Behavior Research Methods, Instruments, & Computers
1984, 16 (2), 96-101

SESSION III
INVITED ADDRESS

Jonathan Vaughan, *Presider*

MINERVA 2: A simulation model of human memory

DOUGLAS L. HINTZMAN
University of Oregon, Eugene, Oregon

An overview of a simulation model of human memory is presented. The model assumes: (1) that only episodic traces are stored in memory, (2) that repetition produces multiple traces of an item, (3) that a retrieval cue contacts all memory traces simultaneously, (4) that each trace is activated according to its similarity to the retrieval cue, and (5) that all traces respond in parallel, the retrieved information reflecting their summed output. The model has been applied with success to a variety of phenomena found with human subjects in frequency and recognition judgment tasks, the schema-abstraction task, and paired-associate learning. Application of the model to these tasks is briefly summarized.

This article summarizes simulation work I have been doing to test the adequacy of some theoretical ideas about human memory. The model, named MINERVA 2, makes just a few basic assumptions; but despite this simplicity, it can be applied to a variety of memory phenomena that have been uncovered in several tasks. This paper is only a preliminary account of the work done to date. More complete reports of the capabilities of MINERVA 2, including the work summarized here, are in preparation.

Of central concern in this research has been the question of whether one needs to assume two different memory systems in order to account for memory for individual experiences, on the one hand, and memory for abstract concepts, on the other. MINERVA 2 represents an attempt to account for data from both episodic and generic memory tasks within a single system.

The theory behind the simulation model is primarily concerned with long-term or secondary memory (SM), although it also assumes that there is a temporary working store or primary memory (PM) that communicates with SM. Interactions between the two stores are limited to two elementary operations: PM can send a retrieval cue, or "probe," into SM, and it can receive a reply, called the "echo." When a probe is sent to SM, a single echo is returned. Information in the echo, and its relation to information in the eliciting probe, are the only clues available to PM regarding what information SM contains.

SM is a vast collection of episodic memory traces, each of which is a record of an event or experience. An experience is assumed to occur when a configuration of primitive properties or features is activated in PM, and a memory trace is a record of such a configuration. Each experience leaves behind its own memory trace, even if it is virtually the same as an earlier one; thus, the effects of repetition are mediated by multiple copies or redundancy, rather than by strengthening. Further, there is no separate conceptual, generic, or semantic store. All information, whether specific or general, is retrieved from the pool of episodic traces that constitutes SM.

When a probe is communicated from PM to SM, it is simultaneously matched with every memory trace, and each trace is activated according to its degree of similarity to the probe. The echo that comes back to PM represents the summed reactions of all traces in SM. That is, there is no process by which individual memory traces can be located and examined in isolation. All SM traces are activated in parallel by the probe and they all respond in parallel, and the echo contains their combined messages. What makes this work is that a trace's contribution to the echo is determined by its degree of activation, so only traces that are relatively similar to the probe make a significant contribution to the echo.

This research was supported in part by National Science Foundation Grant BNS-7824987 and in part by a fellowship to the author from the James McKeen Cattell Fund. Thanks are due also to the MRC Applied Psychology Unit, Cambridge, England, for the use of their facilities in some of this work. The author's mailing address is Department of Psychology, University of Oregon, Eugene, OR 97403.

This idea, that activated traces respond in concert during retrieval, was also central to the memory theory of Semon (1923). Semon argued that abstract ideas arise when individual traces having common properties are activated by the same retrieval cue. He assumed that the contents of consciousness are determined by a kind of resonant state (called "homophony"), in which the individuating or distinctive properties of the activated traces mutually interfere, and their shared properties stand out. MINERVA 2 can be seen as an information-processing analog of Semon's notion of homophony.

THE MODEL

MINERVA 2 bears some similarity to MINERVA 1 (Hintzman & Ludlam, 1980), but is applicable to a much wider variety of tasks. It gains this generality in part by adopting simple feature-list representations of stimuli and their traces. Although they are less powerful than the propositional structures employed in MINERVA 1, feature lists are easily generated and stored, and lend themselves to simple and efficient procedures for computing similarity. Their adoption in the present model is primarily for computational convenience.

