (Figure 5.8b) is essentially the same; the only difference is that the value of the stored digit is included in the node structure.

These declarative structures are processed by productions that represent the different actions subjects perform in completing this task. Table 5.3 presents a list of some of the processes implemented by separate productions in our model. The "read" production applies whenever the goal is to read the digits and a digit is on the screen. After a digit has been read, if it is in the last position, the "store" production applies; this creates a new declarative node

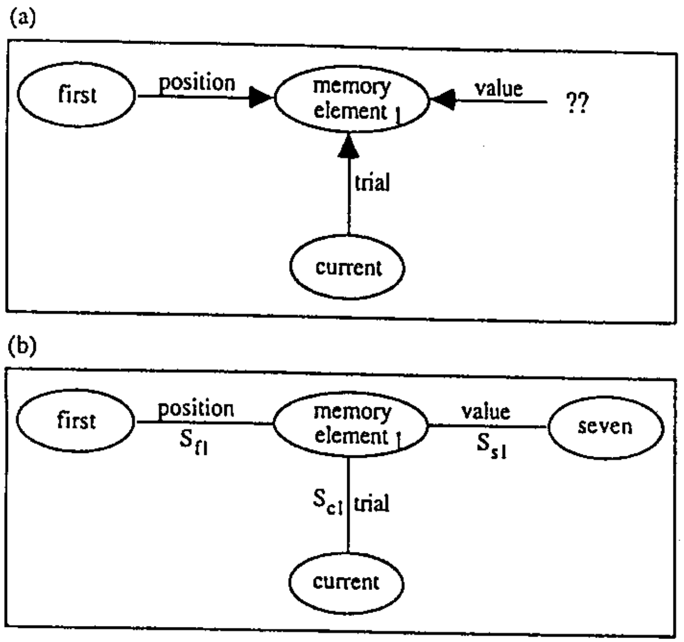

Figure 5.8. (a) A goal to recall the first digit in the digit working memory task. (b) The memory element for the first digit in the digit-working-memory task.

and endows that node with an initial boost of base-level activation. The "store" production also sets a subgoal to rehearse previous memory elements after the current digit is stored. Note that in our model, the production implementing rehearsal attempts to retrieve a particular memory digit. If successful, this retrieval gives a boost to the node's base-level activation, making the node more accessible subsequently (e.g., even easier to rehearse later). The probability of successful retrieval, however, depends on total activation, that is, if the target memory digit is highly activated relative to the others, it will likely be retrieved. Notice that a high value of W would produce high activation for the to-be-rehearsed digit, making rehearsals more likely to succeed and faster. The more successful rehearsals that fit into the time allowed at the end of each string, the greater total activations will be when the "recall" production attempts to retrieve the memory digits. This explains how different values of W can have impact on the amount of learning that goes on within a single trial.

Figure 5.9 shows that different values of W can also directly affect recall of the memory digits at the end of each trial. Here, the goal is to recall a digit for the first position of the current trial. The node representing this target mem-

Table 5.3. *Critical Productions for Digit Working Memory Task*

Read:
IF goal is to read a digit and digit d is on screen
THEN say digit d
Store:
IF goal is to read a digit and digit d is on screen and d is in last column and d has been read
THEN store d and prepare to rehearse
Rehearse:
IF goal is to read a digit and digit d is in the position to be rehearsed
THEN update position to be rehearsed
Recall:
IF goal is to recall digit in position p of trial t and digit d "matches" but has not been recalled
THEN say digit d

ory (memory element$_1$) receives source activation from both goal nodes "first" and "current," whereas other memory digits receive source activation only from the "current" goal node. This difference means that memory element$_1$ is receiving an additional $W_j{\cdot}S_{ji}$ units of source activation.[8] When either W_j or S_{ji} is smaller, this difference in activation will be reduced, and the relative ease and speed of retrieving the target memory element will be reduced, leading to poorer recall. As we described earlier, the link strength S_{ji} is smaller the more nodes linked to goal node j; thus, when the memory list is longer, the S_{ji}'s will be smaller and performance will be worse. Moreover, source activation W_j may differ among subjects. Therefore, subjects with a lower value of W will have a lower W_j, and hence show poorer recall performance. This effect will be even more pronounced on long lists, where the differential activation of target memory digits is the product of both a reduced W_j and a reduced S_{ji}.

To compute the model's predictions, we ran 22 simulations (one for each subject) under each of the task conditions used. These task conditions include all 16 possible combinations of the following factors: number of digits to be

[8] The activation of memory element$_1$ is also greater than that of the other memory elements because it exactly matches the current production's retrieval template (which involves the *first* element of the *current* trial). The other memory elements only partially match this template (they are not in the first position), so they compete for retrieval with an activation level reduced by the partial matching penalty (Equation 5). Because the mismatch penalty is smaller the greater the similarity between a candidate node and the retrieval template, this model produces similarity-based errors where similarity is a function of position in the recall string. Although we do not present them here, the pattern of positional errors produced by the model is very similar to that exhibited by subjects.

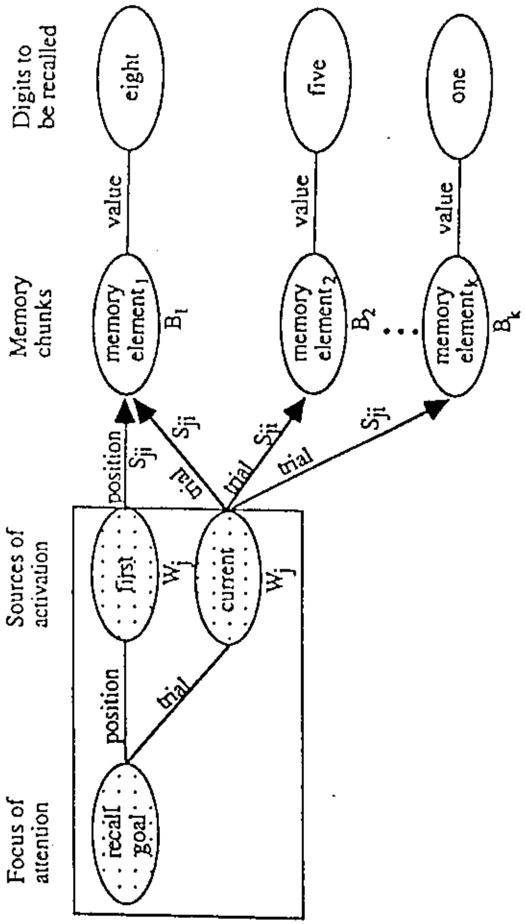

Figure 5.9. Goal and declarative memory representation for the digit working memory task.

