Published in: Modeling Individual Differences in a Digit Working Memory Task. Proceedings of the Cognitive Science Conference. Stanford, CA August 1997.

Modeling Individual Differences in a Digit Working Memory Task

Marsha C. Lovett
lovett@cmu.edu

Lynne M. Reder
reder@cmu.edu

Christian Lebiere
cl@cmu.edu

Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

Individual differences in working memory are an important source of information for refining theories of memory and cognition. Computational modeling is an effective tool for studying individual differences because it allows researchers to maintain the basic structure of a theory while perturbing a particular component. This paper presents a computational model for a digit working memory task and demonstrates that varying a single parameter captures individual differences in that task. The model is developed within the framework of the ACT-R theory (Anderson, 1993), and the continuous parameter manipulated represents attentional capacity for the current goal.

Introduction

Working memory, or the information that people keep active during processing, plays an important role in cognitive processing and performance. Take, for instance, the mental arithmetic problem 134 x 512. To solve this problem without paper and pencil, one must maintain several intermediate results in memory (e.g., 134 x 2 = 268) while continuing to solve the problem. An important result regarding working memory (that the reader may encounter in solving this problem) is that working memory capacity is limited (e.g., Baddeley, 1986). Thus, when a task places extreme demands on working memory, people may have to resort to different strategies for processing (e.g., rehearsal) or they may exhibit performance decrements (e.g., errors). Another important result in the area of working memory is that limitations in working memory capacity vary from individual to individual (e.g., Engle, 1994; Just & Carpenter, 1992; Kyllonen & Christal, 1990). For this reason, different people may experience differential sensitivity to the working memory demands of a task and hence may engage in different processing strategies and exhibit different patterns of errors.

These two working memory phenomena—the limitations and individual differences in working memory capacity—place constraints on theories of working memory. They consequently have important implications for computational models of working memory. First, for a model to accurately depict cognitive processing, it should not be endowed with an unlimited working memory capacity. Second, to account for differences across problems that vary in their working memory demands, a model should incorporate a (functional) limit on working memory that leads to the same performance effects people exhibit (i.e., it should be similarly sensitive to the working memory demands of various tasks). And, third, for a model to account for differences in working memory capacity among individuals, it should be adjustable to reflect different individuals' responses to the same working memory demands.

Computational models have been developed that deal with these issues to varying degrees. For example, several models capture working memory effects aggregated across subjects (e.g., Anderson, Reder, & Lebiere, 1996; Burgess & Hitch, 1992; Lewandowsky & Li, 1994; Norris & Page, 1996). Other models are able to simulate working memory differences between particular subpopulations of people (Just & Carpenter, 1992).

The work presented in this paper goes beyond previous research by modeling individual differences in working memory *directly*. That is, we develop a detailed model of a task that exercises working memory and then quantitatively explore how varying a particular component of the model can account for *subject-to-subject* differences in performance. We fit the model to individual subjects' data, modeling the processes involved in this memory task at an unprecedented level of detail and avoiding the perils of averaging across subjects. Our approach thus addresses the following questions: can an underlying "system parameter" in a computational theory provide enough flexibility to capture the variation in working memory from individual to individual? Does a particular setting of that parameter accurately depict a particular individual's performance? By focusing on individual differences, our approach offers (a) a better understanding of the distribution of working memory capacity across subjects, (b) a more detailed computational account of working memory's relationship to processing, and performance, and (c) a framework for testing whether an individual difference parameter can account for performance patterns for the same individual across tasks. In particular, we show that incorporating individual differences in our model not only enables it to capture the variability in the data but to provide a better overall fit as well. In this way, our work highlights several valuable benefits of incorporating individual differences into computational models.

In the sections below, we begin by describing a task that was designed to exercise working memory to varying degrees while mimimizing the opportunity for people to adapt their strategies to its working memory demands. Then, we describe our model of this task, developed within the ACT-R architecture (Anderson, 1993) and demonstrate that the model can simulate processing in this task at a detailed level. Next, we explore how varying a single, pre-existing parameter in the ACT-R architecture modulates the predictions of our model to account for individual subjects' data. Finally, we discuss some of the implications of our modeling work and make recommendations for modeling individual differences in general.

