21 Ben Murdock and Complexity of Memory

Endel Tulving
University of Toronto

The occasion for the conference in Toronto in June 1990, whose proceedings are reflected in the present volume, was a historic one—celebration of the 65th birthday of a man who for a long time has been playing a decisive role in the science of memory. It is appropriate, therefore, to say a few words on the theme of history, in this case that of the science of memory, and Ben Murdock's role in it, before commenting on the chapters in this section of the volume.

As many of Ben's many fans and friends know by now, he received his graduate training at Yale in clinical psychology, but immediately turned his back on it, and early in his illustrious scientific career found his true love in experimental and mathematical study of memory. His innovative work on short-term memory rapidly established his reputation as a master craftsman whose simple, elegant, and powerful experiments set the standards for others to emulate and helped to change forever the future direction of experimental study of memory.

MURDOCK AND MATHEMATICS

Even before Ben Murdock began publishing his long series of experiments, his penchant for things quantitative showed itself in a paper he published on how to fit exponential curves to data (Murdock & Cook, 1960). In the first paragraph of that paper he proclaimed his credo: The goal of the psychological science is to establish functional relations between antecedent and consequent conditions; these relations are best expressed mathematically, because "a mathematical approach permits an exact and succinct description of the data. More important, a mathematical formulation makes it possible to work out interrelations among variables and deduce specific predictions to be tested experimentally ... Also, such an approach guides theory construction, for a theory must contain within itself all the elements of the mathematical formulation" (Murdock & Cook, 1960, p. 63).

He has remained true to his faith to this day. His devotion to things precise is reflected no better anywhere than in an equation in his famous 1962 paper in the *Journal of Experimental Psychology* whose Figure 1 is one of the most frequently reproduced data sets in experimental psychology, if not the most frequent one (Murdock, 1962). The equation, "an empirical curve, not a rational one," is that of an idealized serial position curve for a 24-word free-recall list. To celebrate Ben Murdock's 65th birthday, and to remind us all of the past glory of the quantitative experimental psychology of memory, it is appropriate to reproduce the equation here:

$$p = 1.00 + .27\, e^{-.77\,(x-1)} - .772\,(.042)^{.555\,(L-x)},$$

where L is list length and x is serial position. Constants are carried to three decimal places, because the precision of the data makes it if not necessary then at least possible!

Ben Murdock's free-recall serial position curves from the 1962 paper were spectacular. It was data like these—beautiful, orderly, and robust—that firmly established a new order in memory research, and spawned a whole generation of dedicated believers in the combined power of experiment and mathematics to lay bare the essence of memory. These data helped to inaugurate the era of optimism, enthusiasm, and confidence that became the hallmark of that domain of experimental psychology that was being transformed into a new powerful movement called "cognitive psychology." Many of those who were there, watching the drama unfold, were swept up by the euphoria of the realization that the human mind was simple and orderly, and that it would yield its secrets readily to determined students armed with the weapons of experimental methods and mathematical symbolism.

Few people looking at Ben's curves could remain unmoved by the striking parallelism of the recency effects from lists varying in length and rate of presentation, and the highly systematic variations in the invariance reflected in each curve's "flat middle section," as Ben modestly put it. Few people looking at these curves could doubt that they reflected anything other than a fundamental regularity of Nature. Orderliness of this kind could not be just a simple consequence of how the experimenter had chosen to tackle a particular problem in his laboratory. Or could it?

MURDOCK AND BABICK

My own favorite Murdock experiment from those early heady days of the golden age of memory research is one that few people know about, an experiment reported in Murdock and Babick (1961). The experiment was simple, elegant, ingeniously conceived; it produced very orderly data; and it carried the Murdock signature of mathematical, as against the conventionally used statistical, treatment of the data. The data showed that mere repetition of an item in the to-be-remembered sets of materials was not a sufficient condition for its enhanced recall. (The reader interested in finding out how Ben did it, over 30 years ago, should feel free to go and read the Murdock and Babick paper.) This at a time when just about everyone still firmly believed that repetition was the major determinant of enhanced performance, and when the forces of the establishment were lining themselves up to fight the heretic notion that associations could be established on a single trial, as had been suggested by Rock (1957) and Estes (1960).