In the simulations done so far, an experience or event is represented as a vector, or ordered list of feature values each of which belongs to the set +1, 0, −1. The values +1 and −1 occur about equally often, so that over a large number of traces, the expected value of a feature is 0. In a stimulus or event description, a feature value of 0 indicates that the particular feature is irrelevant; in an SM trace description, a value of 0 may mean either that the feature is irrelevant or that it was forgotten or never stored.

In learning, active features representing the present event are copied into an SM trace. Each such feature has probability L of being encoded properly, and with probability 1 − L the tract feature value is set at 0. If an item is repeated, a new trace is entered into SM each time it occurs.

Let P(j) represent feature j of a probe or retrieval cue, and T(i,j) be the corresponding feature of memory trace i. The similarity of trace i to the probe is computed as

$$S(i) = \sum_{i=1}^{N} P(j)T(i,j)/N, \qquad (1)$$

where N is the total number of features that are nonzero in either the probe or the trace. S(i) behaves much like a Pearson r, being 0 when the probe and trace are orthogonal and +1 when they perfectly match, and taking on both positive and negative values.

The activation level of a trace, A(i), is a positively accelerated function of its similarity to the probe. In present simulations,

$$A(i) = S(i)^3. \qquad (2)$$

Notice that if trace i was generated randomly (by a process orthogonal to that generating the probe), then the expected value of A(i) is 0 and the variance of A(i) is quite small. Thus, A(i) should be very near 0 unless trace i fairly closely matches the probe. Raising the similarity measure to the third power increases the "signal-to-noise" ratio, in that it increases the number of poorly matching traces required to overshadow a trace that closely matches the probe.

Intensity

When a probe activates the traces in SM, information is returned in the echo. The echo is assumed to have two properties: intensity and content. The intensity of the echo is given by

$$I = \sum_{i=1}^{M} A(i), \qquad (3)$$

where M is the total number of traces in memory. If the probe does not match or resemble any trace, then E(I) = 0. If there are traces in memory matching the probe, then E(I) increases directly with the number of such traces in memory.

Var(I), like E(I), is a function of the number of target traces. If L = 1, then this function is flat, reflecting only the baseline "noise" in I produced by nontarget traces. If L < 1 and is constant, then Var(I) increases linearly with frequency because the A(i) values of the individual target traces vary, and contribute independently to I. Frequency judgments and recognition judgments are assumed to be based on the intensity of the echo, and therefore characteristics of the I distribution are crucial in simulating performance in these tasks.

Content

The content of the echo is the activation pattern across features that is returned from memory following the probe. It is assumed that the activation of each SM trace, i, is passed to each of its constituent features, j, as the product of A(i) and T(i,j). Note that the product will be positive if the signs of A(i) and T(i,j) are the same, and negative if they are different. The contributions of all M traces in memory are summed for each feature; thus, activation of feature j in the echo is given by

$$C(j) = \sum_{i=1}^{M} A(i)T(i,j). \qquad (4)$$

These C(j) values can range from negative to neutral to positive, and their profile or histogram across features is assumed to be immediately available in PM.

Only traces that are similar to the probe become strongly activated; however, those traces can contain information not present in the probe itself, and thus

MINERVA 2 is capable of associative recall. In order to simulate the retrieval of associative information, the set of features can be divided into two segments. For example, to represent face-name pairs, features j = 1 ... 10 might be reserved for the faces, and the remaining features, j = 11 ... 20, for the names. Then a trace of 20 features would represent a single occurrence of a particular pair. Recall of a name upon presentation of a face can be accomplished with a probe having j = 1 ... 10 filled in and j = 11 ... 20 set to 0, and then focusing on C(11) ... C(20) in the echo. Retrieval of a face given a name would be done in the opposite fashion.

It may be worth pointing out here just how this model differs from "holographic" or associative network models (e.g., Eich, 1982; Hinton & Anderson, 1981), which it at least superficially resembles. In those models, there is just one composite (distributed) memory representation in which all experiences are stored. In MINERVA 2, there is a separate representation for each experience. One might say that the network models collapse experiences at the time of encoding, and that MINERVA 2 collapses them at the time of retrieval. This would appear to give the present model more flexibility; but without direct comparisons, the behavioral consequences of this difference are not immediately evident. One advantage that MINERVA 2 may have over associative network models, however, lies in the way it can deal with the problem of "ambiguous recall."