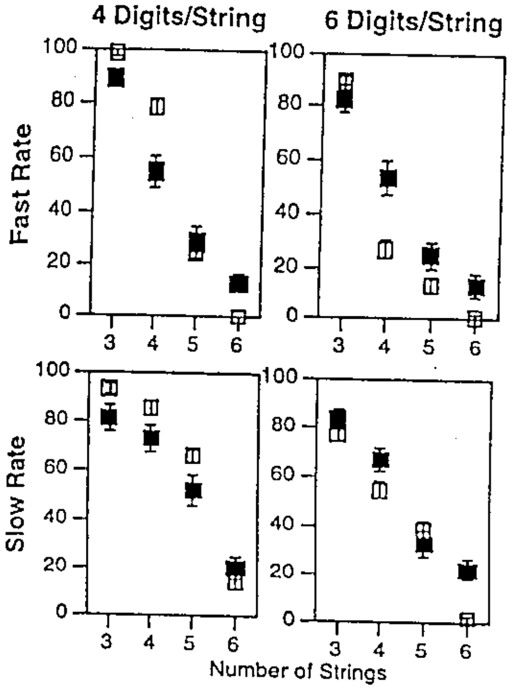

Figure 5.10. Aggregate data (solid boxes) and first-pass model predictions (open boxes) with standard error bars. The dependent measure on the ordinate is percentage of trials perfectly recalled.

recalled (3, 4, 5, or 6), number of digits per string (4 or 6), and presentation rate (0.5 s or 0.7 s). Since the main dependent measure of interest in this task is probability of recall, we computed the proportion of trials of a given type for which all the digits were recalled in the correct position (and averaged across simulations).

Figure 5.10 presents the model's predictions. These predictions are *not* based on an optimal fitting of the parameters: Most parameters were left at their default setting (e.g., $W = 1$), and two parameters, F and s, were adjusted slightly to produce predictions in the appropriate range. Also presented in Figure 5.10 are the aggregated results from 22 subjects who completed 64 trials each (four replications of each of the 16 trial types). Even without optimally tuned parameter values, these predictions show effects of the three main factors that are similar to those in the data. Specifically, our model predicts higher recall probabilities for trials with fewer strings (fewer to-be-recalled digits), for 4-digit strings over 6-digit strings, and for the slower presentation rate over the faster rate. The effect of number of strings is produced by the model because the fewer memory digits, the more source activation that gets spread along each link from a goal node to a declarative node. The effect of number of digits per string is produced by the model because fewer digits to be read shortens the delay between encoding and recall of the memory digits (i.e., base-level activation has had less time to decay). And, finally, the effect of presentation rate is produced by the model because the slower rate allows more rehearsals to be made at the end of each string. Presumably in our model this benefit of extra practice opportunities outweighs the disadvantage of slightly longer delays to recall. This advantage of the slower presentation rate observed in the data also supports subjects' self-report of doing extra rehearsals after each memory digit was presented. Our model would have predicted a disadvantage for the slower rate if rehearsals did not occur during this extra time.

This first-pass model fit is quite encouraging. Even without an optimized parameterization, the best-fitting line between the data and predictions is observed = 0.71 · predicted + 0.16, $R^2 = .88$. Nevertheless, there are two deficiencies. First, it appears that the model tends to overpredict for the 4-digits-per-string trials and underpredict for the 6-digits-per-string trials, and second, the standard error bars for the model's predictions are consistently smaller than those for the data. To address these deficiencies, we next moved to incorporating individual differences into our model.

Modeling Issues II: Including Individual Differences in the Model

ADDING WORKING MEMORY DIFFERENCES TO THE MODEL. We proposed that the *W* parameter in our model, representing an attentional resource, would reflect individual differences in working memory. The model presented earlier, however, took *W* as fixed across all simulations. To incorporate individual differences in *W* in our model, we ran a different set of 22 simulations (one for each subject, as above), but this time each simulation was randomly assigned its own *W* value. (Each *W* was drawn from a Normal with mean 1.0 and standard deviation 0.25.) Here, the different simulations represented different subjects, each with its own limit to source activation. We maintained all other parameter settings, that is, *no* optimal parameter fitting.

Figure 5.11 shows the improved fit attained by the same model as in Figure 5.10 but with randomly varying *W* values. The best-fitting line between the observed data and these predictions is more similar to the line $y = x$ than before: observed = 0.95 · predicted + 0.02, $R^2 = .92$. This reflects the fact that the average predictions of the model are now closer to the average observed accuracies. An important point to note here is that we did not fit each individual *W* parameter to a particular subject's data; rather, we drew the 22 *W* values randomly from a normal distribution. Thus, merely by adding variability to the input of our model via the *W* parameter, we obtained a better fit to the data. Another improvement in this second-pass model fit is that the stan-

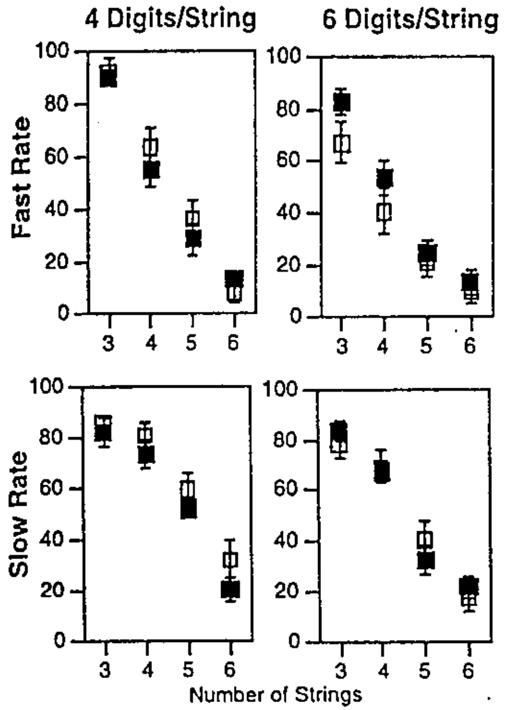

Figure 5.11. Aggregate data (solid boxes) and individual-differences model predictions (open boxes) with standard error bars. The ordinate in each plot is percentage of trials perfectly recalled.

dard-error bars of the predictions now appear similar in size to the error bars of the subjects. In contrast, the previous model did not exhibit enough variability in its predictions. In summary, these results show that by incorporating individual differences in our model, the mean level of the predictions changed as well as the standard error of those predictions. Such changes arose because of the nonlinearities in our model. In fact, adding variability to the input can have such effects in any nonlinear system.

