AFHRL-TP-87-73

AIR FORCE

HUMAN RESOURCES

AD-A198 757

DTIC FILE COPY

SKILL MAINTENANCE: LITERATURE REVIEW AND THEORETICAL ANALYSIS

David W. Fendrich
Alice F. Healy
Lori Meiskey
Robert J. Crutcher
William Little
Lyle E. Bourne, Jr.

Institute of Cognitive Science
University of Colorado
Boulder, Colorado 80309-0345

TRAINING SYSTEMS DIVISION
Brooks Air Force Base, Texas 78235-5601

DTIC
ELECTE
AUG 1 5 1988
S D
D

August 1988
Final Technical Paper for Period September 1985 - September 1986

Approved for public release; distribution is unlimited.

LABORATORY

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235-5601

88 8

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFHRL-TP-87-73
6a. NAME OF PERFORMING ORGANIZATION Institute of Cognitive Science 6b. OFFICE SYMBOL (If applicable)	7. NAME OF MONITORING ORGANIZATION Training Systems Division
6c. ADDRESS (City, State, and ZIP Code) University of Colorado Boulder, Colorado 80309-0345	7b. ADDRESS (City, State, and ZIP Code) Air Force Human Resources Laboratory Brooks Air Force Base, Texas 78235-5601
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Air Force Human Resources Laboratory 8b. OFFICE SYMBOL (If applicable) HQ AFHRL	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER VE5704-022-001
8c. ADDRESS (City, State, and ZIP Code) Brooks Air Force Base, Texas 78235-5601	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO: 62703F; PROJECT NO: 7719; TASK NO: 18; WORK UNIT ACCESSION NO: 40

11. TITLE (Include Security Classification)
Skill Maintenance: Literature Review and Theoretical Analysis

12. PERSONAL AUTHOR(S)
Fenrich, D.W.; Healy, A.F.; Meiskey, L.; Crutcher, R.J.; Little, W.; Bourne, L.E., Jr.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM Sep 85 TO Sep 86	August 1986	34

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	automatism, long-term retention, theoretical analysis, cognitive psychology, permastore, Defense mechanism (Psychology), literature review, skill maintenance,
05	09		
06	09		

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The literature on the topic of skill maintenance is reviewed and a theoretical analysis is provided. The review includes studies on the retention of knowledge, intratask interference, retention of motor skills, discovery learning, and decay of skills. The proposed theoretical mechanisms underlying skill retention include automatism, division of attention, overlapping of processes, task restructuring, proceduralization of knowledge, levels of processing, and transfer-appropriate processing. Also discussed are the generation effect and the role of consciousness in skill retention. Keywords:

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Nancy J. Allin, Chief, STINFO Office	22b. TELEPHONE (Include Area Code) (512) 536-3877 22c. OFFICE SYMBOL AFHRL/TSR

DD Form 1473, JUN 86 Previous editions are obsolete SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

AFHRL Technical Paper 87-73

August 1988

SKILL MAINTENANCE: LITERATURE REVIEW AND THEORETICAL ANALYSIS

David W. Fendrich
Alice F. Healy
Lori Meiskey
Robert J. Crutcher
William Little
Lyle E. Bourne, Jr.

Institute of Cognitive Science
University of Colorado
Boulder, Colorado 80309-0345

TRAINING SYSTEMS DIVISION
Brooks Air Force Base, Texas 78235-5601

Accesion For
NTIS CRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution /
Availability Codes
Dist | Avail and / or Special
A-1

QUALITY INSPECTED 2

Reviewed and submitted for publication by

Hugh L. Burns, Lt Col, USAF
Chief, Intelligent Systems Branch
Training Systems Division

This publication is primarily a working paper. It is published solely to document work performed.

SUMMARY

In contemporary society many resources are devoted to skill training. There is a large body of psychological research aimed at determining the optimal training procedures. However, relatively few of these studies have examined truly long-term retention of skills. The present paper reviews the existing literature on long-term retention and also examines a number of psychological mechanisms which have been demonstrated to be crucial in studies of relatively short-term learning and should have important implications for long-term retention. For example, on the basis of recent research in cognitive psychology, it is conjectured that maintenance of a skill will depend crucially on whether that skill is automatic; that is, can be performed without conscious awareness. Some of the additional mechanisms, including intratask interference and transfer-appropriate processing, point to the importance of the relationship between original learning and final test conditions.

PREFACE

This paper presents a review of the literature on skill maintenance as well as a theoretical analysis of that subject. This paper is meant to provide the framework with which we could begin a program of research on the maintenance of skill. A companion piece (Healy et al., 1986), separately published, provides some sample methodologies we have developed to investigate this opic. We wish to acknowledge the help of Antoinette Gesi and Debbie Aguiar in the preparation of this paper.

TABLE OF CONTENTS

	Page
Review of Studies of Skill Retention	1
Retention of Knowledge	1
Intratask Interference	3
Retention of Motor Skills	6
Discovery Learning	8
Decay of Skills	10
Proposed Theoretical Mechanisms Underlying Skill Retention	11
Automatism	11
Alternatives to Automatism	15
Additional Theoretical Mechanisms	18
Conclusions	21
REFERENCES	22

Table		Page
1	Twelve Selected Experiments on Intratask Interference	4

SKILL MAINTENANCE: LITERATURE REVIEW AND THEORETICAL ANALYSIS

In contemporary society, enormous resources are devoted to education and skill training. The average person spends about 12 years in school, followed by either more formal education, military training, or technical training. Given the magnitude of time and resources devoted to the learning of knowledge and skills, there is considerable merit in efforts to ensure the effectiveness of such training. Regardless of the skill domain, the goal of training is to develop a high level of performance on some category of jobs or tasks, ensuring that the requisite skills are available for application as needed subsequently in the field. Psychological research has provided an extensive database on which effective training procedures can be designed. The generality of this body of research is, however, limited by the fact that relatively few studies have examined truly long-term retention.

The purpose of this paper is to review the available literature on long-term retention. Further, we will examine a number of psychological mechanisms which may have important implications for long-term retention.

Review of Studies of Skill Retention

Retention of Knowledge

The most important recent work in this area has been conducted by Bahrick (1979, 1984) and, therefore, this work will be reviewed in some detail. Bahrick (1979) investigated the acquisition and maintenance of various forms of knowledge over retention intervals considerably longer than those typically used in laboratory experiments. He developed two different methods to evaluate the extent to which maintenance of knowledge depended on periodic access to the learned material. The first method involves a cross-sectional statistical procedure. The use of this method requires the availability of a large number of subjects who acquired the same knowledge at different times in the past. It is also essential that the subjects be able to estimate the amount of rehearsal of that knowledge during the interval between original acquisition and the retention test. The degree of original acquisition of the material for each subject must also be obtainable. Subjects are then assigned to groups depending on when they acquired the information, and a retention function is calculated based on their test performance. This retention function is then corrected, using multiple regression techniques, for factors contributing to the original level of performance and the extent of rehearsal or practice during the retention interval. Bahrick used this technique to study the maintenance of knowledge about the names and spatial locations of buildings and streets within a university city. He concluded from his results that spatial information was lost more rapidly than information about names but that the two types of information could be recovered with equal ease. Specifically, on the average, 7-8 visits per year were necessary to maintain spatial information about street sequence and 8-9 visits per year were necessary for maintaining the street names at the level of graduating seniors. Trade-off estimates were also reported evaluating the relative importance of the recency, duration, and frequency of the visits when the knowledge was rehearsed or refreshed.

In a similar cross-sectional study, Bahrick (1984) administered a large battery of tests of knowledge of Spanish to more than 700 participants whose last exposure to a course in Spanish ranged from 0 to almost 50 years. Participants were also selected and grouped according to their final level of

training in Spanish. Questionnaires concerning the amount and recency of practice and exposure to Spanish were administered to the participants in order to obtain estimates of the amount of rehearsal necessary to maintain knowledge over time. Rehearsal levels during the retention interval were found to be quite low and unrelated to the level of knowledge retention. Reliable predictors of long-term retention were training level, mean grade in Spanish courses, and the level of training in other romance languages. The retention function over 50 years indicated that memory level dropped for about 6 years to a stable asymptote and then remained constant for about 50 years, at which time it again started to drop (possibly because of neurological deterioration attributable to old age). Higher levels of original training and higher mean grades were independently associated with overall higher retention. Because very little rehearsal was reported by the participants, it appeared that much of the information originally learned was retained in a "permastore" and that it remained accessible for many years without periodic maintenance activities. Because the rehearsal predictors were unrelated to retention in this study, Bahrick concluded that most of what influences the amount of long-term retention is determined by acquisition processes and not by rehearsal effects during the retention interval. This conclusion must be viewed as tentative, however, since it is not known whether greater amounts of rehearsal during the retention interval would have increased long-term retention. This rich method is particularly attractive for investigating skill maintenance because it is applicable to skills learned outside the laboratory under ecologically realistic conditions.