The ambiguous recall problem is that information retrieved from memory is sometimes only vaguely similar to what was originally stored, or to any acceptable response. One way of getting around this difficulty is to assume that there is a pattern-recognition system that knows the set of acceptable responses and that takes the memory echo as input to be classified (e.g., Eich, 1982). But MINERVA 2 lends itself to a more elegant solution. If the echo content is "normalized" into the −1 to +1 range and turned into a secondary probe, the subsequent echo is more like an acceptable response than the original echo was. And if this process is repeated, there is further improvement. Typically, only a few iterations of this procedure are necessary to produce a virtually perfect copy of the information that was originally stored. In this way, MINERVA 2 appears able to "bootstrap" its way out of the ambiguous recall problem with no help from a memory system on the outside.

For most of the simulations that have been done, stimuli are generated randomly, within constraints appropriate to the task being modeled. Also, probabilistic learning and forgetting processes are often invoked. When repeated, the simulations produce varying results. For this reason, I generate data for a large number of simulated "subjects," and analyze them as one would data from an actual experiment. Comparisons can then be made between simulated and real data, to determine the extent to which equivalent phenomena appear. In what follows, I summarize briefly the basic results that have been obtained with the model so far, leaving detailed treatment of the findings for later reports. Tasks to which MINERVA 2 has been applied are: judgments of frequency and recognition memory, the schema abstraction task, and paired-associate learning.

FREQUENCY JUDGMENTS AND RECOGNITION MEMORY

Suppose a list is presented in which different items occur different numbers of times, and each presentation is encoded in SM with a learning rate, L, less than 1. As was indicated earlier, the echo intensity, I, will roughly reflect the number of traces matching the probe, and both E(I) and Var(I) will increase with the frequency of the test item. For a wide range of combinations of parameter values, the I distributions of neighboring frequencies overlap. By appropriately partitioning the I scale, therefore, one can model yes-no recognition judgments, confidence ratings, and absolute (numerical) judgments of frequency. The choice of criterion settings is an arbitrary matter, but in general the model seems consistent with signal-detection analyses of confidence ratings in recognition memory, and suggests that frequency judgments can be handled in essentially the same way—that is, by extending the signal-detection model to higher frequencies, as a Thurstonian scaling model. Among other findings that this model explains is the fact that frequency-judgment errors, far from being haphazard guesses, are nearly always close to the correct value (e.g., Hintzman, Nozawa, & Irmscher, 1982).

Forced-choice judgments of recognition and relative frequency are modeled by having MINERVA 2 retrieve I values for both members of each test pair and then choose the member yielding the higher value of I. Forced-choice recognition judgments, then, constitute the lower end of the frequency-discrimination scale (i.e., the discrimination of frequency = 0 from frequency > 0). As one moves up the frequency scale, keeping the frequency difference the same, frequency-discrimination performance deteriorates, due to the increase in Var(I). Simulated forced-choice data look much like human data in this respect (e.g., Hintzman, 1969; Hintzman & Gold, 1983; Hintzman, Grandy, & Gold, 1981).

One consequence of the assumption that recognition memory and memory for frequency are based on the same underlying quantity is that MINERVA 2 predicts that forgetting will affect the two tasks in the same way. Following a forced-choice test of both recognition and frequency discrimination, trace decay was simulated by randomly replacing T(i,j) values of −1 and +1 with 0s, and the test was then repeated. The rates of decline in performance on recognition and frequency discrimination were the same. Two experiments on human subjects comparing forgetting rates in recognition and frequency discrimination have obtained essentially this result (Hintzman & Stern, in press).

Judgments of frequency are always relative to some "time window" or context over which frequency is to be integrated. This is demonstrated clearly by the ability of subjects to give largely independent judgments of frequency for the same word in two different lists (Hintzman & Block, 1971). Such list-specific frequency judgments can be simulated in MINERVA 2 by assuming first of all that certain trace features distinguish between lists. For example, let T(i,1) to T(i,4) be −1, +1, 0, 0 if trace i is from List 1, and 0, 0, −1, +1 if it is from List 2. The remainder of the features (j = 5 ... N) represent the item itself. Now, probes can be constructed that activate features specific to the target list and inhibit those specific to the nontarget list (inhibition is produced by multiplying feature values by −1). Thus, in the present example, j = 1 ... 4 in a probe targeting List 1 and inhibiting List 2 would be −1, +1, +1, −1, with the remaining features representing the test item. The frequency judgment is based on echo intensity, as before.