The Task

The task we have devised to exercise working memory is a variant of the digit working memory task developed by Oakhill and her colleagues (Yuill & Oakhill, 1989). To perform this task, subjects must read a sequence of digits while maintaining in memory a selected subset of those digits. Figure 1 shows the time-stepped presentation of a single trial. Digits of the first string are presented, one at a time, in a row of boxes. Note that the current digit is always erased before the next digit is presented; thus, subjects must keep apace with the presentation rate as they read these digits aloud. Subsequent digit strings begin with a new digit in the leftmost box and continue digit presentation in the same manner. After reading the digits, subjects are prompted to recall the rightmost digit of each string in the order that the strings were presented. (In Figure 1, these to-be-recalled digits are indicated by thick boxes; in the experiment, they were not visually highlighted but were presented for an extra 100ms of "memorizing" time.) Recalling the string-final digits in this task is analogous to recalling the sentence-final words in the Reading Span Task (Daneman & Carpenter, 1980).

Our digit working memory task is distinguished from related working memory tasks in several ways. First, we maintain a precise digit-presentation rate via computer presentation. This reduces the variability from subjects choosing different reading rates. Second, because our chosen presentation rate is quite fast, it reduces the variability due to different rehearsal strategies. Such variability in less constrained tasks can confound "pure" working memory differences. Third, we vary the presentation rate to study its impact on memory performance. Note that a slower presentation rate makes the reading task easier, but it also increases the difficulty of the memory task by elongating the delay between storing and recalling the memory digits.

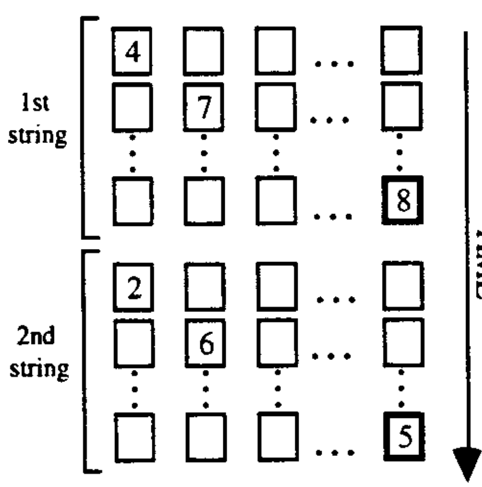

Figure 1 The digit working memory task

Fourth, we present the various trial types in a random order instead of monotonically increasing the difficulty of trials. This randomization eases the assumption that subjects come to each trial with an equal allocation of resources. Moreover, without randomly ordered trials, there may be confounds between various time-based effects (e.g., learning, strategy changes, boredom, fatigue) and trial type. Finally, we include strict recall instructions for our task: the goal is to recall both the identity and position of each memory digit. Specifically, subjects are instructed to repeat *in order* the digits they can recall; i.e., recall must proceed once through the memory list without corrections or backtracking but with the possibility of skipping an unknown digit. This recall procedure reduces recall order variability as well as potential differences in recall strategies.

Empirical Results

Aggregate performance on this task is depicted in Figure 2. Here, the dependent measure is proportion of trials recalled perfectly (i.e., all string-final digits were recalled in exact order of presentation). The factors manipulated in collecting these data were (a) number of strings per trial or number of to-be-recalled digits (3, 4, 5, or 6), (b) number of digits per string (4 or 6), and (c) inter-digit presentation rate (0.5s or 0.7s). All were within-subjects factors; in particular, each of 22 subjects contributed 4 trials to each data point.