The second method, developed by Bahrick (1979) to study the maintenance of knowledge, is conducted solely in the laboratory and involves successive relearning sessions. In the original learning session, the information to be acquired is tested and repeated with a dropout technique which ensures that every item receives the same number of correct responses. Subsequent relearning sessions start with a test of all the items originally learned and then continue with the dropout technique used in the original learning session. Among other materials, Bahrick applied this technique to the learning of English-Spanish word pairs. He systematically varied the intersession interval (from 0 to 30 days) and examined performance during the original learning session, in the two or five subsequent relearning sessions, and at a final test session which followed 30 days after the last relearning session. He found that the information was maintained at a high level across lengthy intersession intervals, and that performance in the final test session depended more on the earlier intersession intervals than on the level of performance reached in the last relearning session. Performance in the last relearning session was greatest when the interstimulus intervals were the shortest. But performance in the final test session was greatest when the earlier interstimulus intervals were the longest so that they matched the interval between the last relearning session and the final test session. Bahrick concluded from these results that for optimum maintenance of knowledge, practice should be spaced at intervals not much shorter than the interval separating practice from test.

The most important point of this study with respect to the maintenance of skills is that people who exhibit the same level of skill proficiency immediately after training can differ substantially on long-term retention, depending on their previous training history. From this result, we must reject the intuitive notion that criterion performance at the end of training alone is a sufficient predictor of long-term retention. Based on studies of intratask

interference, Battig (1979) provided strong convergent evidence for this viewpoint, and his research efforts will be discussed below.

Intratask Interference

On the basis of an earlier extensive review of the literature, Battig (1979) proposed a general intratask interference principle of memory: Greater difficulty or interference at the time of learning produces higher levels of subsequent long-term retention and transfer of what has been learned. There have been enough different demonstrations of this phenomenon, including those cited in Battig's review and others which have since been published (e.g., Shea & Zimny, 1982), to conclude that intratask interference is a potent contributor to permanence in memory.

We cite in Table 1, 12 articles that illustrate the methods and results of previous research on intratask interference. This list is by no means exhaustive. The reader is referred to Battig's (1979) original article for a much more extensive bibliography. The first thing to note is that whereas these studies represent a wide range of problems and procedures, they tend to center on verbal tasks. Only the studies by Hiew (1977), using concept formation problems, and Shea and Morgan (1979), using a serial motor response task, fall outside the strictly verbal domain. Thus, one concern of research is the generality of this principle to motor skills tasks and tasks that have a greater perceptual or problem-solving component.

Table 1 lists these illustrations of intratask interference effects. Each listing includes a brief description of the interference-producing manipulation within each experiment, followed by three numerical values, the first two of which represent percent of maximum retention or transfer under low and high interference conditions. Due to marked variations in overall performance levels across the various conditions of experiments, a derived %MAX measure is given in the final column, which represents the difference between the high and low values divided by the maximum possible difference (100% or perfect performance minus the low value) times 100. In other words, %MAX represents the percentage of maximum facilitation found under the high interference conditions, so that larger %MAX values are indicative of greater magnitudes of the intratask interference phenomenon.

This listing of experiments is intended primarily to document the wide range of experimental evidence consistent with Battig's principle. It should be obvious from even a cursory examination of the various experimental manipulations employed in these experiments that any attempt to summarize or make meaningful comparisons across experiments is premature. Nonetheless, a few tentative statements about these results are possible. First, intratask interference effects are generally sizeable, averaging over one-third of the maximum facilitation that would be possible for high relative to low interference conditions. Second, these studies appear to define two distinct bases for interference, viz., similarity of the items composing the learning unit and variations in the context in which the learning unit appears. For example, Experiments Number 1, 2, 3, and 4 in Table 1 represent variations primarily in the context in which any given item to be learned is presented for study or test. In contrast, Experiments Number 5, 6, 7, and 8 appear to qualify as manipulations primarily of the degree to which items constituting a verbal list or some other set of material to be learned are similar to one another. Although none of these demonstrations included independent manipulations of both

Table 1. Twelve Selected Experiments on Intratask Interference

Description of Experiment	Percentage Retention or Transfer Low	High	%Max
Contextual Variation			
1. Battig (1972) Separation vs. mixing of paired-associate pairings and recall tests	48.3	60.4	23.4
2. Hiew (1977) Blocked vs. mixed training on sets of conceptual rules	15.3	8.5	44.4
3. Nitsch (1977) Constant vs. varied exemplars of verbal contexts	34.1	79.5	68.9
4. Underwood & Lund (1979) Simultaneous vs. sequential list learning	58.5	77.3	45.3
Similarity of Items			
5. Chiesi (1976) Paired-associate response conceptual similarity	50.3	63.2	26.0
6. Pagel (1973) Paired-associate stimulus formal similarity (also stimulus meaningfulness)	8.3	23.8	16.9
7. Pellegrino (1972) Paired-associate response formal similarity	79.9	90.9	54.7
8. Posnansky (1974) Serial intralist similarity	54.6	95.0	89.0
Combined Manipulations			
9. Einstein (1976) Extraneous processing of related words during free recall learning	52.0	68.7	34.8
10. Johnson (1964) Shape-label similarity and congruence	31.9	42.2	15.1
11. Schab (1975) Presentation of related paired-associate lists on alternate trials	57.1	71.4	33.3
12. Shea & Morgan (1979) Blocked vs. random trial sequences in motor skills task	4.3	2.1	51.5
Mean			34.9

the learning unit and its context, Experiments Number 9, 10, 11, and 12 combine these factors.

Comparisons of the %MAX values in the last column of Table 1 show substantial magnitudes of intratask interference for all three classes of experiments. The evidence available fails to give us clear indications, however, of the relative magnitudes of effects attributable to contextual variation and to similarity of the items in the learning unit, nor does it tell us whether a combination of manipulations would produce additive or some other systematic effects. These experiments argue strongly for the need to produce separable systematic variation in both learning unit similarity and contextual variety.

From the foregoing review, intratask interference effects can be conceived as a consequence of (a) greater amounts, elaboration, or distinctiveness of processing and the resulting encodings of items, required under conditions of high interference or difficulty in order to achieve satisfactory short-term performance and/or (b) the development of encodings that are resistant to interference and thus more likely to be remembered under the changed contextual conditions that typically characterize delayed retention or transfer tests. These results imply that whether or not the intratask interference phenomenon will be found is heavily dependent upon degree of encoding-retrieval congruence (the extent to which the memory or transfer tasks are consistent with those of the original learning situation). This, of course, follows directly from Tulving's Encoding Specificity Principle (Tulving & Thomson, 1973), as well as from the results of classical research on intertask transfer showing that any substantive change from the original learning or encoding conditions typically produces marked decrements in memory or transfer (Battig, 1979).

There is a further implication of the research described above that changes in the memory or transfer conditions under which intratask interference is evaluated should show very different magnitudes of effect, depending on the specific sources of acquisition interference or difficulty. It is this line of reasoning that forms the basis for the distinction between variety of the context and similarity in the learning unit as separable sources of the phenomenon. Contextual variety refers to changes over repeated trials (or encounters) in the processing context under which a given task must be learned. Since the primary effects of greater contextual variety should be to produce more elaborate and distinctive encoding, such contextual variety should lead to greater resistance to the normally negative effects of changes at the time of retention or transfer tests. Indeed, Shea, Hunt, and Zimny (1985) reported that the verbal protocols of subjects trained on a serial motor skill under contextual variability contain many more references to distinctive and contrastive encodings than do those of subjects trained under low variability. Stated another way, encoding specificity can most effectively be overcome if the original encodings have taken place under high contextual variety. Contextual variety, however, necessarily implies at least some noncorrespondence between the acquisition and the retention or transfer testing conditions, and thus should be relatively less effective when these conditions are maximally alike. Consequently, increased contextual variety should result in the intratask interference phenomenon primarily when context conditions are changed from the acquisition to the retention or transfer test, with the phenomenon reduced or eliminated when acquisition conditions are maintained on the later test.

Similarity among the items of the learning unit is a source of acquisition interference that induces additional processing consisting primarily of the

formation of organizational and discriminative transformations appropriate to the specific task requirements. Such additional processing, as is induced by item similarity, should be effective primarily under those specific similarity conditions, and is likely to be of little or no value if the learning unit is markedly changed. This leads directly to the hypothesis that increased item similarity should result in better transfer or retention primarily when the learning unit is the same for retention (or transfer) as under acquisition, and the transfer effects should be systematically reduced or eliminated with tests which incorporate major changes from the original acquisition conditions.

Thus, we have one type of intratask interference (contextual variety) which should produce positive effects on transfer and retention primarily under changed retention-transfer conditions, and another (learning unit similarity) which should result in positive effects mainly where there is minimal change from he acquisition to the retention or transfer conditions. Consequently, a combination of both increased contextual variety and learning unit similarity should produce a greater positive effect over a wider range of retention or transfer testing conditions than either of these alone. We can hypothesize, therefore, that intratask interference representing a combination of learning unit similarity with contextual variety should produce a greater positive effect on transfer and retention than should either of these factors alone, under either identical or changed conditions between original acquisition and subsequent testing.