The major message of the Murdock and Babick experiment was not the negative finding of the repetition effect. It was the implication of the null effect: If repetition did not have an effect under the conditions of the Murdock and Babick experiment, but did do so under countless other conditions that had been studied since Ebbinghaus, it must have been so because some of the *other* conditions *necessary* for repetition to have an effect apparently were not satisfied in the Murdock and Babick experiment. In 1961, for students of verbal learning, even for those who were at the threshold of becoming students of memory, the contemplation of the reality of *other* necessary conditions for the repetition effect was a startling event.

A few years later, the influential models of short-term and long-term memory by Waugh and Norman (1965) and by Atkinson and Shiffrin (1968) solved the problem of one-trial learning that had been a source of considerable, even if implicit and usually unmentioned, embarrassment to the previously reigning associative theory. If associations between items are strengthened slowly in the course of trial-by-trial practice, why is the association between two members of a single pair established after a single trial? And given that a typical list consists of single pairs, each of which must be learned on a single trial, is it not the case that a typical learning curve reflects inter-pair interference and the consequent forgetting rather than learning? The problem had been intimated by McGeoch, studiously ignored until Rock (1957) and Estes (1960) came along, and finally solved by the introduction of the concept of primary memory, or the short term-memory store. As it turned out, these short-term models also contained an implicit solution of the puzzle posed by Murdock's and Babick's spectacular finding.

THIRTY YEARS LATER

To at least some of the old-timers in the audience, the conference in Toronto in June 1990 created a certain sense of *déjà vu*: We sensed that we had been there before, but the feeling tone was different. Many speakers went over what seemed to be very much the same ground that the eager mathematical students of short-term memory had been covering over 20 years earlier (Norman, 1970). The models of memory had become more complex, and were now if not "global" then at least comprehensive by the standards of the yesteryear—the SAMs, TODAMs, MINERVAs, and CHARMs. There was discussion of different kinds of priming, text comprehension, holographic memory, and automaticity, topics that were still in the future in 1970; there was even mention of amnesia and different memory systems. What was decidedly absent at the conference was the euphoria, confidence, and sense of purpose of the early days of cognitive revolution.

What happened to the golden age of the study of memory? What happened to the neat data neatly summarized by neat mathematical equations? What had happened to the hope, if not conviction, that Nature will readily yield her secrets about human memory, if questioned intelligently in the form of controlled experiments? Why has the promise of the 1960s not been fulfilled? Where did we, the romantics of the 1960s, go wrong in our expectations?

The real answer to these questions will never be known, because experiments that history performs on humans do not come with necessary control groups. But it is reasonable to assume that an important element in the multiple causation of the developments in the science of memory turned out to be the sheer complexity of the subject matter. We went wrong in our expectations, because we did not quite realize how complex human memory is. Such a realization came as a result of the burgeoning experimental work and theoretical thought lavished upon an ever-expanding set of problems of memory. Thus, only a dozen years after his 1962 paper, Ben ruefully observed that "the human memory system, considered in toto, seems incredibly complex" (Murdock, 1974, p. 1). I vividly remember the quiet after-breakfast walk with Ben at the Psychonomic Society meeting in San Antonio in 1970 when he said, in the course of our spirited conversation about weighty matters concerning our science, that if he could, he would declare a moratorium on all new data gathering, and would make everyone spend some time thinking about how to make sense of the abundant data that were already available. With the relentless onslaught of new findings after findings, it was difficult if not impossible, he said, to come up with theoretical explanations of them.

Ben resolved the conflict between the ever-expanding data base and the need for time for reflection of its deeper meaning in a pure Murdockian fashion: If you cannot stop the world, you can plant your feet firmly on the ground and ignore its whirling. He dedicated the rest of his life to the search of understanding of the part of memory that matters, the part that is or at least seems to be basic and fundamental. The problem of how people learn, retain, and use information about individual items, connections between them, and their serial organization fit the bill.

The complexity of memory, and ensuing difficulties in coming to grips with specific problems is vividly illustrated by Humphreys' and Bain's essay in this volume. Humphreys and Bain discuss, in their characteristically thorough, thoughtful, and balanced fashion, the problem of recognition of a word presented in a study list paired with another word, comparing it with recognition of a word presented by itself.

At first blush, it looks like a straightforward problem. What could be simpler than to establish experimentally the effect of pairwise presentation on recognition of target words, presented at test in different intralist contexts, and then summarize the results in terms of one's favorite model or models? Yet, the upshot of Humphreys' and Bain's discussion is that we are not even nearing what might be regarded as an agreement on what's going on in this kind of a situation, and how it is best explained. What are the processing components of recognition? What components determine the feelings of familiarity? Can the "strength" component be assessed independently of the retrieval environment? Does recognition include recall or retrieval? What kind of retrieval information is matched with what kind of stored information? What role, if any, does the contextual information play in recognition? Different theorists have different answers to these and other similar questions, and the quest for clarification continues unabated.