All three factors show main effects in these data. The most salient effect is a difference in recall performance for the different number of strings, $F(3, 63) = 125$, $MSE = 0.06$, $p < .001$. This result was expected both because it fits with the general finding of a gradual decrement in aggregate performance with increasing memory load and because our dependent measure required perfect recall of all of

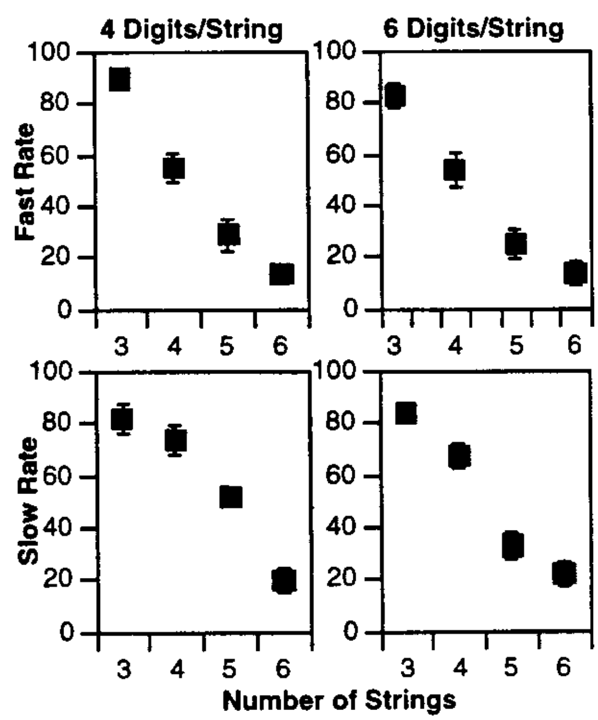

Figure 2 Aggregated data (■) with standard error bars

the to-be-recalled digits. The effect of number of digits per string suggested a slight memory advantage for four-digit strings over six-digit strings, $F(1, 21) = 6.6$, $MSE = 0.03$, $p < .05$. Note that this factor only affects the total number of digits that must be read, not the number of digits that must be recalled. The finding that reading more digits makes the task harder suggests that longer delays and/or more interfering items lead to worse memory. Finally, the effect of presentation rate was informative because it revealed better memory performance with slower rather than faster rates, $F(1, 21) = 20.8$, $MSE = 0.04$, $p < .001$. Finding the effect in this direction suggests that the positive influence of having more time to do the tasks outweighed the negative influence of a longer delay until recall. The only significant interaction was between number of strings and presentation rate, $F(3, 63) = 8.3$, $MSE = 0.03$, $p < .001$. This interaction appears to result from a negligible rate effect for three-string trials, $F(1,21) = 1.63$, $MSE = .02$, n.s., but an advantage for the slow rate on all others, $F(1,21) = 42.6$, $MSE = .03$, $p < .001$.

We also probed subjects, at the end of the experiment, to describe their approach to the task. Subjects mentioned the use of some strategies (e.g., imagery, using number associations) but often commented that they did not find these strategies useful and hence abandoned them. Nevertheless, an interesting and fairly common reported strategy was the limited use of rehearsal (e.g., rehearsing previous memory digits in the time available at the end of each string).[1] We incorporate this information into our model.

The Model

The processes required to perform this task involve reading digits and storing and recalling selected digits. In addition, from subjects' reports, we found a fairly uniform but limited amount of rehearsing digits. We designed our model of the task to reflect all of these processes and to capture the step-by-step activity of subjects.

As mentioned above, the model was developed within the ACT-R architecture. In ACT-R, chunks represent declarative knowledge (facts), and productions represent procedural knowledge (skills). These symbolic knowledge elements are strengthened and deployed according to subsymbolic learning and performance mechanisms specified by the ACT-R architecture. Below, we describe the main symbolic elements in our model and then sketch the ACT-R mechanisms that operate on them.

Our model represents the two main goals of this task separately: reading digits and recalling digits. Each goal is represented as a chunk structure with various pieces of associated information (e.g., trial and position number). The model uses a similar structure for the memory digits in this task. (See Figure 3).

[1]This is consistent with the ANOVA results in that (a) an advantage for the slow rate is consistent with some rehearsal but (b) a disadvantage for longer strings suggests rehearsal did not occur after each digit.

