

The Acquisition and Retention of Knowledge: Exploring Mutual Benefits to Memory Research and the Educational Setting

MOSHE NAVEH-BENJAMIN*

Ben-Gurion University of the Negev, Israel

SUMMARY

This article examines the relationship between laboratory-based memory research and educational research and practice. The first section addresses possible contributions of laboratory research to education. It shows that such research provides meta-theoretical, theoretical, and conceptual frameworks to explain educationally relevant phenomena, as well as solid empirical contributions that can be applied in relevant educational settings. The second section presents several examples of empirical, educationally relevant memory research. Analysis of these studies shows the potential mutual benefits to both educational practice and memory research. In addition to demonstrating the relevance of memory research to education and educational practice, such research can also increase the external validity of memory research, put constraints on the scope of laboratory findings, and help refine theoretical positions. The third section of the article discusses principles to be used in evaluating the applicability and relevance of laboratory research and memory phenomena to education, and looks at ways to increase the applicability of research to educational practice. Among the outlined suggestions are choosing parameters in laboratory-based memory research similar to those existing in educational settings, and considering relevant interactions, laboratory effects size, and variation of relevant variables. Using broad categories of variables, including those of individual differences and non-cognitive ones, will benefit ecologically oriented memory research.

The scientific study of memory has been characterized, at least initially, by the rigorous application of the scientific method in the laboratory. There have been historical reasons for this trend, not the least of which were concerns about establishing a new field modelled after the natural sciences (Ebbinghaus, 1885/1963). As a result, the new field adopted some of the methodological tools from the natural sciences. These tools include careful observations and experiments in a controlled environment—namely, the psychological laboratory. Such an environment has helped to establish a large body of knowledge, including concepts, empirical findings and generalizations, and theories.

Memory researchers have recently begun to question the wisdom behind some of the earlier decisions. Particularly, questions have been raised about the type of research dictated by the laboratory. Neisser (1978) commented that the first hundred years of scientific memory research in the laboratory told us little about everyday

*Correspondence to: Moshe Naveh-Benjamin, Department of Behavioral Sciences, Ben-Gurion University of the Negev, P O Box 653, Beer-Sheva 84120, Israel.

uses of memory. Many of the empirical generalizations (effects of interference, of study time and of meaningfulness) are quite trivial. Some of the real discoveries are about memories that last up to a few seconds (sensory and short-term memories). It is also not clear whether the theories proposed, based on findings and generalizations established in the laboratory, tell us anything meaningful about memory in real life.

Of course, memory researchers clearly think their work is relevant to everyday behaviour. This article evaluates the relationship between laboratory memory research and everyday memory in one domain—education, particularly higher education. The article is organized according to the following sections. The first section evaluates the possible contributions of laboratory memory research to educational practices, comparing these possible contributions with contributions actually made. The second section analyses some specific research examples to illustrate the potential contributions of applied, educationally oriented memory research both to practice and theory. The third section suggests principles to be used in evaluating the potential applicability of basic laboratory phenomena to education. In this section some modifications in laboratory memory research are discussed, modifications which would make laboratory memory results more relevant for educational settings.

EVALUATING POSSIBLE CONTRIBUTIONS OF LABORATORY-BASED MEMORY RESEARCH TO EDUCATIONAL PRACTICE

There are probably few contexts in which the application of memory research should be more obvious than in the context of higher education. The macro-level tasks involved in higher education, including learning new information and retrieving it on occasion (e.g. in tests), are particularly related to questions of acquiring, storing, and retrieving knowledge—questions at the centre of laboratory memory research in the past century.

The same is true of micro-level analysis. For example, it might be difficult to get anything out of a lecture without making extensive use of a vast base of knowledge readily available in memory. To comprehend the content of a lecture, we need to call up from memory the information needed to interpret the message of the speaker. This goes from understanding each individual word (using our working vocabulary store), through combining words into sentences, up to forming a theme. All these interpretive activities require us to draw on past knowledge about a topic. Note-taking, too, requires these abilities.

What then are the possible contributions of memory research to education?

Past meta-theoretical, theoretical, and conceptual contributions

Cognitive psychology in general, and memory research in particular, have provided useful classifications and theoretical frameworks to many aspects of both learning and instruction—the two complementary facets of formal education. The detailed analyses of components of learning and instruction put forward by Gagne (1970), Glaser (1976), and Entwistle (1981) clearly parallel the cognitive theoretical framework developed from basic laboratory research.

The influence of conceptual frameworks in cognitive psychology on this research can hardly go unnoticed. The influence of the organizational school (e.g. Miller, 1956; Bower, 1972) and the levels of processing approach (Craik and Lockhart, 1972) are particularly noticeable in the analyses of student learning.