Retention of Motor Skills

Practice methods. There is considerable interest in the realms of both verbal and motor learning in the effects blocked and random acquisition trials have on retention. In the study of these two practice methods, a curious paradox surfaced. As reviewed above, Battig, as early as 1966, found that in verbal learning tasks, contextual interference, caused by randomly ordering the subjects's practice regimen, tended to hamper the acquisition of a task, yet facilitate its retention. This paradox was also found to be present in motor learning tasks (Shea & Morgan, 1979). Shea and Morgan's task, however, confounded practice schedule effects (blocked vs. random) with reaction-time paradigm effects (simple—inherent in blocked practice vs. choice—inherent in random practice).

With this problem in mind, Lee and Magill (1983) devised three experiments to avoid these confounding effects yet came up with the familiar paradoxical results. The task used was a hinged barrier course similar to that used by Shea and Morgan. Possible reaction-time confounding was eliminated by using cued and uncued practice trials and also by adding a third serial practice group. The previous findings of Battig (1966) and Shea and Morgan (1979) were upheld in all three experiments, with both serial and random practice outperforming blocked practice.

The suggestion here is that a substantially weaker memory results from blocked practice compared to random and serial. The reason for this inferior memory might lie the relative predictability or unpredictability of the practice schedule. In a predictable (blocked) regimen, the subject relies primarily on the predictability as the prime reinforcer, whereas in an unpredictable (random) regimen, the subject must utilize other sources of information to accomplish the task. With predictability removed, as in retention trials, the random practice group will excel and the blocked practice group, without their "crutch," will flounder. This explanation for the locus of contextual interference is viable and is right in line with Schmidt's (1975)

constant or variable knowledge of results (KR) theory as a predictor of motor task retention. Here, KR represents feedback from the experimenter as to how close the subject came to achieving the task goal. This theory predicts that subjects receiving KR on every trial, or on a constant schedule, come to depend too heavily on this source of feedback for learning. Subjects who do not receive KR or every trial but rather, on a variable schedule, are forced to utilize other forms of feedback (e.g., visual, auditory, kinesthetic). Consequently those groups that have learned to use forms of feedback other than KR perform better in retention tasks when KR has been withdrawn.

Schema theory. In assessing skill retention, one person's performance or learning (as measured by an appropriate transfer test) is usually compared to another person's performance or to some standardized performance criterion. Unfortunately, such comparisons provide no information as to what training variables lead to performance differences. Using different types of training equipment, for example, may produce different learning outcomes. But more importantly to those in the business of training, given the same training equipment, which training methods result in superior retention?

One important issue concerning type of training is whether training should be constant or variable. Some (e.g., Adams, 197 aintain that constant training or overpractice of a criterion movement leads to stronger memory and perceptual traces, thereby ensuring that later the learner will be able to consistently and accurately perform the criterion movement. Others (e.g., Schmidt, 1975) maintain that variable practice or training is more effective and invoke a schema theory to support their claim. According to Schmidt, in learning a movement skill, an individual stores dif erent types of information (e.g., initial conditions, response specifications, sensory consequences, and response outcomes) and uses this information to form a relational schema, or cognitive representation of the task. Later, this schema can be used to produce the required movements in a task. The larger or more diverse the base of information forming the schema, the stronger are the informational relationships, the schema itself, and hence, the probability of producing the criterion task. To achieve this larger or more diverse base of information, Schmidt recommends increasing the variability of training or practice around the criterion movement--an approach opposite to that recommended by Adams of practicing the criterion movement only.

Long-term retention of motor skills. Because comprehensive literature reviews on the long-term retention of motor skills are available (e.g., Naylor & Briggs, 1961; Schendel, Shields, & Katz, 1978), this section will focus on summarizing the major findings of this field and relating them to the retention of perceptual and cognitive skills.

Schendel et al. (1978) have concluded that the level of original training is the single most important determinant of motor skill retention. This general conclusion was also reached by Bahrick (1984) in his study on the long-term retention of knowledge concerning Spanish. Additionally, mastery training (overlearning) can increase long-term retention of motor skills beyond the level attained with proficiency (criterion) training. Mastery training has also been shown to counteract the negative effect of high levels of arousal or anxiety, an important consideration for skills needed in emergency situations.

Refresher training and mental rehearsal have been shown to aid the long-term retention of motor skills. The effect of re-exposure to a skill not only reinforces knowledge previously learned but can also produce new learning. Although Bahrick (1984) did not find evidence for an effect of rehearsal on the

retention of the Spanish language, the levels of rehearsal reported may have been too low to produce an enhancement of retention.

The individual ability of subjects in a training program can be a good predictor of speed of learning and retention. That is, subjects with high ability will learn a skill to a performance criterion faster and retain it at a higher level than will subjects with lower ability levels. It should be noted, however, that the better retention of high ability subjects may be due to differences in the degree of original learning, and that if all subjects are trained to the same degree, long-term retention may be unrelated to initial ability differences.

An important consideration in the evaluation of skill retention is the manner in which it is measured. If retention is measured from the first performance of a skill at the end of the retention interval, subjects will, in general, show poor retention. However, performance improves quickly after the first retention testing and if retention is indexed by the speed of relearning to criterion performance, subjects will show a higher level of retention. In other words, after a long period of disuse, skill performance will initially be "rusty," but after a short period of practice, performance will return to its original level. Thus, an important consideration for retention training is whether or not a period of relearning is feasible. If relearning is not feasible, as in emergency situations, then a more intensive retention training program should be considered to ensure a high level of performance in the first execution after a long retention interval.

The marked difference between retention measured from the first re-exposure and the speed of relearning forces us to qualify the concept of permastore. Bahrick's conception of permastore refers to the former measure of retention; that is, the information which can be accessed without any relearning. Thus, permastore represents the "tip of the iceberg" of overall retention because much of what is retained manifests itself only after some relearning.

Schendel et al. (1978) concluded that the distribution of practice (massed vs. spaced) does not appear to influence the retention of motor skills. This finding is in marked contrast to those from studies showing strong effects of practice distribution in many cognitive skills and, in particular, to the work of Bahrick (1979). A possible explanation for this discrepancy pertains to the conclusion of Bahrick that the spacing of learning sessions should be as long as the retention interval. Many of the studies cited by Schendel et al. as not showing a practice distribution effect for motor skills compare immediate trial repetitions to ones spaced only seconds apart. Retention was then tested after a relatively long retention interval (e.g., 1 month). Perhaps the spacing of the learning trials was too short relative to the retention interval to produce a detectable difference in retention. Further research is needed to examine this possibility before practice distribution is discarded as a consideration in a skill retention program.

Discovery Learning

Are procedures better retained or more likely to appear in permastore after discovery or after expository learning? Which method of learning would provide for better transfer of knowledge to new problems? These are questions which address the long-debated issue of whether discovery or expository learning is better (i.e., makes information easier to retrieve from memory and to transfer to new situations). Bruner (1961) argued that discovery learning is beneficial for a variety of reasons, such as the fact that discovery causes the learner to organize the material, which would lead to better retrievability than if the material were rote memorized. Also, discovery would allow the learner to form

an hierarchically higher-level conceptualization of the material; it also might cause the learner to find alternate solutions to a problem rather than merely the "right way to do it," which would provide flexibility in problem-solving ability and would aid in transfer to different situations. Friedlander (1965), on the other hand, provided a number of arguments against the superiority of discovery learning. He noted that there are skills and facts which must be memorized before a larger body of knowledge can be learned. Skills and facts, after being memorized, may yield higher-level conceptualizations. Also, discovery learning usually takes more time than does expository learning, and the learner may be led to wrong or inappropriate conclusions.

Richard Mayer, James Greeno, and their colleagues (see Mayer, 1975, for a review) attempted to test this question systematically and found that both Bruner and Friedlander are correct. In most of these studies, subjects learned to solve various probability problems, which could be represented in equation form (e.g., binomial probability, Bayes theorem). Egan and Greeno (1973) had subjects solve problems and generalize with few instructions (the discovery group), or had subjects solve problems and gave them instructions (the expository, or rule, group). Before training, all subjects were given pre-tests to test their knowledge of probabilistic concepts and their ability to solve computations and use permutation strategies. Subjects were tested after training. Measures of performance were number of errors and speed of solution. There were no overall differences between the discovery and rule groups on the post-test. However, subjects in the rule group performed uniformly well on the post-test, but subjects in the discovery group performed better on the post-test if they had higher ability (as measured by the pre-tests). Also, low ability subjects performed better in the rule group than in the discovery group. Egan and Greeno postulated that the discovery subjects were forced to invoke past knowledge in order to solve the problems, and subjects with greater ability were better able to do this. On the other hand, subjects in the rule group were adding a piece of information which was relatively independent of their past knowledge, so previous abilities and knowledge were less important in determining their performance.