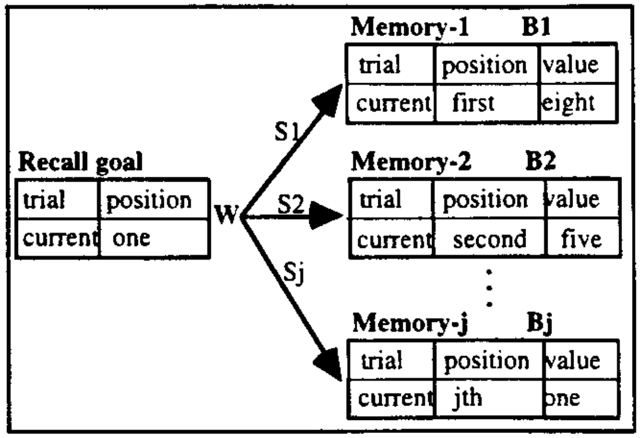

Figure 3 Goal and memory units in our model

The goals and memory elements in our model are acted upon by productions of the form "IF <conditions> THEN <actions>". Figure 4 presents a list of some of the processes implemented by separate productions in our model. The "read" production fires whenever the goal is to read the digits. After a digit has been read, if it is in the last position, the "store" production will fire to create a new memory element for that to-be-recalled digit. This gives the new memory element an initial boost of 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 does not need to take any outward action (such as saying the digit out loud). Instead, by virtue of having retrieved a digit in this production's conditions (i.e., the digit d must be recalled in order to be identified as the digit in position p), ACT-R naturally increases that memory element's activation and hence its likelihood to be recalled later. Finally, the "recall" production retrieves digits at the end of each trial. As described below, the memory element that is retrieved will tend to (but does not necessarily) represent an exact match to the element specified in the current recall goal.

READ:
IF goal is to read a digit & digit d is on screen
THEN say digit d

STORE:
IF goal is to read a digit & digit d is on screen
& d is in last column & d has been read
THEN store d & prepare to rehearse

REHEARSE:
IF goal is to read a digit & 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 &
digit d "matches" but has not been recalled
THEN say digit d

Figure 4 Some productions from our model

In ACT-R, activation is the main unit of "currency" for processing. That is, learning and performance functions are specified in terms of how various elements' activations change and impact performance.

The ACT-R learning mechanism plays a role in our model by specifying how each memory digit's activation is increased. When a memory digit is first stored, it is endowed with an initial activation that decays with time. Each time the memory digit is accessed, it receives an additional activation "boost"; as time passes, however, these activation "boosts" also decay as a power function of the timelag since access. The sum of these decaying activations produces the memory element's *base-level activation*, B_i:

$$B_i = \log(\Sigma t_j^{-d}),$$

where t_j is the time lag since the j^{th} access and d is the decay rate.

For performance, the ACT-R theory posits that a memory element to be retrieved by a particular production gets an additional activation from the current goal. This *source activation*, denoted W, can be conceptualized as the amount of attention directed from the current goal: we take W as the individual difference parameter in our model based on the work of Anderson et al. (1996) and on our own related work. W affects performance by spreading its source activation from the goal to the to-be-retrieved memory element, increasing that element's total activation (Figure 3). A memory element with higher total activation will be more likely to be retrieved. Thus, the model will produce better recall under higher values of W. There is one additional constraint, however: the proportion of source activation that reaches a given memory element, S_i, is reduced as more memory elements are connected to the current goal. In the case of our digit working memory task, this means that source activation will be spread more thinly the more memory digits in the current trial, $S_i \sim$ -log(number of memory elements). Thus, the *total activation* of memory element i is:

$$A_i = W \cdot S_i + B_i + N(0,\sigma^2),$$

where $N(0,\sigma^2)$ represents the Gaussian noise added to each element's activation.

This total activation value is then transformed into a performance measure according to:

$$P(\text{retrieve } i) = (e^{Ai/s})/(\Sigma_j e^{Aj/s}),$$

where $s = \sqrt{6}\sigma/\pi$ and the denominator sums over the competing memory elements.[2] This performance function specifies the model's predictions in terms of retrieval probability for a given item. The critical feature of this function is that probability of recall is a *nonlinear* function of activation and hence a nonlinear function of W.