FITTING INDIVIDUAL SUBJECT'S WORKING MEMORY DIFFERENCES. Although the foregoing model takes into account individual differences, the predictions portray only performance in the aggregate. It is possible that a model (even one that takes into account individual differences) can capture aggregate results but not be able to fit the data of any individual subjects. Thus, we next fit the parameter *W* to the data for each subject individually, keeping the other parameters fixed. Our model predicts that the higher an

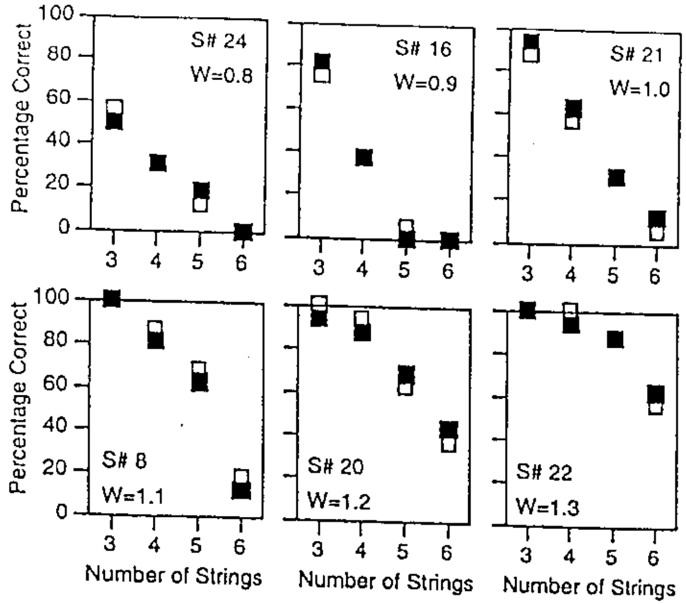

Figure 5.12. Individual participant data (solid boxes) and predictions (open boxes).

individual's attentional capacity W, the more likely correct retrievals will be. This influence of W on retrieval is twofold: (a) higher W values lead to faster retrieval latencies, which means that more rehearsals (learning) can be fit into a fixed amount of time, and (b) higher W values lead to the spreading of more source activation to memory elements that are strongly linked to the goal nodes, making those memory elements more accessible.

As Figure 5.12 shows, the model accounts well for individual subject's recall performance, even matching the shape of individual subject's data. In the figure we display the data only by number of strings to maintain a sufficient number of replications per data point. Although error bars are not plotted for the model predictions, activation noise leads to stochasticity in retrieval; even for a fixed value of W, the model's predictions vary somewhat from simulation to simulation. The six subjects in Figure 5.12 were chosen to represent the range of estimated W values; the model provided a good fit for all of the 22 subjects. Again, it is important to note that for these individual subject fits, the global parameters of the model were maintained at the same fixed (and not optimized) values. The only parameter we specially estimated for these fits was a particular W value for each subject.

It is interesting to note that this procedure led to a bell-shaped distribution of estimated W values for our sample (Figure 5.13): a few subjects were best fit

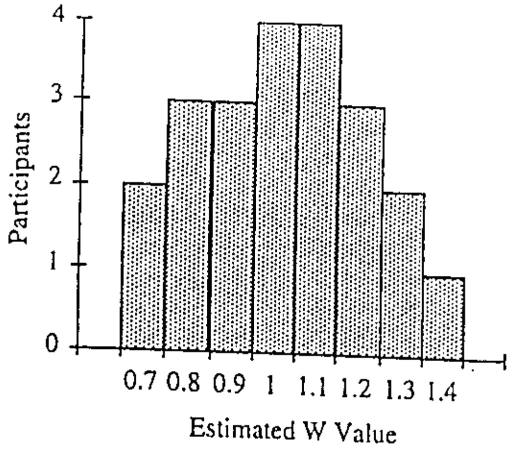

Figure 5.13. Histogram of number of participants with different estimated values for W.

by high or low W, and most subjects were fit by $W \sim 1$. Thus, these W estimates tell us something about the subject-to-subject variability in the quantity that W represents. Moreover, according to our model, each participant's W value represents a fixed quantity of source activation for that individual, which should be reflected across repeated performance of the same task and across performance in different tasks. We take up these issues of fixed W for each subject in the next two subsections.

Modeling Issues III: Modeling Performance Across Sessions

To begin to test our hypothesis that a subject's working memory capacity is fixed across time and tasks, we asked several subjects to perform the digit working memory task multiple times (each time with different stimuli). This allows us to measure the test–retest reliability of our paradigm and to evaluate how well our model – with a single W value per subject – can fit multisession data.

The declarative and procedural knowledge of the model is the same as that presented earlier. The one important difference from the previous treatment, however, is that here we are exploring the model's performance across multiple sessions. This means that we must consider the learning implications of getting practice at the task. In the general description of the ACT-R theory presented earlier, we focused exclusively on declarative learning mechanisms. However, repeating the digit working memory task with different stimuli does not allow for much benefit from declarative learning – nodes are not repeated or practiced across trials let alone sessions. The most reasonable kind of learning in which subjects likely engage as they repeat this task is *procedural learn-*

ing. In ACT-R, procedural learning involves the strengthening of productions with practice. This strengthening mechanism is analogous to the declarative learning mechanism; instead of increasing a node's base-level activation with each use, it involves increasing the strength of a given production with each use. As in declarative learning, these strength boosts decay with time since each use. Equation 8 specifies the strength of production p in terms of its uses at time lags t_k:

$$S_p = \log (\Sigma t_k^{-d}), \tag{8}$$

Procedural learning has a similar effect to honing a useful tool: The more often the tool is used, the better it is used. That is, a production that has been used more frequently and more recently will require less time to do its processing. This allows us to extend the latency function of Equation 3 to describe the latency of retrieving declarative node i using production p:

$$T_{ip} = Fe^{-(S_p + A_i)}, \tag{9}$$

where S_p is the strength of the production and A_i is the total activation of the declarative node.

In terms of our digit working memory model, the basic prediction is that subjects will get faster at the processes involved in the task (e.g., doing rehearsals) as they gain experience. What implications does this have for memory performance? As we demonstrated, having more time for rehearsals leads to improved recall. Therefore, when subjects get faster because of procedural practice, they will have time for more rehearsals so greater declarative learning will produce an increase in their proportion of correct recalls. This prediction allows us to expect reliability of performance across repetitions of this task to be of a certain form: Subjects' recall scores will not necessarily hover at the same level across sessions but will tend to improve with practice.

To generate quantitative predictions regarding memory improvement across sessions, we analyzed our model in light of ACT-R's basic mechanisms to determine the effects of production practice on retrieval probabilities. By drawing on the ACT-R equations presented up to this point, we obtained the following approximation (see Appendix B for details):

$$\text{LogOdds}(i) = C + D \ln (i), \tag{10}$$

where the equation refers to the log odds of correctly retrieving individual nodes at session i. When retrievals are plotted against session number on log–log coordinates, Equation 10 predicts a linear relationship between retrieval odds and session number. The value C in Equation 10 would be the y-intercept of that line; it is a constant related to the log odds of correct retrieval at session 1, which is a function of W as shown in the preceding model fits. Therefore, individual differences in W should lead to differences in the y-intercept of different subjects' retrieval functions. The value D, on the