Using this method, quite realistic list-specific frequency judgments are obtained. As with human subjects, there is a small amount of generalization from frequency of the test item in the nontarget list. In both the simulated and the human data (Hintzman & Block, 1971), mean judgments increase in an approximately linear fashion with both the target-list and the nontarget-list frequencies, but the target list receives much greater weight. This same kind of selective-activation scheme can be applied to source-specific and context-specific frequency and recognition judgments (e.g., Johnson, Raye, Wang, & Taylor, 1979). The success of MINERVA 2 on this problem demonstrates the model's flexibility in being able to selectively activate different subsets of memory traces—even among very similar ones—depending on the structure of the retrieval probe.

It should be added that MINERVA 2 provides an account of why and in what sense memory for frequency is "automatic" (Hasher & Zacks, 1979). An analog representation of frequency exists in SM as an inevitable consequence of attending to repetitions; but frequency is inferred not from a count of the relevant traces, but from echo intensity, which is influenced by the degree to which each trace—target or nontarget—is activated by the probe. Apparent frequency is not a product of trace frequency alone.

SCHEMA ABSTRACTION

MINERVA 2 provides an "exemplar" approach to generic memory—but a more extreme form of the exemplar view than theorists usually entertain (cf. Smith & Medin, 1981). A general concept is represented in SM not by representations of subvarieties (which are themselves generalizations), or even by representations of individual members, but by a large number of traces of individual experiences with individual category members. How the concept is represented in PM will vary, depending on which of these episodic traces have been activated strongly by the probe. The active representation in PM should be highly context specific, since contextual information contributes importantly to the probe.

Among the phenomena that a theory of generic memory should explain are that general concepts are typically experienced as stripped of information about particular contexts (Brewer & Pani, 1983), that people tend to agree in their descriptions of a prototypical member of a category, that category exemplars vary in their perceived degree of membership or "typicality," and that category membership seems to be more a matter of family resemblances than of defining features (Rosch & Mervis, 1975; Smith & Medin, 1981).

Schema-abstraction experiments, in which subjects learn to classify unfamiliar exemplars (e.g., dot patterns generated by perturbing the positions of dots in category prototypes) have been seen as reflecting the same underlying processes as those by which natural categories are learned. A number of replicable results have been generated using this paradigm, and these are data with which the output of MINERVA 2 can be compared.

Associative learning is required in the schema-abstraction task, since category labels must be given when exemplars are presented. This was accomplished in the present simulations in the way that was previously described: One subset of features was taken to represent the stimulus (e.g., dot pattern) and the remaining features to represent the category name. A prototype stimulus was generated for each category as a random −1, +1 string, and then distortions of the prototypes were produced by changing either a large proportion (high-level distortion) or a small proportion (low-level distortion) of the features chosen at random. Each memory trace encoded one of these exemplars together with the appropriate category name. Testing was done by presenting category names as probes to retrieve stimulus features in the echo and by presenting various stimulus patterns as probes to retrieve name features in the echo. In the latter case, the name features in the first echo were correlated with the acceptable category names, and the name giving the highest correlation determined categorization. Test stimuli included the exemplars that had been stored, the prototypes (which had not been stored), new exemplars, and new random patterns. Following a bout of forgetting in which 75% of the trace features were set to zero, the same testing procedure was repeated.

Results of simulations that have been done to date support the following conclusions:

(1) When the category name is used as the probe, the echo closely resembles the category prototype. The larger the number of exemplars stored and the lower their distortion level, the greater is the resemblance. Thus, MINERVA 2 can retrieve an "abstract idea" even though an abstraction as such was never stored.

(2) Classification of old exemplars declines more

rapidly with forgetting than does that of prototypes. This has been the routine finding with humans (e.g., Posner & Keele, 1970). In MINERVA 2, as in MINERVA 1 (Hintzman & Ludlam, 1980), this "differential forgetting" reflects a statistical advantage that accrues to the prototype. After much information has been lost from memory, it is better for a pattern to be moderately similar to all the original category exemplars than for it to be identical to only one.