Errors. The complement of the above retrieval probability (i.e., 1-P(retrieve)) gives the model's prediction for errors of omission (i.e., when a memory digit is not retrieved). Errors of commission (i.e., when an incorrect digit is retrieved instead of the correct one) also occur in systematic ways that need to be reflected in our model. The ACT-R architecture provides a way to predict errors of comission through its partial-matching mechanism. Once the similarity between various items is specified (e.g., how similar are the third and fourth positions in a list? how similar are the first and fourth positions in a list?), this mechanism lowers the total activation of a given memory element with respect to how closely it matches the current goal. Since elements with higher activation are more likely to be retrieved, there is still a bias to retrieve the correct memory digit (if it is above threshold). Nevertheless, with partial matching, similar memory elements (e.g., those in neighboring positions) also have some chance of being retrieved in place of the correct digit. Although it is not the focus of this paper, our model is thus able to capture various error patterns in the data.

Modeling Results

As a first exploration of the model described above, we produced aggregate model predictions (Figure 5). These predictions were based on the default parameter settings prescribed by the ACT-R theory, e.g., the W parameter's default setting is 1.0. To obtain these predictions, we ran the same exact model through our task 22 times, to simulate each of the 22 subjects.

This first-pass fit demonstrates that our model can produce behavior in the range of that exhibited by subjects. The best-fitting line between the data and predictions is *observed = 0.71*predicted + 0.16*, $R^2 = .88$. However, there are two noticeable deficiencies in this first-pass model fit: it appears that the model tends to overpredict for the four-digit strings

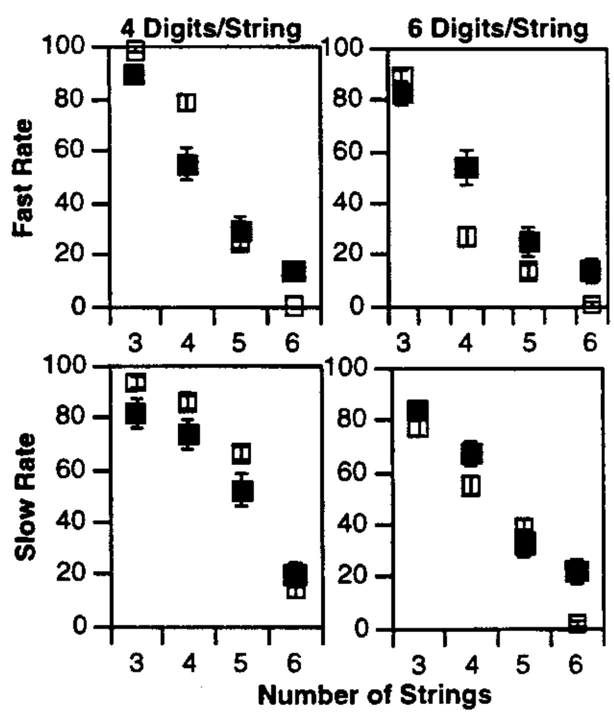

Figure 5 Aggregate data (■) and first-pass model predictions (□) with standard error bars

[2]A retrieval threshold is included as one of the competing elements so that when a memory element's total activation is not above threshold, it is unlikely (depending on σ) to be retrieved.