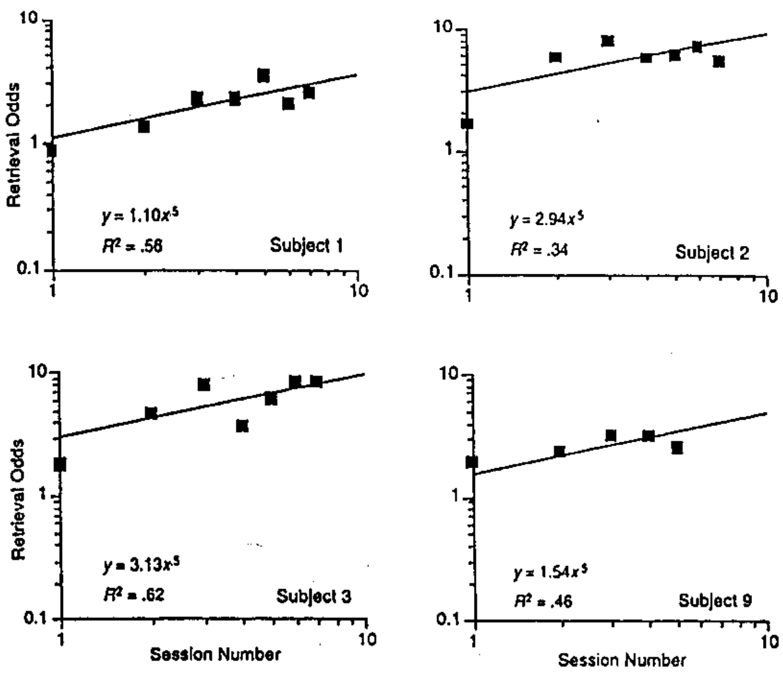

Figure 5.14. Individual subjects' average odds of retrieving memory digits, by session (in log-log coordinates).

other hand, represents the slope of the retrieval function in Equation 10; it is a constant related to the decay rate d. In all the models we have discussed, d was taken as fixed at 0.5 (the ACT-R default). Therefore, we would not expect differences in the slope of different subjects' retrieval function. This is consistent with other evidence that people do not differ in decay rates (e.g., Salthouse, 1994).

Figure 5.14 shows the data of four subjects who participated in five or more sessions. These data are plotted in log–log coordinates with odds of digit retrieval against session number. As the model predicted, for each subject, there is improvement across sessions. Moreover, the subjects differ in their session 1 retrieval performance, which is consistent with the predicted between-subject variability in W. The predictions of Equation 10 are represented by the solid line in each panel. These lines were fit with a common slope ($D = 0.5$) and estimated values for C. Because C is a function of an individual's W value, the different intercept values reflect our model's account of working memory differences in this task.

The main point of these individual-subject, multisession plots is to show that a single W value does not preclude improvement across sessions. Indeed, the fitted lines in Figure 5.14 are based on a stable W parameter for each individual and yet show improvement across sessions that is consistent with the data.[9] These results also suggest that it is helpful to view test–retest reliability and stability of individual differences in terms of a model of cognitive processing. In this way, learning effects can unfold naturally, and the model can provide mechanistic interpretations of both the stable components of behavior (e.g., relatively fixed W for each subject) and the variable ones (e.g., learning and practice effects).

Modeling Issues IV: Generalizing to Another Task

Thus far, we have shown that varying the amount of source activation (W) captures individual differences in a digit working memory task and that these differences appear stable across repeated sessions. One of the primary strengths of the ACT-R theory, however, is its broad applicability across different tasks. Therefore, in this section we show that varying the W parameter in a model of a different task produces the observed individual differences in that task as well. This supports the notion that W is general enough to account for memory differences across tasks. The next step then will be to explore the hypothesis that the same value for W predicts a given individual's performance on multiple tasks. Here, we describe our preliminary results on the way to that goal.

The task that we model has previously been studied in its exact form by Zbrodoff (1995) and others (Logan, 1988; Logan & Klapp, 1991; Rabinowitz & Goldberg, 1995). It is called an alpha-arithmetic task because the goal of each trial is to verify whether or not a given statement involving letters and numbers is true. For example, the statement A + 2 = C is true because C is two letters after A in the alphabet. Correspondingly, the statement A + 2 = D is false. In this task, a statement is presented, and the subject presses one of two keys to indicate whether the statement is true or false. Subjects' latency and accuracy are measured for each trial. We used stimuli from Zbrodoff's (1995) Experiment 1.[10] These include 12 different alpha-arithmetic statements, half true and half false, with numerical addends of 2, 3, and 4. During the experiment, the 12 unique statements were repeated (in random order) 48 times, giving subjects a chance to learn the facts associated with each.

Past results using this task suggest that subjects initially solve the problems by counting up from the initial letter a number of times specified by the numerical addend. More counts are required for problems with larger numer-

[9] Our procedural learning account of this memory improvement is also consistent with subjects' reports. When asked about their strategies after each session, participants tended to maintain the same basic store–rehearse–retrieve strategy mentioned earlier.

[10] We thank Jane Zbrodoff for sharing her experiment software with us.

Table 5.4. *Critical Productions for Alpha-Arithmetic Task*

Retrieve alpha-arithmetic
IF the goal is to analyze START + ADDEND = RESULT and the fact START + ADDEND = RESULT has truth value *X*
THEN set a goal to respond *X*

Compute alpha-arithmetic
IF the goal is to analyze START + ADDEND = RESULT
THEN set a goal to compute START + ADDEND and compare with RESULT

ical addends, producing what is called the "problem-size effect": Latencies for 4-addend problems are greater than for 3-addend problems, which are greater than for 2-addend problems. With practice, however, subjects gradually learn the alpha-arithmetic facts required to solve the problems and thus no longer need to count to verify each statement but can simply retrieve the relevant fact. This retrieval phase is signaled by a reduction in the problem size effect (i.e., the difference in latency across the 4-, 3-, and 2-addend problems decreases or disappears because subjects are no longer counting up the alphabet).[11] The processes involved in this task are, for the most part, different from those in the digit working memory task, and this alpha-arithmetic task does not explicitly involve dual tasks. For our purposes, the task allows for the study of working memory effects (and individual differences therein) in a relatively simple learning task.

Our model implements two productions for deciding whether the current statement is true (Table 5.4): (a) a production to retrieve a relevant alpha-arithmetic fact and (b) a production to initiate a sequence of steps for counting up the alphabet. The retrieval production will tend to be preferred,[12] but if the retrieval attempt fails, the counting production will be initiated. Given these two productions, performance is mainly driven by the activation of alpha-arithmetic facts in declarative memory because these activations determine the probability and latency of retrieval (i.e., the first production).

Figure 5.15 displays a node structure corresponding to the fact "A + 3 = D is true." Suppose this structure is already in declarative memory (i.e., the fact has already been encountered and stored on a previous trial); it will have a certain base-level activation. When the current goal involves processing the

[11] Notice, however, that in ACT-R retrieval latencies are a function of practice, so if the different sized problems occur with different frequencies, some problem-size effect is predicted to remain.

[12] Another related approach involves using the actual activation values of the problem representation to select among strategies, i.e., adopt a retrieval strategy if enough activation from the elements of the problem intersect, but otherwise use a calculation strategy (see Schunn, Reder, Nhouyvanisvong, Richards, & Stroffolino, 1997).