(3) On delayed as well as immediate tests, classification of old exemplars exceeds that of new exemplars of the same distortion level. Again, this is the routine finding with human subjects.

(4) The prototypes are classified better than their new low-level distortions, which are in turn classified better than new high-level distortions. Thus, the model's classification performance, manifestly based on family resemblance, since the simulations involve no defining features, displays the "typicality" effect characteristic of both natural and artificial concepts.

(5) Transfer to prototypes and new exemplars increases with category size (i.e., the number of stored exemplars). In several experiments by Homa and his colleagues (e.g., Homa, Cross, Cornell, Goldman, & Schwartz, 1973), this has emerged as a reliable result. Also in accord with the literature, random patterns are more likely to be assigned to the larger than to the smaller categories.

(6) Learning with low-level distortions produces better overall classification performance than does learning with high-level distortions. This finding seems to directly contradict conclusions of Homa and Vosburgh (1976) and Posner and Keele (1968). However, both of those studies employed a performance criterion in learning, and subjects who learned with high-level distortions required many more trials to reach the criterion than did those who learned with low-level distortions. Simulations show that if a confounding between distortion level and number of traces per exemplar is allowed, then the direction of the effect of distortion level on the model's performance can be reversed. This inconsistency with human data, therefore, seems to have been only apparent, and not real.

(7) Distortions of the exemplars that were used in learning are classified much as are new distortions of the prototypes. In particular, the effects of distortion level and category size are essentially the same. This finding, obtained with human subjects, was used by Homa, Sterling, and Trepel (1981) to argue against exemplar-based models of classification; but the argument apparently does not apply to MINERVA 2.

(8) If MINERVA 2 is augmented to allow trace activation to be a function not only of similarity to the probe, but also of the similarity of activated traces to each other (as suggested by a "resonance metaphor"), then classification performance improves as a function of the similarity among category exemplars. Elio and Anderson (1981) obtained this result with human subjects, and interpreted it as evidence of abstraction during learning—the idea being that, in addition to traces of the exemplars, a separate representation of their common properties must have been stored. Simulation of the result by the modified version of MINERVA 2 demonstrates another way it might be explained.

PAIRED-ASSOCIATE LEARNING

MINERVA 2 was developed with the frequency-judgment and schema-abstraction tasks in mind; however, since the model is capable of associative learning, it can be applied with little modification to the learning of paired-associate lists. I have explored the model's paired-associate learning performance by having it learn an eight-pair list. The stimulus terms were orthogonal to one another rather than random, and the response terms were also constructed in this way. Fifty learning protocols were generated, using a learning rate of L = .20. Learning was by the anticipation method, and was carried to a criterion of two errorless trials on the list.

To evaluate the model's performance, I fitted Bower's (1961) one-element model to MINERVA 2's output. This simple, all-or-none learning model is ideal for this purpose because it is known to provide excellent fits to many aspects of human paired-associate data and to fail to fit certain others. Thus, by using the one-element model as an intermediary, a fairly rigorous if indirect comparison between simulated and human learning can be obtained.

The findings can be considered to be only preliminary, but they have been most interesting. The fit of the one-element model to MINERVA 2's output was excellent on the following statistics: the learning curve, distribution of total errors per item, distribution of trial of last error, distribution of error run lengths, and mean number of errors following an error on trial n. These are precisely the statistics on which the one-element model most closely matches human data. Poor fits to the simulated data were obtained on the Vincentized "precriterion learning curve" (the one-element model predicts a flat curve, whereas in the simulated data the curve rises and then drops slightly), and on the probability of repeating the same error on successive trials. The one-element model fails to fit human data on these same statistics. The nature of the discrepancy with the one-element model is not exactly the same for the simulated and human data, however. MINERVA 2 perseverates too much. Unlike human subjects, MINERVA 2 may make the same error for 10 or more trials in a row.