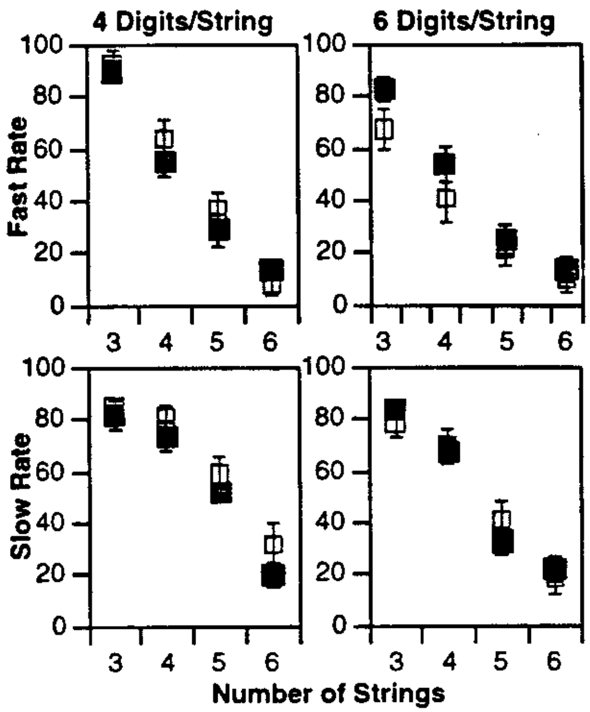

Figure 6 Aggregate data (■) and individual differences model predictions (❑) with standard error bars

and underpredict for the six-digit strings, and 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.

First, we ran the model through 22 simulations of the experiment as above, but this time, each simulation had a different, randomly distributed value for W. We kept the same basic parameter values (i.e., no optimal parameter fitting) but took the W parameter as normally distributed with mean 1.0 and variance 0.0625. Figure 6 shows the improved fit attained (best-fitting line: *Observed* = *0.95*Predicted + 0.02*, $R^2 = .92$). Indeed, the error bars for the model and data in Figure 6 overlap in every case except one. Moreover, by incorporating individual differences in our model, the standard error bars of the predictions now appear more similar to those of the subjects.

While the above model fit suggests that varying W parameter can lead to performance variability that is consistent with the individual differences in our sample, it still suffers from aggregating over subjects. In other words, it is possible that a model (even one that takes into account individual differences) can capture aggregate data but not be able to fit data of individual subjects. Thus, we next fit the parameter W to the data for each subject individually. As Figure 7 shows, the model can account for individual subject's recall performance and even matches the shape of individual subject's data. Note that here we only break down by number of strings to maintain a sufficient number of replications per data point. The six subjects' data presented in Figure 7 were chosen to represent the range of W values; the model provided a good fit for all of the 22 subjects.

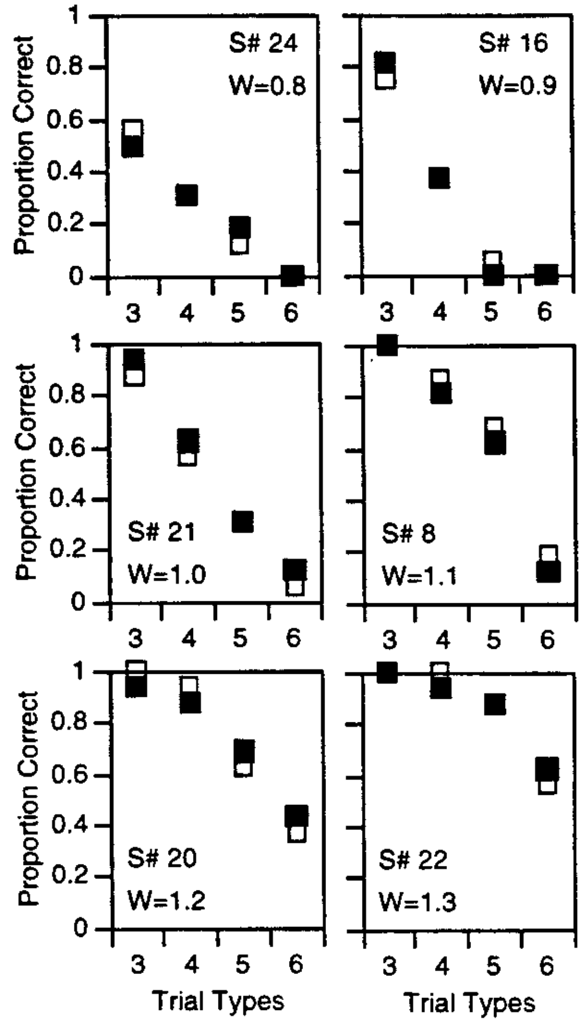

Figure 7 Individual participant data (■) and predictions (❑)

Indeed, this fitting procedure produced a bell-shaped distribution of W values for our sample (See Figure 8): a few subjects were best fit by high or low W, and most subjects were fit by W ~ 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 in other tasks we can also model.