AUTOMATICITY AND MEMORY

Logan reports some interesting ideas on the general theme that automaticity is memory retrieval, and some interesting data on the automatization of performance on the alphabet arithmetic task. In this task, Logan reminds us, there are at least two ways of answering the questions posed, by computation performed on a previously existing data base in the (semantic) memory store, which is slow and laborious, and by look-up in a newly constructed data base in the (episodic) memory store. As look-up on this task is faster than computation, and a typical subject likes to take the path of least resistance, the look-up method is used whenever possible, that is, whenever there is relevant information in the memory store, and whenever the subject realizes that this is so.

The automatization of the performance in the course of learning a task has an interesting parallel in the memorial consequences of automaticity (Kolers, 1975). Earlier research on this issue can be seen as including the study by Keppel, Postman, and Zavortink (1968) who found that general learning-to-learn effects in the mastery of paired-associate lists, observed over many days of practice, were accompanied by lowered 24-hour retention scores of the learned lists. The suggestion is that increased automaticity of requisite encoding of the material, that is, decreased employment of algorithmic computations, leads to less memorable memory traces. Similar findings, in a different paradigm, have also been reported by Kolers (1975). Another example is found in a report of Jacoby (1978) who contrasted the memorial consequences of computation and look-up in a short-term cued recall task. He found that subjects' subsequent retention of computed information was higher than that of information gained by look-up, and that the effects of distributed repetitions of the same information in a list were greater when subjects had to rely on computation. Comparable results were reported by Slamecka and Graf (1978) in a series of experiments demonstrating what they referred to as the generation effect: the retention of to-be-remembered items constructed out of computations based on existing (semantic) data base at study was superior to that of items "looked up" in lexical memory.

Yet another possible example has to do with the use of mediators in performing tasks such as memorizing pairs or series of words. One can regard the use of mediators at encoding, too, as computations performed on an existing (semantic) data base. Now, an interesting fact about mediators is that they "drop out" with practice. An early example was reported by Balaban (1910); a more recent one comes from the well-known study by Ericsson, Chase, and Faloon (1980) describing the subject S. F. with an "immediate" memory span of 80 digits. It is as if the subject did not need to rely on the mediators any more, even when confronted with what is, from the experimenter's point of view, new material. An alternative interpretation would be that the subject still relies on the mediators but their use has become sufficiently automatized so that the subject is no more consciously aware of doing so.

Now, if look-up, however we conceptualize it, is faster than computation, and its consequences are less memorable, there is a trade-off of the kind that was initially proposed by Craik and Lockhart (1972) in their depth of processing account of memory. The new idea suggested by the studies comparing computations and look-ups, however, is that the processing "depth" is not determined by, or quantitatively specified in terms of, "layers" consisting of surface features, phonological properties, and semantic features of the to-be-remembered items, but rather by and in terms of the overall extent of computation involved in the encoding operation. This idea helps to explain not only data such as those reported by Hyde and Jenkins (1969), Craik and Tulving (1975), and other experiments comparing retention across different encoding domains, but also data such as those described by Mathews (1977), Johnson-Laird, Gibbs, and de Mowbray (1978), and McClelland, Rawles, and Sinclair (1981) in which levels-of-processing effects were analyzed only in the semantic domain.

Logan's work on automaticity, that is, the transformation of (memorable) computation into (less memorable) look-up, when linked to the work on memorial consequences of automaticity, could be seen as suggesting that processing depth varies with the amount of prior practice in computing a given chunk of information. This idea further suggests a possible method of pre-experimentally determining and measuring processing "depth," thereby helping one to escape from the circularity of reasoning that seems to bother some students of memory (e.g., Baddeley, 1978). Thus, for instance, one would expect that repeated prior computations of the answer to a specific semantic question about a particular target item would lead to the automatization of the computations involved and hence a less memorable memory trace, in relation to comparable nonsemantic computations, in experiments of the kind reported by Craik and Tulving (1975).