MINERVA 2 is not a one-trial learning model, since it adds new information to memory on each trial. Therefore, the close agreement between MINERVA 2 and the Bower (1961) model on so many fine-grained statistics is surprising. With regard to one-trial learning, the lesson seems to be that learning can appear in many respects to be all-or-none even though it is not. With regard to

MINERVA 2, the preliminary conclusion I draw from this exercise is simply that the way the model learns associations should not be rejected out of hand.

CONCLUSION

Definite conclusions based on this brief summary would be premature. However, MINERVA 2 appears to point in promising directions for memory theory. MINERVA 2 is a multiple-trace model, assuming that each experience is encoded in memory no matter how similar it may be to an earlier one. Thus, it is patently a model of episodic memory. However, the model demonstrates a way in which phenomena of episodic memory and generic memory might be explained with a single system. The key assumption is that although a separate trace of each episode is stored, the traces are activated in parallel by a retrieval cue, and abstractions can be retrieved through the summed responses of those traces that are most strongly activated by the cue. Although a theory proposing basically the same process was presented long ago by Semon (1923), among current theories of memory this assumption appears to be unique to MINERVA 2. The model is very simple, and therefore limited in its application. Nevertheless, the promise of the approach is suggested by the range of findings, in recognition and frequency judgments, schema abstraction, and associative learning, which the model seems to explain coherently.

REFERENCES

Bower, G. H. (1961). Application of a model to paired-associate learning. *Psychometrika*, **26**, 255-280.

Brewer, W. F., & Pani, J. R. (1983). The structure of human memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 17). New York: Academic Press.

Eich, J. M. (1982). A composite holographic associative recall model. *Psychological Review*, **89**, 627-661.

Elio, R., & Anderson, J. R. (1981). The effects of category generalizations and instance similarity on schema abstraction. *Journal of Experimental Psychology: Human Learning and Memory*, **7**, 397-417.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, **108**, 356-388.

Hinton, G. E., & Anderson, J. A. (1981). *Parallel models of associative memory*. Hillsdale, NJ: Erlbaum.

Hintzman, D. L. (1969). Apparent frequency as a function of frequency and the spacing of repetitions. *Journal of Experimental Psychology*, **80**, 139-145.

Hintzman, D. L., & Block, R. A. (1971). Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology*, **88**, 297-306.

Hintzman, D. L., & Gold, E. (1983). A congruity effect in the discrimination of presentation frequencies: Some data and a model. *Bulletin of the Psychonomic Society*, **21**, 11-14.

Hintzman, D. L., Grandy, C. A., & Gold, E. (1981). Memory for frequency: A comparison of two multiple-trace theories. *Journal of Experimental Psychology: Human Learning and Memory*, **7**, 231-240.

Hintzman, D. L., & Ludlam, G. (1980). Differential forgetting of prototypes and old instances: Simulation by an exemplar-based classification model. *Memory & Cognition*, **8**, 378-382.

Hintzman, D. L., Nozawa, G., & Irmscher, M. (1982). Frequency as a nonpropositional attribute of memory. *Journal of Verbal Learning and Verbal Behavior*, **21**, 127-141.

Hintzman, D. L., & Stern, L. D. (in press). A comparison of forgetting rates in frequency discrimination and recognition. *Bulletin of the Psychonomic Society*.

Homa, D., Cross, J., Cornell, D., Goldman, D., & Schwartz, S. (1973). Prototype abstraction and classification of new instances as a function of number of instances defining the prototype. *Journal of Experimental Psychology*, **101**, 116-122.

Homa, D., Sterling, S., & Trepel, L. (1981). Limitations of exemplar-based generalization and the abstraction of categorical information. *Journal of Experimental Psychology: Human Learning and Memory*, **7**, 418-439.

Homa, D., & Vosburgh, R. (1976). Category breadth and the abstraction of prototypical information. *Journal of Experimental Psychology: Human Learning and Memory*, **2**, 322-330.

Johnson, M. K., Raye, C. L., Wang, A. Y., & Taylor, T. H. (1979). Fact and fantasy: The roles of accuracy and variability in confusing imaginations with perceptual experiences. *Journal of Experimental Psychology: Human Learning and Memory*, **5**, 229-240.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, **77**, 353-363.

Posner, M. I., & Keele, S. W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology*, **83**, 304-308.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, **7**, 573-605.

Semon, R. (1923). *Mnemic psychology*. London: George Allen & Unwin.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.