Discussion

In the modeling work above, we have shown that a single, continuous-valued parameter of the ACT-R theory can produce individual differences similar to those displayed by a sample of adults performing a digit working memory task. This parameter (W) modulates the amount of *source activation* spreading from the current goal to associated memory elements and thereby affects their probability of retrieval. While this parameter has been used in other ACT-R models to represent a global attentional resource that is common across individuals, our model takes the parameter as fixed for a given individual but potentially varying across

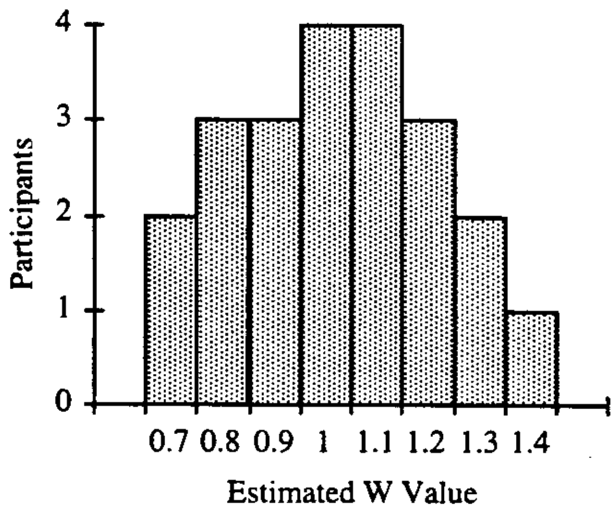

Figure 8 Histogram of number of participants with different estimated values for W, the attentional capacity parameter.

individuals. This individual-differences approach enabled our model to reflect the variability in an aggregated data set and to capture the individual performance curves of each subject. Moreover, our approach predicts that differences in this parameter, as measured by our digit working memory task, will produce systematic individual differences in a variety of different tasks. We are currently testing such across-task predictions of our model.

Our modeling work also highlights the fact that incorporating individual differences in a nonlinear model can have important implications for the model's predictions. First, adding variability to a single parameter in a nonlinear model not only changes the variability of the model's performance but its average predictions as well. This effect was particularly evident in the two aggregate model fits presented in this paper (Figures 5 & 6); here, changing the W parameter from a constant to a random variable impacted the standard errors of the model's predictions and the predicted values themselves. Second, our approach suggests that, in a nonlinear model, a single parameter can have systematic effects on performance across tasks even while performance across the different tasks does not show a strong linear relationship. For instance, depending on the working memory demands of different tasks, a particular individual's performance may not look very similar across tasks. Thus, linear-based analyses of performance may not be able to uncover the common source of individual differences in a nonlinear system. In contrast, our approach uses pre-specified (nonlinear) functions to predict performance and thus is able to estimate a common parameter setting for a given individual and simulate performance across tasks.

Conclusions

Computational models provide an effective tool for studying individual differences in working memory capacity because they allow researchers to maintain the basic structure of a theory while perturbing any given component. As we have demonstrated above, one can then rigorously and quantitatively explore how varying a particular component of the model can account for individual differences, leading to a better understanding of the phenomena at hand and refinements of one's theory. This approach also provides a framework for studying the impact of a single individual difference parameter across tasks and for using computational models to predict individuals' performance on a new task based on the individualized parameter value estimated from their performance on a previous task.

References

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

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.

Burgess, N., & Hitch, G. J. (1992). Toward a network model of the articulatory loop. *Journal of Memory and Language, 31,* 429-460.

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. Sternberg (Ed.), *Encyclopedia of Intelligence.* New York: Macmillan.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122-149.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity. *Intelligence, 14,* 389-433.

Lewandowsky, S., & Li, S. C. (1994). Memory for serial order revisited. *Psychological Review, 101,* 539-543.

Norris, D., & Page, M. (1996). *The primacy model of serial recall,* submitted.

Yuill, N, Oakhill, J., & Parkin, A. (1989). Working memory, comprehension ability and the resolution of text anomaly. *British Journal of Psychology, 80,* 351-361.

Acknowledgments

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.