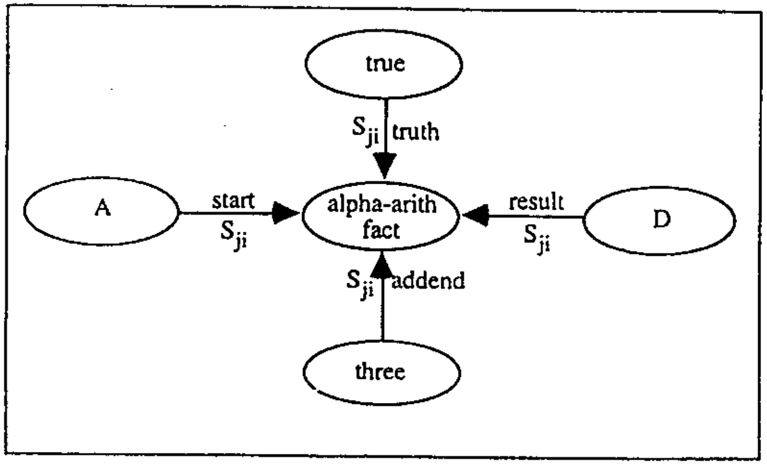

Figure 5.15. Node structure corresponding to the fact "A + 3 = D is true."

statement "A + 3 = D," the goal nodes "A," "three," and "D" will all spread source activation to this alpha-arithmetic fact, thereby increasing its total activation along all three links. This fact had been used only once many trials ago, and so its base-level activation would have decayed to a low value. These two components – received source activation and base-level activation – combine to determine a node's current accessibility.

As this model progresses through the alpha-arithmetic task, an important change occurs in its processing: Initially, it will tend to compute its response by counting (because the necessary alpha-arithmetic facts are not yet present or have very low base-level activations), and later (once these facts are present and have been practiced), the model will be able to consistently respond via retrieval. This general transition is consistent with aggregate results obtained for this task.

But what are the consequences of variation in *W* for this model? As in the digit working memory task, the higher *W*, the faster all retrievals will be and hence the shorter observed latencies will be. This effect applies both to the counting approach (because alphabet retrievals are required) and to the retrieval approach (because alpha-arithmetic fact retrievals are required). One prediction is that a person with a higher value for *W* will be faster overall than a person with a lower value of *W*. Recall from Equation 3 that latency is a power function of total activation. Thus, ACT-R predicts a linear relationship between log latencies and log number of node accesses (base-level activation component of total activation) and a shift in these lines for individual differences in *W* (source activation component). Figure 5.16a shows this relation-

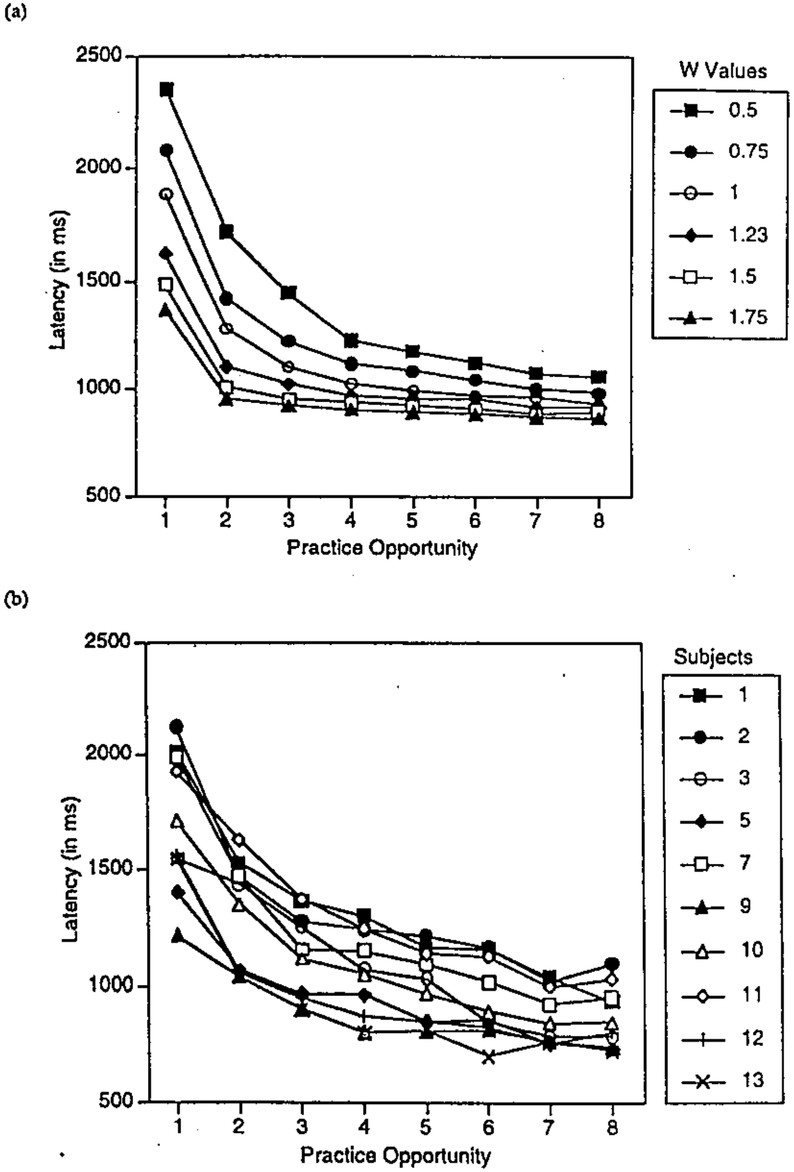

Figure 5.16. (a) Latency predictions for alpha-arithmetic as a function of node practice and *W*. (b) Observed practice effects for alpha-arithmetic experiment.

ship for several values of W. These values correspond quite well with the observed latencies for 10 different subjects (Figure 5.16b).

Another prediction of our model involves how quickly people will learn the alpha-arithmetic facts and thus make the transition from counting to retrieving. Recall that the higher W is, the more source activation will be propagated from the goal nodes to the alpha-arithmetic facts that are being learned throughout the course of the experiment. This makes the alpha-arithmetic facts more accessible than they would otherwise be with just their base-level activations alone. Therefore, an individual with high W will be more likely to retrieve a node with a certain base-level activation than will an individual with low W. This means that a person with high W will be able to retrieve the alpha-arithmetic facts earlier on in the experiment (when they have been less practiced) than will a person with a low W value. In terms of the two approaches for processing each trial, computation should phase out earlier for high-W subjects than for low-W subjects. We can gain an indirect measure of the relative amounts of computation and retrieval by inspecting the problem-size effect that the model produces in the first, second, and third blocks of the experiment. The problem-size effect here is measured as the average increase in trial latency per increase in numerical addend (i.e., how much slower is the response for +4 trials vs. +3 trials vs. +2 trials?). Figure 5.17a plots the model's problem-size effect for different values of W. Note that all the curves show a decrease in the problem-size effect, suggesting more retrieval and less computation. Moreover, the curves with high W (faster retrieval latencies) end up with very small problem-size effects, suggesting that there is almost no effect of addend size (and that retrieval is being used universally) in block 3. Figure 5.17b plots the corresponding observed data, with a separate curve for each subject, and shows similar effects. In summary, varying the W parameter in our model of this alpha-arithmetic task produces variations in performance (as measured by both latencies and problem-size effects) that correspond well to those exhibited across our sample.