Other experiments (Mayer, 1974; Mayer & Greeno, 1972; Mayer, Stielh, & Greeno, 1975) used two expository training groups. In one group, general concepts familiar to the subjects were used to teach subjects to solve the probability problems, and thus, their general knowledge was invoked in the learning process. In the other group, subjects were given the appropriate probability formula and practiced using this formula; the formula was a piece of information which subjects did not need to integrate with past knowledge. The results were the same as for the discovery and rule groups in the Egan and Greeno (1973) experiment. Previous ability was important for subjects in the general concept group's performance, and ability was not significantly related to performance for subjects in the rule group. They also found that subjects trained with the formula performed better than subjects trained with general concepts on near transfer problems (problems similar to training problems), but subjects in the general concept group performed better on far transfer problems (problems requiring interpretation in order to be able to use the information gained in training). They found that these differences in transfer ability occurred early in learning and remained consistent throughout greater levels of training. Mayer (1974) also found that subjects performed as well with a closed book test (notes not available) as they did with an open book test (notes available). Thus, the knowledge necessary to solve the problem is stored in memory, but subjects seem to have difficulty in knowing when to apply this knowledge. Subjects who were taught general concepts were better at applying their knowledge to conceptual problems and sub-problems of those solved in

training. Subjects who were taught the formula were better at applying their knowledge to problems similar to those solved in training.

In these studies, and in most of the relevant literature reviewed, tests were given the same day as, or soon after, the training period. Few long-term effects due to discovery and expository training have been studied. Two exceptions are studies by Solter and Mayer (1978), who studied children learning the concept of one-to-one correspondence, and Singer and Pease (1976), who studied the learning of a serial motor task. The same general pattern emerged from both studies. Subjects in discovery and expository groups were trained to the same criterion, and given a transfer test at 1 and 3 weeks (Solter & Mayer) or after 2 days (Singer & Pease). Discovery subjects performed better in the first transfer test than did guided subjects, but this difference disappeared on the next similar transfer test. The elimination of this difference between the groups after one test seems odd if, as Mayer and his colleagues claimed, and as the concept of permastore implies, different types of training cause different types of cognitive structures to form. However, it may be that subjects do initially have different structures, but the test in which subjects must discover the solution is enough to get the guided group to invoke their past knowledge, as the discovery group had done in training.

McDaniel and Schlager (1985) studied discovery and expository learning in the domains of water-jar and river-crossing problems. They found that strategies can be learned by discovery training. Discovery subjects solved transfer problems faster than did expository subjects, but only if the same kind of information (e.g., general strategies or procedures) had to be discovered in transfer as was discovered in training. Knowing when to activate past relevant knowledge seemed to aid in transfer. Discovery subjects seemed to have learned how to search efficiently and where to stop a search.

The apparent positive effect of discovery on transfer performance is important, but perhaps not as important as another effect of discovery. Many researchers have noted that subjects trained by discovery are highly motivated to learn, whereas subjects trained expositorily find the task uninteresting and monotonous. Kersh (1958) reported that some discovery subjects later tested their friends on the material they learned or went to the library to look up further information. Most expository subjects, on the other hand, complained that the task was boring, and they did not try to remember the rules because no one told them to do so. Not surprisingly, discovery subjects retained the rules they learned better than did the expository subjects.

In conclusion, it seems that discovery learning is beneficial in its motivating effects and in getting learners to integrate new information with their past knowledge, which aids in transfer. But subjects must first know the prerequisite concepts. However, if the learner needs to use only the specific knowledge learned in training (e.g., using a certain equation), then expository training is more efficient, and application of that knowledge will be more efficient. There is no strong evidence that either type of learning causes better retention of information; long-term studies are necessary to determine what remains of the information learned by each method.

Decay of Skills

A review of the relevant psychological literature reveals that there has been a great deal of work on the acquisition of skills, but relatively little on the maintenance of skills once they have been acquired. For example, much is known about the acquisition of simple short-term motor skills, of both a discrete (e.g., placing small cylindrical blocks into holes; Kimble & Bilodeau, 1949) and continuous (e.g., pursuit rotor; Adams, 1968) nature; about the

development of long-term mnemonic strategies, such as memorizing lengthy sequences of digits (see, e.g., Ericsson & Chase, 1982); and, from a more practical standpoint, about learning information in the classroom, such as how to solve mathematics problems (see e.g., Schoenfeld, 1979).

Although there is considerably less in the literature on skill decay, there are several recent studies which provide important insights into the question of how to counteract decay, both during and after skill acquisition. One study by Geoffrey Loftus (1985) addressed the crucial preliminary issue of how to measure decay and how to compare decay rates for differing degrees of original acquisition (although the retention intervals considered by Loftus were not particularly long). More specifically, Loftus reviewed work by Slamecka and McElree (1983) in which subjects learned verbal material to differing degrees of acquisition. Subjects were later tested after acquisition at retention intervals varying from 0 to 5 days. Slamecka and McElree found that the degree of acquisition did not interact with retention interval (so that the difference in performance levels on the items learned to differing degrees of acquisition was as great after long retention intervals as after short retention intervals), and they concluded that decay was independent of degree of acquisition. Loftus, on the other hand, examined the same data from a different perspective and reached the opposite conclusion. He assessed how much decay time is required for performance to fall from any given level to some lower level, and he diagnosed decay rates as different whenever such decay times differ. Since he found that it took less time for a drop from one level of performance to a lower level of performance when the initial acquisition degree was lower (the amount of overlearning was less), he concluded that decay is slower under high acquisition conditions than it is under low acquisition conditions. This hypothesis was originally proposed by Jost, who formulated the following law: "If two associations are now of equal strength but of different ages, the older one will lose strength more slowly with the further passage of time" (Woodworth & Schlosberg, 1954, p. 730). This work by Loftus not only provides important measurement tools for assessing decay rate, but it also reaffirms the important rule that skills can be acquired to differing degrees and those differences will influence the extent to which the skills are maintained.

Proposed Theoretical Mechanisms Underlying Skill Retention

Automatism

Our discussion of proposed theoretical mechanisms underlying skill maintenance will focus in part on a conjecture that follows from recent studies reported in the psychological literature, even though these studies have not directly addressed the issue of how to maintain skills. This conjecture is that maintenance of a skill will depend crucially on whether that skill is automatic; i.e., can be performed without conscious awareness. In particular, we propose that automatic skills will persist in time with little need for rehearsal or refreshing. This conjecture derives from recent research in cognitive psychology which has drawn an important distinction between automatic and controlled processes (e.g., Schneider & Shiffrin, 1977). Skills that require only minimal cognitive capacity and attention to perform are either fully or partly automatic; those that require resources and cognitive effort involve controlled processes. The classification of a skill as automatic or controlled depends in large part on the degree of prior skill acquisition. Many skills employ controlled processes during the initial stages of acquisition, but these processes become automatic with extensive practice. Consider, for example, the skills involved in riding a bicycle. Much attention and effort are expended by those initially learning this task; learners must concentrate on every movement they make in order to keep on track and avoid falling down. However, after

considerable practice, those riding bicycles can concentrate their thoughts and attention to other activities, such as the scenery or conversation (see Reed, 1982, pp. 50-51).

In order to test our conjecture concerning the persistence of automatic skills, we need a clear definition of and a set of criteria for automatism. Although the concept of automatism has been widely used by psychologists in recent years, there has been considerable debate concerning its defining attributes. Three prominent lines of research are particularly relevant to this issue. The first line derives from the seminal work of Schneider and Shiffrin (1977; Shiffrin & Schneider, 1977). This work has largely been focused on the role of attention in simple perceptual tasks like target detection. The second line derives from the influential work by LaBerge and Samuels (1974), which has been addressed in large part to the development of automatic subprocesses in the complex task of reading text. The third line of research was initiated by Hasher and Zachs (1979) and is centered on memory activities. Different, though overlapping methods and criteria for automatism have been developed by each of these sets of researchers.

Schneider and Shiffrin. After providing a review of the recent studies using their approach to automatism, Schneider, Dumais, and Shiffrin (1984) proposed a two-part definition for automatic processes: First, the process must not make use of or deplete the general, non-specific cognitive resources. Second, the process must be carried out in response to the presence of the relevant external stimuli even when subjects attempt to ignore the stimuli. In other words, capacity reductions do not affect automatic processing, and automatic processes are not subject to conscious control. How do automatic processes develop? Extensive practice is needed to ensure automatism, but not all types of practice are sufficient. Consistency of practice is the major factor. Schneider and Shiffrin compared two types of practice situations differing only in the amount of consistency in training; with "consistent mapping," the subject makes the same response each time a particular stimulus or a particular class of stimuli occurs, whereas with "varied mapping," the responses to stimuli change across training trials.

One paradigm used extensively by Schneider and Shiffrin (197) to compare consistent and varied mapping was the "multiple frame visual search" task. In this task, subjects are presented a series of frames successively, one after another with virtually no delay between frames. Each frame is presented very briefly for a duration called the "frame time." Before a trial, or sequence of frames, is started, the subjects are given a set of items, called the "memory set," and are told to make a "yes" response to any item from the memory set that occurs in the following sequence of frames. In the experiments reported by Schneider and Shiffrin, each trial consisted of the presentation of 20 successive frames to which the subject was to respond "yes" or "no" depending on whether or not the sequence included the presence of a memory set item; the dependent variable was detection accuracy (the percentage of hits and false alarm responses); and the independent variables included frame time, frame size (the number of characters occurring in each frame, typically ranging from one to four), the size or number of items included in the memory set, and (most crucially) the type of mapping, consistent or varied. For example, in one consistent mapping condition, the memory set items were always digits which occurred embedded within letters, whereas in the analogous varied mapping condition, the memory set consisted of some random subset of the letters. The differences between the two mapping conditions were striking: Subjects required a frame time of 120 msec or less to achieve a hit rate of approximately 95% in the consistent mapping condition, whereas even frame times as long as 800 msec did not always lead to such a high level of performance in the varied mapping

condition. Further, performance in the varied mapping condition was strongly influenced by frame size and memory set size, but the effects of those variables on consistent mapping were minimal. Hence, consistent mapping, but not varied mapping, satisfied a specific rendition of the first part of the definition of automatism: Changes in attention and memory load had essentially no effect on detection accuracy.