PRIMACY AND RECENCY IN TEXT COMPREHENSION

Walter Kintsch has made a name for himself as perhaps the most perceptive, inventive, and systematic student of text comprehension of our time. In the chapter that he and David Welsch have contributed to the present volume, the attractive features of his approach are once more skillfully demonstrated. His construction-integration model is constructive and integrative in more than one sense. In one sense, the terms in the appellation of the model refer to the major theoretical concepts that play a major role in Kintsch's search for the theoretical understanding of text comprehension. In another, it could also be seen as characterizing Kintsch's overall approach to the daunting problems that such understanding entails. Kintsch is a great believer in the continuity of the cognitive science and his modeling of text comprehension has always reflected this belief. He makes good use of relevant facts provided by previous research, even if the research may have been conducted under a "wrong" label, and he builds on top of past achievements, rather than assiduously ignoring them, as some others are wont to do.

Kintsch's and Welsch's treatment of the data on first mention versus recency in sentence processing nicely illustrates the continuity between list-learning experiments and text comprehension, between the past and the present. It was known from list learning experiments that the ubiquitous recency effects in immediate free recall or immediate paired-associate recall would be transformed into negative recency effects after a longer retention interval (Craik, 1970; Madigan & McCabe, 1971). Thus, over the last few list items, immediate recency effects are converted into delayed primacy effects. This phenomenon of recency-to-primacy conversion is especially clearly demonstrated in the studies by Wright, Santiago, Sands, Kendrick, and Cook (1985), who measured serial position curves for yes/no recognition of lists of four verbally noncodable items, and found the conversion in pigeons, monkeys, and people. In the studies reported by Gernsbacher, Hargreaves, and Beeman (1989) a similar conversion was observed in the processing of the first-mentioned and second-mentioned participant of an action described by the sentence. Although the immediate versus delayed shift in the sentence processing covered a much shorter time interval (perhaps attributable to the priming measure), the main structural features of the phenomenon are identical with those reported in list learning studies. Inasmuch as the Wright et al. (1985) experiment clearly suggests that the phenomenon can occur independently of language, it seems appropriate to conclude that it represents a rather fundamental feature of the biological memory systems that, among other things, seems to manifest itself in the processing of meaningful sentences.

This kind of empirical generality must be regarded as the ultimate objective of experimental research; to the extent that it is achieved, the researchers are justified in believing that they are really communicating with Nature. The fact that the construction-integration model can account for the phenomenon, and do so naturally, is a theoretical achievement of which the theorists can be justifiably proud.

LIST STRENGTH IN RECALL AND RECOGNITION

Shiffrin and Murnane's essay in this volume tosses the gauntlet to the authors of composite memory models: These models, by postulating inter-item interference at encoding, cannot, Shiffrin and Murnane claim, account for the patterns of results showing prominent list-strength effects in free recall, detectable albeit weaker effects in cued recall, but little trace of the effect in recognition, under conditions where encoding conditions are held constant (Ratcliff, Clark, & Shiffrin, 1990). It would be great, of course, if we possessed firm data that clearly exclude certain classes of theories, because nothing helps purify theory more than such exclusion (Platt, 1964). But psychological theories of memory, as all psychological theories, tend to be flexible, their fate depending more on their keepers' faith and scientific style than on empirical facts. This is why it is only possible to recommend that interested observers of the scene remain tuned in for future developments on the dialogue between the believers in separate storage and in distributed composition.

Shiffrin and Murnane express mild surprise about the fact that in a number of experiments *negative* list-strength effects have been observed in recognition, that is, that recognition of "weak" items is relatively higher in mixed lists than in pure lists. To at least some of those who participated in the fun of memory research in an earlier era, however, an equally mildly surprising finding would be the *absence* of similar negative list-strength effects in free recall, in experiments as yet not done in which the strength of the list items has been manipulated in terms of differential within-list study times. Such a negative effect could be predicted on the basis of the data reported by Waugh (1970) before the list-strength effect was discovered. Waugh, in Experiment 3, presented individual study-list items (monosyllabic words) at within-list rates varying from 1 sec to 8 sec, and found no effect whatsoever: Item presented for 1 sec was recalled as well as all other items, including those presented at 8 sec. Given this lack of any strength differences in a mixed list, and traditional strength effect in pure lists, the conclusion follows that a Waugh-type comparison of mixed and pure lists would yield a large negative list-strength effect even in free recall. Of course, the experiments with which Shiffrin and Murnane are concerned did differ from Waugh's in a number of ways, and the precautions taken against "rehearsal borrowing" in those experiments may have had the desired effect of creating some strength differences between weak and strong items. Nevertheless, it would probably not be amiss in future research to consider possible differences in item strengths created through manipulations of study time versus frequency of presentation.