Conclusions, Future Issues, and Speculations

We believe that the approach described here is a promising one. Table 5.5 gives a summary of our answers to the designated questions; however, to summarize more generally we list here what we take as some of our modeling accomplishments: (a) We developed several models in the ACT-R framework that produced working memory results similar to those exhibited by subjects. (b) We showed that varying the W parameter in those models improved model fits and produced a similar range of performance to that exhibited across subjects. (c) We fit individual simulation runs to individual subjects' data by estimating a single individual difference parameter, W. (d) Given the procedural learning mechanism built into the ACT-R framework, our models captured – with a stable W value for each individual – working memory

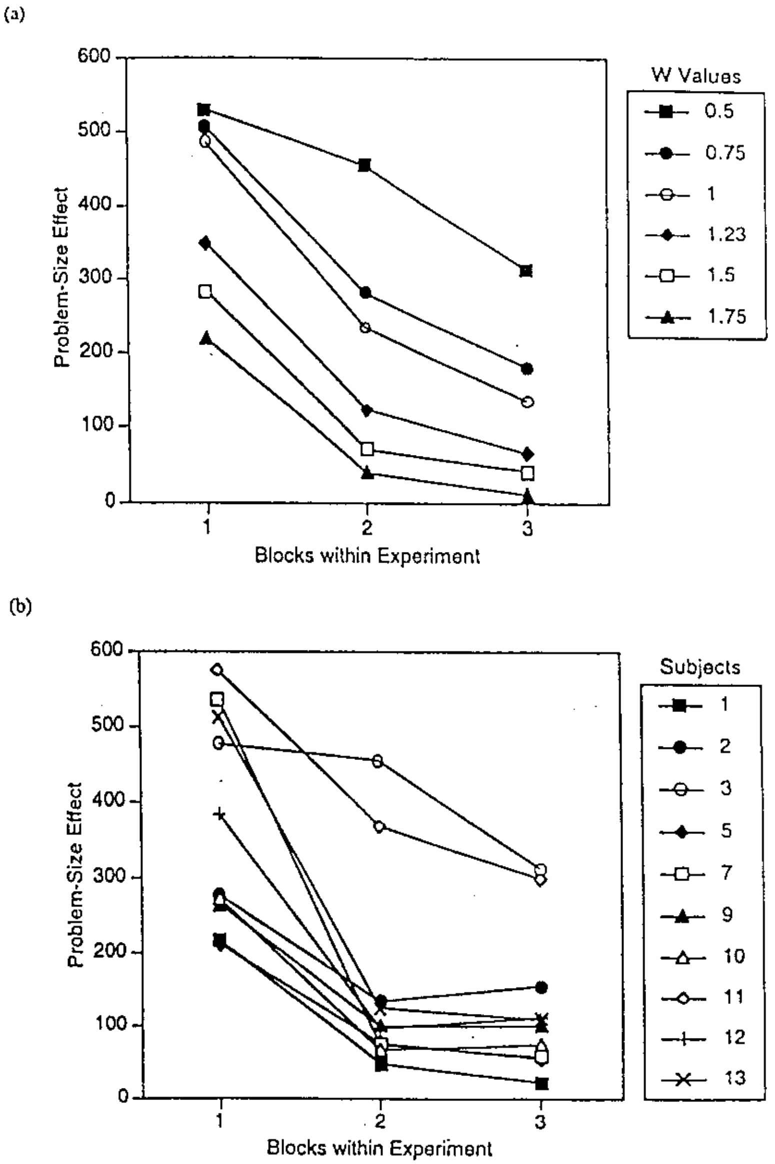

Figure 5.17. (a) Predicted and (b) observed problem-size effect in alpha-arithmetic task.

Table 5.5. *Summary of Answers to Designated Questions*

(1) **Basic Mechanisms and Representations in Working Memory**
Because working memory is conceived of as the highly activated subset of declarative memory, its representations are the same as those for declarative information. Declarative knowledge is represented as nodes in an interconnected network, with different types of information using different combinations of links (see Figure 5.1a). The basic mechanisms in working memory are (a) the spreading of source activation (a kind of "attentional energy") from the current goal to related nodes in declarative memory and (b) the learning and decay of declarative nodes' base-level activations. These two kinds of activation combine to determine a declarative node's accessibility: More highly activated nodes are more likely to be retrieved and tend to have shorter retrieval latencies.

(2) **The Control and Regulation of Working Memory**
The mechanisms described in answer 1 control the processing of information. Source activation propagating from the goal increases the total activation of goal-relevant nodes. This leads to context effects. However, there is an imposed limit to the amount of source activation focused on the goal, so this resource must be shared. The base-level activation of nodes reflects their accessibility independent of context. Base-level activation is not constrained by a fixed limit; it keeps accumulating the more a node is used. This leads to practice effects. However, base-level activation decays with time since each use, producing forgetting effects and sensitivity to time delays.

(3) **The Unitary Versus Non-Unitary Nature of Working Memory**
Because we conceive of working memory as the activated subset of declarative memory (i.e., it can be composed of all kinds of information and structures), it may be considered a non-unitary construct. On the other hand, ACT-R posits a limit to the source activation propagating from the current goal. This limited resource is fixed across all tasks, making working memory processing appear unitary. This common limit to source activation, however, is just one contributor to working memory performance; knowledge and strategy are other variables that influence performance.

(4) **The Nature of Working Memory Limitations**
Working memory limitations are imposed by the constraint that a limited amount of *source activation* (attentional energy) is directed from the current goal. This source activation must be shared among the elements of the goal (the more complex the goal, the smaller the shares) and then spread to neighboring nodes in declarative memory (the more neighbors, the smaller the subdivided amounts). This propagation of source activation to declarative memory serves to differentiate goal-relevant nodes from other nodes, making the former more accessible.

(5) **The Role of Working Memory in Complex Cognitive Activities**
The same working memory processes apply across all cognitive tasks. Because of the limit to source activation, more complex tasks produce greater sharing of that limited resource and hence lead to degraded performance. These effects of limited source activation apply in both dual-task situations and

Table 5.5, continued

memory-loaded situations. ACT-R models of a variety of complex cognitive tasks have provided good quantitative and qualitative fits to observed latencies and probabilities of recall.

(6) **The Relationship of Working Memory to Long-Term Memory and Knowledge**
Working memory is conceived of as the more activated subset of long-term memory. Working memory elements will tend to be those nodes in declarative memory that are strongly related to the current goal (i.e., those receiving extra source activation from the goal). Other nodes for which this is not true may still be in working memory, however, because of their high base-level activation (e.g., highly familiar concepts).

(7) **The Relationship of Working Memory to Attention and Consciousness**
A conception of working memory complementary to that in answer 6 emphasizes processes over contents. The primary working memory process is the spreading of source activation from the goal. This process represents the notion that attention is focused on elements of the goal that thereby modulate other information processing (e.g., retrieval). In this sense, we liken limited source activation to limited attentional resources. Working memory is related to consciousness in that elements of the goal and nodes in declarative memory are accessible to conscious awareness because of their heightened activation. Declarative nodes below this threshold of awareness can still vary in total activation; therefore, they vary in terms of the processing required to bring them into awareness.