LaBerge and Samuels. LaBerge and Samuels (1974) proposed a theory of reading based on the automatization of its components. Reading is assumed to proceed in a series of hierarchical stages beginning with the identification of individual letters, combining letters into words, activating the meaning of the words, and combining these meanings into successively larger groups of comprehended discourse. Because attention is limited, reading skill is constrained by the degree to which the "lower level" components are automated. When children are first learning to read, their attention is directed toward identifying letters and combining them into words. Through practice, these components become automated, and attention can be focused on identifying the words as units. Through further practice, successively higher levels of processing become automated up to the point where attention can be focused solely on the gist of the text. The readers also have the ability to focus their attention on lower levels of processing when, for example, a high degree of accuracy is required or they are reading aloud. In support of their theory, LaBerge and Samuels (1974) showed that performance on identification and matching tasks benefited from practice for novel letters but not for familiar letters; presumably this processing stage had become automated for familiar letters, but initially required attention and then became automated for the novel letters.

This theory of automatism represents somewhat of a departure from that of Shiffrin and Schneider (1977). Although it assumes that some components of reading are truly automated, it also allows for the flexibility to focus attention on an automatic component if the task requires it.

One important consequence of automatism, as stressed by LaBerge and Samuels (1974) in their work on reading, is that the size of the processing units increases as tasks become automatic. Much research has been conducted in recent years that employs a simple detection task to investigate the size and nature of the processing units used when reading printed text. These investigations rest on the assumption that once individuals have abstracted a unit, they no longer concern themselves with its constituent parts. The detection task is thus used to indicate the size of an individual's processing units by revealing which constituents are ignored by the individual. For example, in one version of this task, subjects have been asked to read a passage of text and circle every instance of a given target letter (e.g., the letter t). It has been found that more errors are made on very common words, like the word the, than on rare words, like thy (Healy, 1976). These results have been explained by a unitization model (see Drewnowski & Healy, 1977; Healy, 1980; Healy & Drewnowski, 1983), according to which subjects miss letters on familiar words like the because they process such words automatically in units larger than the letter without completing processing at the letter level.

Hasher and Zacks. Hasher and Zacks (1979) proposed that memory processes vary along a continuum of the amount of limited-capacity attentional resources these processes consume. At one end of the continuum are purely automatic processes. These processes require virtually no attentional capacity and can therefore be performed in parallel with other processes. These skills are unaffected by contextual variables such as arousal levels, and are minimally changed by development or practice. Examples of such automated processes are

the encoding of spatial, temporal, and frequency information. Processes at this end of the continuum are presumed to be innate whereas some less polarized automatic skills can be acquired through practice, such as the encoding of the meaning of a word. At the other end of the attention consumption continuum are "effortful" processes. These tasks require a great deal of attentional resources and are likely to interfere with other resource-consuming activities operating in parallel. Effortful processes are assumed to be influenced by contextual factors and practice. Examples of effortful processes are rehearsal and mnemonic elaboration.

Hasher and Zacks (1979) used this framework to account for the effects of instruction, practice, interference, arousal, and development on memory tasks which are predicted to be either automatic or effortful. The predictions are, in general, consistent with the findings cited by Hasher and Zacks. One notable inconsistency is an aging difference in memory for frequency, a presumably automatic process.

This framework makes some interesting predictions about the maintenance of skills. For example, automated components of a skill which were not developed through practice should not be expected to deteriorate when the skill is not used. It should therefore be unnecessary to include them in a maintenance program. Additionally, less-polarized automated components developed through practice would be expected to show a strong benefit from a maintenance program. This is predicted because practice should help to ensure that the process does not return to an effortful state and reduce the cognitive capacity necessary for other ongoing operations.

Assessing automatism. After reviewing much of the research on automatism, Jonides, Naveh-Benjamin, and Palmer (1985) proposed two principles that should be followed in the study of automatism on cognitive processes. The first of these principles was that the concept of automatism is best applied to component processes of a complex task, not to the task as a whole. The concerns here are twofold. One, in judging a task to be automatic, the experimenter might overlook some component processes that are not themselves automatic. Conversely, the second concern is that by judging a task not to be automatic, the experimenter may overlook component processes that are, by themselves, automatic. To avoid these pitfalls, Jonides et al. suggested constructing sensitive tests utilizing an explicit model of the processes involved in the task.

The second principle Jonides et al. proposed was that the criteria chosen for the evaluation of automatism should be motivated by the processes involved in the task, not by individual criteria (such as minimal attentional demands and lack of voluntary control) attributed to generic automatic tasks. The example used to fortify this principle comes from the research comparing maintenance and elaborative rehearsal (Naveh-Benjamin & Jonides, 1984). Three criteria were used to assess the automatism of both forms of rehearsal: demand on capacity, susceptibility to interruption, and stereotype. The importance of listing these criteria is that though stereotype is atypical of the criteria that usually appear in the automatism literature, it proved to be particularly effective in determining the relative automatism of elaborative and maintenance rehearsal methods.

The conclusions drawn are that before an automatism study is conducted, an explicit model of the component process of a task should be formulated and that once formulated, the criteria for automatism should be relevant to those specific component processes, not tasks in general.

Central to the study of automatism through attention is the belief that the human is a limited capacity processor. That is, there is a limited amount of resources from which control processes can draw in attempting to perform a task. Automatic-control processing theory assumes that capacity limitations arise out of competition between concurrent control processes. Hence, when concurrent performance of tasks causes the control process demands to exceed the capacity limitations, there is a decrement in performance in one or both of the tasks due to lack of resources. In contrast to control processes, automatic processes can occur in parallel, free from the limitations of control processes, without any noticeable decrement in performance (Wickens, 1980). A middle ground approach to this control-automatic dichotomy proposes that as a task becomes better and better learned, its resources (attentional) demands gradually decline until ultimately demands are minimal and the task is assumed to be automatic (Logan, 1979).

With capacity theory (e.g., Kahneman, 1973) in hand, the concern switches to the measurement of attention. The paradigm most used to measure the resource utilization of a task is the dual-task paradigm (e.g., Posner & Boies, 1971). The dual-task paradigm permits the comparison of two tasks in terms of common units. This comparison is achieved through the use of a subsidiary task setup. In effect, the subjects are given two tasks, a primary task which they are instructed to do as well as they can and a secondary task which they are to attempt only after successfully completing the primary task. In this method, the measure of attention required for the primary task is inversely related to the performance of the secondary task. As performance increases in the secondary task, it is presumed that the attentional demand for the first task is decreasing. Conversely, should performance on the secondary task be low, attentional demands for the first task are presumed to be high. Examples of a dual-task setup would be to have the subject read a passage while simultaneously detecting letters (Proctor & Healy, 1985) or read a passage while simultaneously transcribing dictation (Spelke, Hirst, & Neisser, 1976).

Finally, caution should be taken in comparing primary tasks in a dual-task paradigm, as the disruption of the secondary task depends not only on the difficulty of the primary task but also on its structure. Secondary tasks may interfere differently on differing primary tasks, as would be the case in the comparison of a task with a high motor load to a task with a high conceptual or perceptual load. Perhaps the safe road to take would be a modification of the dual-task paradigm, where a battery of secondary tasks is used instead of just one such task.

Alternatives to Automatism

Spelke, Hirst, and Neisser. Although the notion of automatism has been widely accepted, some data have suggested the need for alternative approaches. Spelke, Hirst, and Neisser (1976), for example, trained two people to copy dictated words while reading stories. After considerable practice, the participants were able to do both tasks as well simultaneously as they could do them alone. Subsequent experiments provided evidence against the hypothesis that participants were alternating their attention between the two tasks. Within the same paradigm, Hirst, Spelke, Reaves, Caharack, and Neisser (1980) showed that redundancy of the reading material did not influence the time necessary to reach criterion performance on the dual task and that transfer from more to less redundant material was equivalent to transfer from less to more redundant material. If participants were alternating their attention between the reading and the writing tasks, then the more redundant reading material should have facilitated this process and thus increased dual-task performance. Additionally, if participants learned to exploit the redundancy of the text,

then their performance should have dropped when they transferred to the less redundant material. Hirst et al. also provided findings which they interpreted as evidence against the hypothesis that the word-copying task becomes automated and thus does not involve attentional resources. They found that fewer copying errors were made when the to-be-copied materials were sentences relative to random word strings and that recognition memory for integrated sentences indicated that the participants had remembered implications of the sentences. These results strongly suggest that the participants understood the sentences they copied and therefore the copying task required attention and was not automatic. These researchers concluded that the notion of a fixed attentional capacity is incorrect; with sufficient practice, attention can be divided without automatism or the loss of conscious control.