CONCLUSION

Tempora mutantur et nos mutamur in illis.

AUTHOR NOTE

Endel Tulving's research is supported by Grant A8632 of the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory, Vol. 2* (pp.89-105). New York: Academic Press.

Baddeley, A. D. (1978). The trouble with levels: A re-examination of Craik and Lockhart's framework for memory research. *Psychological Review, 85*, 139-152.

Balaban, A. (1910). Über den Unterschied des logischen und mechanischen Gedächtnisses. *Zeitschrift für Psychologie, 56*, 356-377.

Craik, F. I. M. (1970). The fate of primary memory items in free recall. *Journal of Verbal Learning and Verbal Behavior, 9*, 143-148.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*, 268-294.

Ericsson, K. A., Chase, W. G., & Faloon, S. (1980). Acquisition of a memory skill. *Science, 208*, 1181-1182.

Estes, W. K. (1960). Learning theory and the new "mental chemistry." *Psychological Review, 67*, 207-223.

Gernsbacher, M. A., Hargreaves, D. J., & Beeman, M. (1989). Building and accessing clausal representations: The advantage of first mention versus the advantage of clause recency. *Journal of Memory and Language, 28*, 735-755.

Hyde, T. S., & Jenkins, J. J. (1969) Differential effects of incidental tasks on recall and organization of highly associated words. *Journal of Experimental Psychology, 82*, 275-292.

Jacoby, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning and Verbal Behavior, 17*, 649-667.

Johnson-Laird, P. N., Gibbs, G., & de Mowbray, J. (1978). Meaning, amount of processing, and memory for words. *Memory & Cognition, 6*, 372-375.

Keppel, G., Postman, L., & Zavortink, B. (1968). Studies of learning to learn: VIII. The influence of massive amounts of training upon the learning and retention of paired-associate lists. *Journal of Verbal Learning and Verbal Behavior, 7*, 790-796.

Kolers, P. A. (1975). Memorial consequences of automatized encoding. *Journal of Experimental Psychology: Human Learning and Memory, 1*, 689-701.

Madigan, S. A., & McCabe, L. (1971). Perfect recall and total forgetting: A problem for models of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 10*, 101-106.

Mathews, R. C. (1977). Semantic judgments as encoding operations: The effects of attention to particular semantic categories on the usefulness of interitem relations in recall. *Journal of Experimental Psychology: Human Learning and Memory, 3*, 160-173.

McClelland, A. G. R., Rawles, R. E., & Sinclair, F. E. (1981). The effects of search criteria and retrieval cue availability on memory for words. *Memory & Cognition, 9*, 164-168.

Murdock, B. B., Jr. (1962). The serial position effect in free recall. *Journal of Experimental Psychology, 64*, 482-488.

Murdock, B. B., Jr. (1974). *Human memory: Theory and data.* Potomac, MD: Lawrence Erlbaum Associates.

Murdock, B. B., Jr., & Babick, A. J. (1961). The effect of repetition on the retention of individual words. *American Journal of Psychology, 74*, 596-601.

Murdock, B. B., Jr., & Cook, C. D. (1960). On fitting the exponential. *Psychological Reports, 6*, 63-69.

Norman, D. A. (Ed.). (1970). *Models of human memory.* New York: Academic Press.

Platt, J. R. (1964). Strong inference. *Science, 146*, 347-353.

Ratcliff, R., Clark, S., & Shiffrin, R. M. (1990). The list-strength effect. I. Data and discussion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 163-178.

Rock, I. (1957). The role of repetition in associative learning. *American Journal of Psychology*, *70*, 186-193.

Slamecka, N., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, *4*, 592-604.

Waugh, N. C. (1970). On the effective duration of a repeated word. *Journal of Verbal Learning and Verbal Behavior*, *9*, 587-595.

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

Wright, A. A., Santiago, H. C., Sands, S. F., Kendrick, D. F., & Cook, R. G. (1985). Memory processing of serial lists by pigeons, monkeys, and people. *Science*, *229*, 287-289.

VI DISTRIBUTED APPROACHES TO MEMORY

Relating Theory and Data: Essays on Human Memory in Honor of Bennet B. Murdock

Edited by
William E. Hockley
Wilfrid Laurier University
Stephan Lewandowsky
University of Oklahoma