(8) **The Biological Implementation of Working Memory**
Although connecting the theoretical construct of working memory to a biological implementation requires bridging a wide gap, neuropsychological and brain-imaging results suggest a biological implementation that is consistent with our theoretical position and computational models. First, damaging an ACT-R model's ability to spread source activation (an "ability" tightly linked to working memory and the W parameter) produced behavioral impairment similar to that of frontal-lobe patients (Kimberg & Farah, 1993). Second, brain-imaging data (Cohen et al., 1994) show that both context-sensitive processing and the maintenance of symbolic information produce activation in the prefrontal cortex. These brain functions have been modeled by a connectionist framework in which retrieval processes (represented as links between PFC and other brain areas) are modulated by the current goal (represented in the PFC). Third, brain-imaging studies have revealed individual differences in the parameters describing brain activation during working memory tasks. Mapping these results onto our framework suggests the goal-based processing (i.e., spreading of W) in our models describes brain functions associated with the PFC and the PFC's connection to other brain areas.

improvements across repeated sessions of the same task. (e) We showed preliminary simulations and data suggesting that these results generalize across tasks.

Clearly there is much more to do. Our next major goal is to explore the stability of W across tasks, that is, to explore whether the W parameter will predict individual differences for a given subject across tasks. This will test the hypothesis that W reflects a stable individual difference in working memory processes. Other research in this area suggests that the question is still an open one. Cases of both stability and change in individual differences across tasks have been found (e.g., Cantor & Engle, 1993; Shah & Miyake, 1996).

Our approach may help clarify these different results because it will explore individual differences in working memory in the context of computational models of cognitive processing. Our work has already shown that a complete cognitive model can clarify issues of test–retest reliability (by modeling test–retest effects in terms of procedural learning) and issues of individual differences (by incorporating variability in our modeling efforts and reducing important sources of strategic variability in our empirical efforts).

APPENDIX A

Take source activation of goal i at time T that has not been in focus since time t_i to be decaying according to the function $e^{-(T-t_i)/\tau}$ (Equation 7).

The sum of all such goals' source activations is bounded by

$$\int_0^T e^{-(T-t)/\tau}\,dt = e^{-T/\tau} \cdot \tau[e^{-T/\tau} - 1] = \tau\,[1 - e^{-T/\tau}].$$

As the current time T increases, this expression goes to τ, which is to say the total amount of source activation in the system, with exponential decay rate τ upon goal switching, is capped by τ. In our framework, W represents that same cap, suggesting that our W parameter and Lü, Williamson, and Kaufman's (1992a,b) τ represent the same individual difference variable.

APPENDIX B

When practice opportunities are equally spaced, strength of production p, S_p, is approximated by $\ln\left[\frac{NL^{-d}}{1-d}\right]$, where N = number of practice opportunities, L = lifetime of the system, and d = decay rate. Let i = number of sessions and n = number of practice opportunities per session. Then, $S_{p,i} = S_0 + \ln\left[\frac{inL^{-d}}{1-d}\right]$, and the change in strength from session i to session $i+1$ is given by

$$S_{p,\Delta i} = S_{p,i+1} - S_{p,i} = \ln\left[\frac{(i+1)nL^{-d}}{1-d}\right] - \ln\left[\frac{inL^{-d}}{1-d}\right] = \ln\left[\frac{i+1}{i}\right]. \qquad \text{(B1)}$$

Using Equations 3 and B1, the ratio of latencies to retrieve node k with production p at session $i+1$ to session i is

$$\frac{T_{k,p,i+1}}{T_{k,p,i}} = \frac{Fe^{-(A_{k,p}+S_{p,i+1})}}{Fe^{-(A_{k,p}+S_{p,i})}} = e^{-(S_{p,i+1}-S_{p,i})} = e^{-S_{p,\Delta i}} = e^{\ln\left[\frac{i+1}{i}\right]^{-1}} = \frac{i}{i+1}.$$

Let H_i = the average number of rehearsals per digit that are completed at session i. Then, $H_{i+1} = \frac{H_i(i+1)}{i}$, which can be used to determine the base-level activation of a node at sessions i and $i+1$. Applying the production-strength learning approximation to base-level activation learning, the activation of node k is $\ln\left[\frac{NL^{-d}}{1-d}\right]$, where N = number of accesses of the node (rehearsals), L = lifetime of the system, and d = decay rate. Substituting H_i for N, we obtain

$$A_{k,p,i} = \ln\left[\frac{H_i L^{-d}}{1-d}\right] \text{ and } A_{k,p,i+1} = \ln\left[\frac{H_{i+1}L^{-d}}{1-d}\right] = \ln\left[\frac{H_i(i+1)L^{-d}}{i(1-d)}\right].$$

Using these activations in Equation 4 and converting to odds, we get $\text{Odds}(\text{retrieve}_{k,p}) = \frac{e^{A_{k,p}/s}}{e^{\tau}}$, with retrieval threshold τ as the competing node. From this, we derive the ratio for session $i+1$ to i of the odds of retrieving node k with production p:

$$\frac{\text{Odds}(\text{retrieve}_{k,p,i+1})}{\text{Odds}(\text{retrieve}_{k,p,i})} = \frac{e^{\ln\left[\frac{H_i(i+1)^{1-d}L^{-d}}{i^{1-d}(1-d)}\right]^{\frac{1}{s}}}}{e^{\ln\left[\frac{H_i L^{-d}}{1-d}\right]^{\frac{1}{s}}}} = \left[\frac{i+1}{i}\right]^{\frac{1-d}{s}}.$$

By telescoping this odds ratio, we get $\text{Odds}(\text{retrieve}_i) = O_1 \cdot i^{\frac{1-d}{s}}$, where O_1 = Odds of retrieval at session 1, a function of W. Taking the log of both sides of this equation produces Equation 10 in the text, with $D = (1-d)/s$.

REFERENCES

Anderson, J. R. (1993). *Rules of the mind.* Hillsdale, NJ: Erlbaum.

Anderson, J. R., & Matessa, M. (1997). A production system theory of serial memory. *Psychological Review, 104,* 728–748.