Salthouse. A very similar argument has been made based on the performance of skilled typists. Salthouse (1984) and others have argued that skilled typing is the result of learning to overlap the performing of its component processes (see the cascade model of McClelland, 1979, for a similar idea). Evidence for overlapping comes from the finding that typing speed is greatly decreased when the size of a preview text is reduced below about eight characters and reaches the speed of a choice reaction time task with a preview window of only one character. It is easy to see how reducing the size of the preview window would impede the overlapping of component processes, but this manipulation should have been inconsequential if the typists were performing each of the component processes sequentially. The finding from eye movement studies of the eye-hand span (i.e., typists fixate several characters ahead of the one being typed; Butsch, 1932) provides convergent evidence for the overlapping processes notion. Salthouse showed that typing skill is positively correlated with eye-hand span, suggesting again that typing skill is related to the degree to which component processes can occur in parallel.

Although different typing processes can presumably occur simultaneously, there is good evidence that they are not automatic. For example, typing errors are detected almost immediately, as shown by an increased latency to press a key following an error (Salthouse, 1984; Shaffer 1976) and by the fact that a key is pressed with less force when it is an error (Rabbitt, 1978 Wells, 1916). Additional evidence against the claim that typing is automatic is the finding that when participants were told to stop typing whenever they heard a tone, they typically stopped within one or two letters rather than typing to the end of the word (Logan, 1982).

The results from the study of skilled typing are consistent with the dual-task studies described above. Through practice, multiple tasks or multiple components of a task can be performed in parallel without the loss of conscious control. These results are at some level inconsistent with (or at least suggest limitations of) the theories of attention proposed by Shiffrin and Schneider (1977) and Hasher and Zacks (1979). Explanations for the typing results have been proposed which claim that a process can be automatic but still available to conscious control. However, such explanations can be held only at the expense of making the definition of automatism more vague and less susceptible to experimental tests.

Salthouse (1984) also provided data which make a simple component explanation of typing difficult. For example, choice reaction time, presumably a component of typing, is uncorrelated with typing skill but is known to be slower for older people. However, old and young typists are equivalent in typing speed. It appears, then, either that choice reaction time is not a component of typing or that older typists are able to compensate in some way for a deficit in the reaction time component. Regardless of which explanation of

this paradox is true, this finding suggests that examining component processes out of context may yield misleading information concerning performance on the global task.

Cheng. As an alternative to automatism, Cheng (1985) provided an interpretation of skill improvement based on task restructuring. She argued that through practice, the components of a task can be reorganized such that they operate in a more coherent, integrated manner. Task performance can be improved by replacing less efficient modes of processing with more efficient ones. In other words, one can learn to perform a task in a new way, rather than learning to do the old way faster or more automatically. As an example, Cheng contrasted the difference between solving the arithmetic problem of finding the sum of ten 2s by addition and by multiplication. Although the solution can be obtained by either method, clearly multiplication is more efficient. Indeed, students learn how to restructure the addition problem into a multiplication problem, rather than learning to do numerous addition operations automatically.

This account of skill improvement is appealing and is consistent with the theorizing on skilled typing (Salthouse, 1984) and dual-task performance (Hirst et al., 1980). The relationship between automatism and restructuring may be best thought of as supplementary rather than contradictory, since both processes are probably operating to some extent in the acquisition of all skilled behavior.

Anderson. Still another important alternative approach to the problem of skill acquisition is Anderson's (1982) approach based on knowledge proceduralization. Anderson claimed that during the initial stage of skill learning, knowledge is in a declarative form. In order to use knowledge in this form, it must be retrieved from long-term memory and held active in working memory. Consequently, the use of declarative knowledge is associated with slow retrieval times and the inhibition of concurrent operations which require working memory capacity.

Through practice, declarative knowledge is compiled or proceduralized. Proceduralized knowledge does not have to be entered into working memory to be acted upon. Consequently, it can be retrieved quickly and its use requires virtually no working memory capacity. Whereas declarative knowledge is flexible and can be modified to remove incorrect information, proceduralized knowledge is much more difficult to change. After a skill has been proceduralized, further learning and improvement are accomplished by fine-tuning the application of the operations.

Anderson's theory provides insight into the utility of the skill acquisition process. When a skill is first being acquired, much of what is gained involves learning to perform the skill correctly before learning to perform it quickly or more efficiently. At this stage, it is advantageous that the task relies on working memory, thus enabling the close monitoring and adjustment of skill knowledge. Once the task can be performed correctly, the knowledge does not need to be changed and it can be proceduralized to crystallize it and increase the efficiency of its operation.

This model is similar to Cheng's (1985) restructuring hypothesis in that skill improvement is seen as the result of developing new, more efficient operations. It is also similar to automatization theories (e.g., Shiffrin & Schneider, 1977; Hasher & Zacks, 1979) in that both approaches stress that an important difference between a proceduralized (automatic) skill and a nonproceduralized (controlled) skill is the gain in speed and efficiency at the cost of a loss in control and flexibility. Both theories are also similar in

assuming that working memory is strongly involved in learning a new skill and its role is greatly diminished as the skill becomes well learned.

Additional Theoretical Mechanisms

Levels of processing and transfer-appropriate processing. Levels of processing theories attempt to explain how different types of processing performed on stimuli affect the retention of those stimuli. The original levels of processing framework (Craik & Lockhart, 1972) attempted to explain differences in retention of words based on the kind of processing performed on those words. Words which were processed at a "deep" (more semantic) level were better remembered than words which were processed at a "shallow" (structural or phonemic) level. Craik and Lockhart hypothesized that depth of processing could be indexed by the amount of time the subject spent processing the word, with deeper levels of processing taking more time. Craik and Tulving (1975), however, found that in a complex structural task, subjects took longer to process a word than in a semantic task, but words processed semantically were still better remembered. They attempted to find an alternative measure of depth of processing. It was discovered that subjects remembered the words to which they had responded "yes" in the learning task better than words to which they had responded "no." The hypothesis was that words which fit the context in which they were learned (words responded to with "yes") were more tightly integrated with that context than were words which did not fit the context. When the word fit the context, the context could serve as a retrieval cue for that word. Craik and Tulving proposed that a process of elaboration, in which the word is tightly integrated with the learning context or the event is specified more uniquely, could explain the retention results. They did not, however, specify how to measure the level of elaboration; so, this theory did not have any more predictive power than the original Craik and Lockhart framework.

There have been many other criticisms, besides lack of predictability, of the depth of processing framework (see Baddeley, 1978, for a review of some of these criticisms). Morris, Bransford, and Franks (1977) showed that if an appropriate test were used, a supposedly shallow level of processing would produce better retention than would a semantic level of processing (on this test). In particular, they found that subjects trained with a rhyming task showed better retention on a rhyming test than those trained with a semantic task (filling in a word in a sentence frame). They hypothesized that the previous benefits of semantic processing over "shallower" processing were found because subjects were given a semantically oriented test. Stein (1978) found results similar to those of Morris et al. by examining the retention of semantic versus letter case information. Subjects given case questions during acquisition performed better on a subsequent case recognition test than did subjects given semantic questions during acquisition. Also, subjects given semantic acquisition questions performed better on a semantic recognition test than did subjects given case acquisition questions. Stein concluded that the ability to remember particular aspects of an input depends crucially upon how well the learner was able to detect that information during encoding and was able to encode that information distinctly from other aspects of the input.

It therefore seems that an important conclusion from this work is that the form in which knowledge will be used should be foremost in determining how the knowledge should be taught, in order for optimal retention to be demonstrated. For example, if the knowledge will be used phonetically, it should be taught phonetically. Within a particular "level" of training, the more integrated the information is with the learning context, the more distinct the information can

be made from other information (the more elaborated), and the better it will be remembered.

The generation effect. The generation effect refers to the observation that information which is generated is better remembered than information which is simply read. The typical experimental paradigm used to examine this effect provides the subject with a stimulus item and either a rule for generating a response item or the response item itself. The memorial advantage of item generation has been obtained with a variety of generating tasks, such as providing an opposite (McElroy & Slamecka, 1982), transposing letters (Gardiner & Hampton, 1985), completing word fragments (Glisky & Rabinowitz, 1985), and even providing the product of a multiplication problem (Gardiner & Rowley, 1984). Generation effects have been found with a number of memory tests, such as free recall (Gardiner & Hampton, 1985), cued recall (Graf, 1980), and recognition (McElroy & Slamecka, 1982). Further, the effect holds for several types of verbal material, with the exception of nonword letter strings.

The generation effect is clearly a robust phenomenon, but a satisfactory theoretical account of it has not yet been provided. McElroy and Slamecka (1982) broadly divided the proposed explanations into two categories, those which point to the involvement of semantic memory which results from the act of generation and those which stress the inherent difference between generating and reading. Basically, the former type of explanation predicts that a generation task must involve semantic memory to provide a generation effect, and the latter type of explanation predicts that any generation task will result in a generation effect.