Anderson, J. R., Reder, L. M., & Lebiere, C. (1996). Working memory: Activation limitations on retrieval. *Cognitive Psychology, 30*, 221–256.
Baddeley, A. D. (1986). *Working memory.* New York: Oxford University Press.
Baddeley, A. D., & Lieberman, K. (1980). Spatial working memory. In R. S. Nickerson (Ed.), *Attention and performance VIII* (pp. 521–539). Hillsdale, NJ: Erlbaum.
Brooks, L. R. (1968). Spatial and verbal components of the act of recall. *Canadian Journal of Psychology, 22*, 349–368.
Byrne, M. D. (1998). Taking a computational approach to aging: The SPAN theory of working memory. *Psychology and Aging, 13*, 309–322.
Cantor, J., & Engle, R. W. (1993). Working-memory capacity as long-term memory activation: An individual-differences approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 1101–1114.
Caplan, D., Rochon, E., & Waters, G. S. (1992). Articulatory and phonological determinants of word length effects in span tasks. *Quarterly Journal of Experimental Psychology Human Experimental Psychology, 45A*, 177–192.
Chase, W. G., & Ericsson, K. A. (1981). Skilled memory. In J. R.. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 141–189). Hillsdale, NJ: Erlbaum.
Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 16, pp. 1–58). New York: Academic Press.
Cohen, J. D., Forman, S. D., Braver, T. S., Casey, B. J., Servan-Schreiber, D., & Noll, D. C. (1994). Activation of prefrontal cortex in a nonspatial working memory task with functional MRI. *Human Brain Mapping, 1*, 293–304.
Connelly, S. L., & Hasher, L. (1993). Aging and the inhibition of spatial location. *Journal of Experimental Psychology: Human Perception & Performance, 19* (6), 1238–1250.
Conway, A. R. A., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General, 123*, 354–373.
Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466.
Engle, R. W. (1994). Individual differences in memory and implications for learning. In R. J. Sternberg (Ed.), *Encyclopedia of intelligence.* New York: Macmillan.
Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*, 211–245.
Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and new view. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 22, pp. 193–225). San Diego: Academic Press.
Healy, A. F., Fendrich, D. W., Cunningham, T. F., & Till, R. E. (1987). Effects of cueing on short-term retention of order information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 413–425.
Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122–149.
Just, M. A., Carpenter, P. A., Keller, T. A., Eddy, W. F., & Thulborn, K. R. (1996). Brain activation modulated by sentence comprehension. *Science, 274*, 114–116.
Kane, M. J., Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Connelly, S. L. (1994). Inhibitory attentional mechanisms and aging. *Psychology & Aging, 9* (1) 103–112.
Kimberg, D. Y., & Farah, M. J. (1993). A unified account of cognitive impairments following frontal lobe damage: The role of working memory in complex, organized behavior. *Journal of Experimental Psychology: General, 122*, 411–428.
Lebiere, C., & Anderson, J. R. (1993). A connectionist implementation of the ACT-R production system. *Proceedings of the Fifteenth Conference of the Cognitive Science Society* (pp. 635–640). Hillsdale, NJ: Erlbaum.
Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*, 492–527.
Logan, G. D., & Klapp, S. T. (1991). Automatizing alphabet arithmetic: I. Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 179–195.
Logie, R. H. (1986). Visuo-spatial processing in working memory. *Quarterly Journal of Experimental Psychology, 38A*, 229–247.
Logie, R. H. (1995). *Visuo-spatial working memory.* Hove, UK: Erlbaum.
Lovett, M. C., & Anderson, J. R. (1996). History of success and current context in problem solving: Combined influences on operator selection. *Cognitive Psychology, 31*, 168–217.
Lü, Z.-L., Williamson, S. J., & Kaufman, L. (1992a). Behavioral lifetime of human auditory sensory memory predicted by physiological measures. *Science, 258*, 1668–1670.
Lü, Z.-L., Williamson, S. J., & Kaufman, L. (1992b). Human auditory primary and association cortex have differing lifetimes for activation traces. *Brain Research, 572*, 236–241.
Newell, A. (1990). *Unified theories of cognition.* Cambridge, MA: Harvard University Press.
Plaut, D. C. (1996). Relearning after damage in connectionist networks: Toward a theory of rehabilitation. *Brain and Language, 52*, 25–82.
Rabinowitz, M., & Goldberg, N. (1995). Evaluating the structure process hypothesis. In F. E. Weinert & W. Schneider (Eds.), *Memory performance and competencies: Issues in growth and development.* Hillsdale, NJ: Erlbaum.
Reder, L. M., & Gordon, J. S. (1997). Subliminal perception: Nothing special, cognitively speaking. In J. Cohen & J. Schooler (Eds.), *Cognitive and neuropsychological approaches to the study of consciousness* (pp. 125–134). Hillsdale, NJ: Erlbaum.
Salthouse, T. A. (1994). The nature of the influences of speed on adult age differences in cognition. *Developmental Psychology, 30*, 240–259.
Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*, 403–428.
Schunn, C. D., Reder, L. M., Nhouyvanisvong, A., Richards, D. R., & Stroffolino, P. J. (1997). To calculate or not calculate: A source activation confusion (SAC) model of problem-familiarity's role in strategy selection. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1–27.
Sejnowski, T. J. (1981). Skeleton filters in the brain. In G. E. Hinton & J. A. Anderson (Eds.), *Parallel models of associative memory.* Hillsdale, NJ: Erlbaum.
Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General, 125*, 4–27.
Stoltzfus, E. R., Hasher, L., Zacks, R. T., Ulivi, M. S., & Goldstein, D. (1993). Investigations of inhibition and interference in younger and older adults. *Journal of Gerontology, 48*, 179–188.

Waters, G. S., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *Quarterly Journal of Experimental Psychology, 49A*, 51–79.
Yuill, N., Oakhill, J., & Parkin, A. (1989). Working memory, comprehension ability and the resolution of text anomaly. *British Journal of Psychology, 80*, 351–361.
Zacks, R. T., & Hasher, L. (1994). Directed ignoring: Inhibitory regulation of working memory. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory mechanisms in attention, memory, and language* (pp. 241–264). New York: Academic Press.
Zbrodoff, N. J. (1995). Why is 9 + 7 harder than 2 + 3? Strength and interference as explanations of the problem-size effect. *Memory & Cognition, 23*, 689–700.

6 Insights into Working Memory from the Perspective of the EPIC Architecture for Modeling Skilled Perceptual-Motor and Cognitive Human Performance

DAVID E. KIERAS, DAVID E. MEYER,
SHANE MUELLER, AND TRAVIS SEYMOUR

FIVE CENTRAL FEATURES OF THE THEORY

Computational modeling of human perceptual-motor and cognitive performance based on a comprehensive detailed information-processing architecture leads to new insights about the components of working memory. To illustrate how such insights can be achieved, a precise production-system model that uses verbal working memory for performing a serial memory span task through a strategic phonological loop has been constructed with the Executive-Process/Interactive-Control (EPIC) architecture of Kieras and Meyer. EPIC is characterized by five central features that may be compared and contrasted with those of other theoretical frameworks in this volume. These features include:

(1) Formal implementation with multiple component mechanisms for perceptual, cognitive, and motor information processing (cf. Barnard, Chapter 9; Lovett, Reder, & Lebiere, Chapter 5; Young & Lewis, Chapter 7; Schneider, Chapter 10).

(2) Representation of procedural knowledge in terms of a production system whose condition-action rules are all applied simultaneously and repeatedly during the cyclic operation of a central cognitive processor (cf. Lovett et al., Chapter 5; Young & Lewis, Chapter 7; O'Reilly, Braver, & Cohen, Chapter 11).

(3) Executive control procedures that schedule task activities efficiently and coordinate the use of limited-capacity peripheral perceptual-

This research was supported by grant N00014-92-J-1173 from the Cognitive Sciences Program of the Office of Naval Research. The authors thank members of the Brain, Cognition, and Action Laboratory (David Fencsik, Jennifer Glass, Leon Gmeindl, Cerita Jones, and Eric Schumacher) at the University of Michigan for helpful suggestions and criticisms. Many helpful comments from our colleagues, the editors, and other authors of the present book are also gratefully acknowledged.