The "generate only" theory is contradicted by the lack of a generation effect for nonwords. Because nonwords are not a part of semantic memory, this finding has been taken as evidence for the "semantic memory" explanation. However, a strong version of the semantic memory account, which states that the generated item must be represented as a unit in the lexicon (mental dictionary), is also problematic because the effect has been obtained with digits (Gardiner & Hampton, 1985; Gardiner & Rowley, 1984) and word pairs (Gardiner & Hampton, 1985). In a recent series of experiments, Glisky and Rabinowitz (1985) showed that performing a generation task at the time of test enhances the generation effect if the item was generated during study but not if it was only read at study. Additionally, they found that the enhancement could be obtained if the same specific operation was repeated but not if the same generation rule was applied in a different manner.

Regardless of which mechanism produces the generation effect, its potential application for the retention of skills is evident. Unfortunately, no research has been conducted to determine whether or not a generation effect can be obtained at a long retention interval. If the generation effect can be extended to long-term retention, then it would be beneficial to incorporate generation tasks in a skill training program.

The role of consciousness in skill retention. It is commonly assumed that information retrieval and/or skill elicitation somehow involves consciousness or awareness. This assumption derives largely from traditional studies of memory involving recall or recognition primarily of verbal information encoded at some earlier time. The assumption is not consistent with some recent results from studies that have assessed performance with tests that are not closely tied to any particular prior experiences (Tulving, Schacter, & Stark, 1982). In these cases, subjects exhibit memory in the form of facilitated performance on a given task but without any accompanying conscious recollection of the experiences that contributed to that facilitation. For example, Cofer (1967), among others, has reported that subjects are more successful on a word completion test (i.e., a

test on which the subject is presented with two or three letters of a word and has to fill in the missing letters) when the target word has recently been presented than when presented a new word. This facilitation of completion performance has been called a direct priming effect.

Direct priming effects have also been demonstrated in word identification tasks and lexical decision tasks (Graf & Schacter, 1985; Tulving, 1983). A variety of constructs have been used to distinguish between the type of memory that is tapped by priming tests, on the one hand, and by traditional recall and recognition tests on the other. One distinction that seems particularly appropriate is the distinction between episodic, semantic, and procedural memory (Tulving, 1983). Almost all examples of episodic memory, requiring as they do a reference to certain prior experiences, involve conscious, effortful recollection. Procedural memory and semantic memory, in contrast, are often revealed without any apparent conscious effort to retrieve. According to Mandler (1980), effortless, automatic memory is based on the activation of specific pre-existing memory representations. Study materials used in many memory experiments are familiar, individual words, already represented in long-term, semantic memory prior to their appearance in the study list. Mandler argued that these pre-existing representations are activated as a result of presenting the study items, and that activation occurs automatically and thus independently of the processes that mediate conscious remembering. Such an interpretation is supported by studies of amnesic patients. Amnesic patients are characterized by poor memory for recent events but relatively normal retention of older knowledge and skill (Squire, Cohen, & Nadel, 1984). One of the hallmarks of amnesia is a patient's inability to acquire and remember new associations. Nevertheless, amnesic patients show relatively normal priming or effortless retention when the study materials are familiar items that have a pre-existing memory representation (Graf, Squire, & Mandler, 1984). The finding of normal priming effects, in conjunction with the observation that amnesic patients seem unable to acquire new associations, is consistent with the view that effortless memory is mediated by the activation of pre-existing representations. Learning new episodes means there is no pre-existing memory representation. Thus, activation cannot produce correct remembering. Further, whatever the mechanism is for establishing new representations, it seems to be defective in amnesic patients. But priming derives from already-existing representations which can be activated automatically to produce a memory-like effect.

We take these studies to have a further implication. When a test is arranged for a skill learned earlier, it is typically found that skilled performance falls below the level achieved during original practice. This suggests that either (a) the skill as a whole or all of its components deteriorated, in some sense, over time without exercise; or (b) skill components differ in their persistence, some of them being available for retrieval in full-blown form at a later time while others have significantly decayed. On the basis of results that distinguish between automatic and effortful retrieval of items or components from memory, we suggest the hypothesis that some components do indeed persist unaltered over time and can be activated by the reintroduction of the appropriate stimulus to their mental representation. Other components, in contrast, which are generally connected to the changed circumstances between original training and subsequent test, or are heavily context dependent, might not be so readily available. Wherever a component lacks mental representation, for whatever reason, automatic activation will not work. The subject must engage in a conscious search of memory and possibly the establishment or reestablishment of mental representation in support of that component. Available evidence suggests that memory for motor skills, especially highly integrated, continuous motor skills, is excellent because retrieval of the

appropriate representation tends to be effortless. Skill components tied to particular episodes or experiences require effortful search and possible re-encoding, with the implication that their performance may be relatively deteriorated after prolonged periods of time.

Conclusions

In conclusion, the literature shows that the long-term retention of knowledge and skills is a complex multifaceted problem. The research conducted to date on long-term retention has already uncovered some of the psychological principles involved in its promotion. Further, investigations of the mechanisms believed to mediate skill acquisition have provided principled accounts of how and when permanent memory should occur. Additional investigations of such mechanisms for very long-term memory will determine whether or not these mechanisms are applicable to skill retention. It is our belief that a principled investigation of skill maintenance will provide useful applications for improving retention and provide a major contribution to both the military and civilian communities.

REFERENCES

Adams, J.A. (1968). Response feedback and learning. Psychological Bulletin, 70, 486-504.

Adams, J.A. (1971). A closed loop theory of motor learning. Journal of Motor Behavior, 3, 111-150.

Anderson, J.R. (1982). Acquisition of cognitive skill. Psychological Review, 89, 369-406.

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

Bahrick, H.P. (1979). Maintenance of knowledge: Questions about memory we forgot to ask. Journal of Experimental Psychology: General, 108, 296-308.

Bahrick, H.P. (1984). Semantic memory content in permastore: Fifty years of memory for Spanish learned in school. Journal of Experimental Psychology: General, 113, 1-29.

Battig, W.F. (1966). Facilitation and interference. In E.A. Bilodeau (Ed.), Acquisition of skill. New York: Academic Press.

Battig, W.F. (1972). Intratask interference as a source of facilitation in transfer and retention. In R.F. Thompson & J.F. Voss (Eds.), Topics in learning and performance. New York: Academic Press.

Battig, W.F. (1979). The flexibility of human memory. In L.S. Cermak & F.I.M. Craik (Eds.), Levels of processing and human memory (pp. 23-44). Hillsdale, NJ: Erlbaum.

Bruner, J.S. (1961). The act of discovery. Harvard Educational Review, 31, 21-32.

Butsch, R.L.C. (1932). Eye movements and tye eye-hand span in typewriting. Journal of Educational Psychology. 23, 104-121.

Cheng, P.W. (1985). Restructuring verses automaticity: Alternative accounts of skill acquisition. Psychological Review, 92, 414-423.

Chiesi, H.L. (1976). Retroactive inhibition as a function of the locus of categorized response sets. Memory & Cognition, 4, 514-518.

Cofer, C.N. (1967). Conditions for the use of verbal associations. Psychological Bulletin, 68, 1-12.

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.

Drewnowski, A., & Healy, A.F. (1977). Detection errors on the and and: Evidence for reading units larger than the word. Memory and Cognition, 5, 636-647.

Egan, D.E., & Greeno, J.G. (1973). Acquiring cognitive structure by discovery and rule learning. Journal of Educational Psychology, 64, 85-97.

Einstein, G.O. (1976). Effects of simultaneous interference upon free recall learning and retention. Memory & Cognition, 4, 701-708.

Ericsson, K.A., & Chase, W.G. (1982). Exceptional memory. American Scientist, 70, 607-615.

Friedlander, B.Z. (1965). A psychologist's second thoughts on concepts, curiosity, and discovery in teaching and learning. Harvard Educational Review, 35, 18-38.

Gardiner, J.M., & Hampton, J.A. (1985). Semantic memory and the generation effect: Some tests of the lexical activation hypothesis. Journal of Experimental Psychology: Learning, Memory, and Cognition, 11, 732-741.

Gardiner, J.M., & Rowley, J.M.C. (1984). A generation effect with numbers rather than words. Memory & Cognition, 12, 443-445.

Glisky, E.L., & Rabinowitz, J.C. (1985). Enhancing the generation effect through the repetition of operations. Journal of Experimental Psychology: Learning, Memory, and Cognition, 11, 193-205.

Graf, P. (1980). Two consequences of generating: Increased inter- and intra-word organization of sentences. Journal of Verbal Learning and Verbal Behavior, 19, 316-327.

Graf, P., & Schacter, D.L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. Journal of Experimental Psychology: Learning, Memory and Cognition, 11, 501-518.

Graf, P., Squire, L.R., & Mandler, G. (1984). The information that amnesic patients do not forget. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 164-178.

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

Healy, A.F. (1976). Detection errors on the word the: Evidence for reading units larger than letters. Journal of Experimental Psychology: Human Perception and Performance, 2, 235-242.

Healy, A.F. (1980). Proofreading errors on the word the: New evidence on reading units. Journal of Experimental Psychology: Human Perception and Performance, 6, 45-57.

Healy, A.F., & Drewnowski, A. (1983). Investigating the boundaries of reading units: Letter detection in misspelled words. Journal of Experimental Psychology: Human Perception and Performance, 9, 413-426.

Healy, A.F., Mieskey, L., Fendrich, D.W., Crutcher, R.J., Little, W., & Bourne, L.E., Jr. (1988, August). Skill maintenance: Specific sample methodologies (AFHRL-TP-87-72). Brooks AFB, TX: Training Systems Division, Air Force Human Resources Laboratory.

Hiew, C.C. (1977). Sequence effects in rule learning and conceptual generalization. American Journal of Psychology, 90, 207-218.

Hirst, W., Spelke, E.S., Reaves, C., Caharack, G., & Neisser, U. (1980). Dividing attention without alternation or automaticity. Journal of Experimental Psychology: General, 109, 98-117.

Johnson, R.B. (1964). Recognition of nonsense shapes as a function of degree of congruence among components of a pretraining task. Unpublished doctoral dissertation, University of Virginia, Charlottesville, VA.

Jonides, J., Naveh-Benjamin, M., & Palmer, J. (1985). Assessing automaticity. Acta Psychologica, 60, 157-171.

Kahneman, D. (1973). Attention and effort. Englewood Cliffs, NJ: Prentice-Hall.

Kersh, B.Y. (1958). The adequacy of "meaning" as an explanation for the superiority of learning by independent discovery. Journal of Educational Psychology, 49, 282-292.

Kimble, G.A., & Bilodeau, E.A. (1949). Work and rest as variables in cyclical motor learning. Journal of Experimental Psychology, 39, 150-157.

LaBerge, D., & Samuels, S.J. (1974). Toward a theory of automatic information processing in reading. Cognitive Psychology, 6, 293-323.

Lee, T.D., & Magill, R.A. (1983). The locus of contextual interference in motor-skill acquisition. Journal of Experimental Psychology: Learning, Memory and Cognition, 9, 730-746.

Loftus, G.R. (1985). Evaluating forgetting curves. Journal of Experimental Psychology: Learning, Memory, and Cognition, 11, 397-406.

Logan, G.D. (1979). On the use of concurrent memory load to measure attention and automaticity. Journal of Experimental Psychology: Human Perception and Performance, 5, 189-207.

Logan, G.D. (1982). On the ability to inhibit complex movements: A stop-signal study of typewriting. Journal of Experimental Psychology: Human Perception and Performance, 8, 778-792.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. Psychological Review, 87, 252-271.

Mayer, R.E. (1974). Acquisition processes and resilience under varying testing conditions for structurally different problem-solving procedures. Journal of Educational Psychology, 66, 644-656.

Mayer, R.E. (1975). Information processing variables in learning to solve problems. Review of Educational Research, 45, 525-541.

Mayer, R.E., & Greeno, J.G. (1972). Structural differences between learning outcomes produced by different instructional methods. Journal of Educational Psychology, 63, 165-173.

Mayer, R.E., Stiehl, C.C., & Greeno, J.G. (1975). Acquisition of understanding and skill in relation to subjects preparation and meaningfulness of instruction. Journal of Educational Psychology, 67, 331-350.

McClelland, J.L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. Psychological Review, 86, 287-330.

McDaniel, M.A., & Schlager, M. (1985). Discovery learning and transfer of problem solving skills: A cognitive approach. Manuscript submitted for publication.

McElroy, L.A., & Slamecka, N.J. (1982). Memorial consequences of generating nonwords: Implications for semantic-memory interpretations of the generation effect. Journal of Verbal Learning and Verbal Behavior, 21, 249-259.

Morris, C.D., Bransford, J.D., & Franks, J.J. (1977). Levels of processing versus transfer appropriate processing. Journal of Verbal Learning and Verbal Behavior, 16, 519-533.

Naveh-Benjamin, M., & Jonides, J. (1984). Maintenance rehearsal: A two-component analysis. Journal of Experimental Psychology: Learning, Memory and Cognition, 10, 369-385.

Naylor, J.C., & Briggs, G.E. (1961). Long-term retention of learned skills: A review of the literature (ASD-TR-61-390,AD-267 943) Wright-Patterson AFB, OH: Advanced Systems Division.

Nitsch, K.E. (1977). Structuring decontextualized forms of knowledge. Unpublished doctoral dissertation, Vanderbilt University, Nashville, TN.

Pagel, J.C. (1973). A Markov analysis of transfer in paired-associate learning with high intralist similarity. Journal of Verbal Learning and Verbal Behavior, 12, 456-470.

Pellegrino, J.W. (1972). Effects of intralist response formal similarity upon paired-associate transfer and retroactive inhibition. Journal of Experimental Psychology, 92, 134-141.

Posnansky, C.J. (1974). An investigation of serial learning and retention processes in children. Journal of Experimental Child Psychology, 18, 127-148.

Posner, M.I., & Boies, S.J. (1971). Components of attention. Psychological Review, 78, 391-408.

Proctor, J.D., & Healy, A.F. (1985). A secondary-task analysis of a word familiarity effect. Journal of Experimental Psychology: Human Perception and Performance, 11, 286-303.

Rabbitt P. (1978). Detection of errors by skilled typists. Ergonomics, 21, 945-958.

Reed, S.K. (1982). Cognition: Theory and applications. Monterey, CA: Brooks/Cole.

Salthouse, T.A. (1984). Effects of age and skill in typing. Journal of Experimental Psychology: General, 113, 345-371.

Schab, J.M. (1975). Intratask interference and long term memory. Unpublished master's thesis, University of Colorado, Boulder, CO.

Schendel, J.D., Shields, J.L., & Katz, M.S. (1978). Retention of motor skills: Review (U.S. Army Technical Paper, No. 313). Alexandria, VA.

Schmidt, R.A. (1975). A schema theory of discrete motor skill learning. Psychological Review, 82, 225-260.

Schneider, W., Dumais, S.T., & Shiffrin, R.M. (1984). Automatic and control processing and attention. In R. Parasuraman & D.R. Davies (Eds.), Varieties of attention (pp. 1-27). Orlando, FL: Academic Press.

Schneider, W., & Shiffrin, R.M. (1977). Controlled and automatic human information processing: I. Detection, search and attention. Psychological Review, 84, 1-66.

Schoenfeld, A.H. (1979). Explicit heuristic training as a variable in problem-solving performances. Journal for Research in Mathematics Education, 10, 173-187.

Shaffer, L.H. (1976). Intention and performance. Psychological Review, 83, 375-393.

Shea, J.B., Hunt, J.P., & Zimny, S.T. (1985). Cognitive aspects of movement production. Unpublished manuscript, mimeo, Pennsylvania State University, University Park, PA.

Shea, J.B., & Morgan, R.L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. Journal of Experimental Psychology: Human Learning and Memory, 5, 179-187.

Shea, J.B., & Zimny, S.T. (1982). Context effects in memory and learning movement information. In R.A. Magill (Ed.), Memory and motor control. Amsterdam: North Holland Press.

Shiffrin R.M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. Psychological Review, 84, 127-190.

Singer, R.N., & Pease, D. (1976). A comparison of discovery learning and guided instructional strategies on motor skill learning, retention, and transfer. The Research Quarterly, 47, 788-796.

Slamecka, N.J., & McElree, B. (1983). Normal forgetting of verbal lists as a function of their degree of learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 9, 384-397.

Solter, A., & Mayer, R.E. (1978). Broader transfer produced by guided discovery of number concepts with preschool children. Journal of Educational Psychology, 70, 363-371.

Spelke, E.S., Hirst, W.C., & Neisser, U. (1976). Skills of divided attention. Cognition, 4, 215-230.

Squire, L.R., Cohen, N.J., & Nadel, L. (1984). The medical temporal region and memory consolidation: A new hypothesis. In H. Weingarner & E. Parker (Eds.), Memory consolidation. Hillsdale, NJ: Erlbaum.

Stein, B.S. (1978). Depth of processing re-examined: The effects of the precision of encoding and test appropriateness. Journal of Verbal Learning and Verbal Behavior, 17, 165-174.

Tulving, E. (1983). Elements of episodic memory. Oxford, England: Clarenden Press.

Tulving, E., Schacter, D.L., & Stark, H.A. (1982). Priming effects in word fragment completion are independent of recognition memory. Journal of Experimental Psychology: Learning, Memory and Cognition, 8, 336-342.

Tulving, E., & Thomson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. Psychological Review, 80, 352-373.

Underwood, B.J., & Lund, A.M. (1979). Retention differences as a function of the number of verbal lists learned simultaneously. Journal of Experimental Psychology: Human Learning and Memory, 5, 151-159.

Wells, F.L. (1916). On the psychomotor mechanisms of typewriting. American Journal of Psychology, 27, 47-70.

Wickens, C.D. (1980). The structure of attentional resources. In R. Nickerson & R. Pew (Eds.), Attention and performance VIII. Hillsdale, NJ: Erlbaum.

Woodworth, R.S., & Schlosberg, H. (1954). Experimental psychology (rev. ed.). New York: Holt, Rinehart and Winston.