More specifically, the literature on student learning contains many familiar memory-related concepts from active learning, through various processing modes (e.g. surface and deep processing), up to organization and structure. For example, the idea of deep versus surface processing in memory research, starting with the work of Craik and Lockhart (1972), has parallels in educational research in the discussions of deep versus surface approaches of students to learning. The aim of this research was to distinguish different types of students according to their approach to learning. It resulted in the identification of several major types of students (Marton, 1975; Pask, 1976; Entwistle, 1981; Entwistle and Ramsden, 1983). Using either qualitative (interviews) or quantitative (factor analyses of psychometric measures) approaches, educational researchers found that the surface approach is characterized by intention to complete required tasks, memorization, and failure to distinguish principles from examples. The deep approach is characterized by intention to understand, relating new ideas to previous knowledge, relating evidence to conclusion, and examining the logic of the argument (Marton, 1975).

One further example can be found in the differentiation between hierarchical and serial organization established in memory research. The former is characterized by higher-order relationships, while the latter is characterized by single-dimension relationships (see Bower, 1972). In a similar vein, Pask (1976) distinguished between two general categories of learning strategy that could be identified in cognitive tasks: 'Serialists learn, remember and recapitulate a body of information in terms of string-like cognitive structures where items are related by simple idea links ... holists, on the other hand, learn, remember and recapitulate as a whole' (Pask and Scott, 1972, p. 218). This distinction supposedly reflects qualitative differences among students in their approaches to academic learning.

Clearly, theories of memory coming out of laboratory research have provided useful taxonomies and theoretical frameworks for understanding student learning in education. It should be noted, however, that the adoption of such ideas and frameworks was mostly done on a meta-theoretical level where these general ideas, which were part of the cognitive *zeitgeist*, influenced conceptual frameworks both in memory research and in education. For example, the role of prior knowledge and its effects on encoding of new materials was recognized in memory research starting with the seminal work of Bartlett (1932). Most specific investigations of the effects of prior knowledge, however, started in the 1970s (e.g. Bransford and Johnson, 1972, showing the effects of providing advance information about labels to paragraphs and stories, on memory). In education, Ausubel (1968) suggested the importance of advance organizers as a vehicle to help students incorporate and assimilate new information by providing them with an advanced structure of the materials. Since there was little specific memory research at that time on the effects of prior knowledge, Ausubel was probably influenced by these meta-theoretical cognitive ideas rather by any specific application of them.

To summarize, it appears that, at least in terms of some general approaches and meta-theoretical conceptualizations, there are clear parallels in the development of ideas in laboratory memory research and in educational research on student learning.

In some of these cases laboratory-based memory research provided some of the conceptualization used in education.

Past empirical contributions

To what extent has memory research developed any specific principles and empirical generalizations relevant to educational settings? There are various ways to evaluate this question. One way is to look at practical study guides for college students. The chapter on forgetting and remembering in Pauk's *How to Study in College* (1984) outlines 10 principles to help learning and remembering. Among them are motivated interest, selectivity, intention to remember, basic background, meaningful organization, recitation, consolidation, distributed practice, imagery, and association. Such a list, while probably only partial, is quite impressive. (See also McKeachie, 1986, for similar suggestions for instruction in college.)

If memory researchers are asked for findings relevant to education, they will come up with a respectable list. Bjork (1979), using an information processing analysis of college teaching, derived four principles from laboratory research—the spacing effect, variable encoding, levels of processing, and the importance of structure. In light of numerous laboratory results, Bjork shows that each could be used to help students study, to help instructors design courses, and to help both during lectures. The latter two principles are among the general conceptual ideas mentioned by Pauk.

Such claims, however, should not be accepted automatically. The question is how solid these principles for classroom application are, since they are based primarily on laboratory research? More generally, to what degree are specific empirical generalizations and findings from basic memory research relevant and replicable in applied educational research and settings? This is quite important since we have no assurance in advance that findings in a confined laboratory context will be relevant and actually work in everyday educational settings.

ANALYSIS OF POSSIBLE CONTRIBUTIONS OF APPLIED EDUCATIONALLY ORIENTED MEMORY RESEARCH BOTH TO PRACTICE AND TO THEORY: SOME EXAMPLES

To elaborate on the point, I would like to take a sample of specific principles and findings established in laboratory memory research and evaluate both empirically and conceptually their likelihood of being applied in college learning and teaching. Specifically, I will use some examples from studies in which these principles and findings were applied in real-life situations and were properly evaluated. Furthermore, in analysing these examples I will consider the role of such studies and their contributions to memory research and theory as well as to practice.

The first principle, the spacing effect, is established in laboratory memory research. It is one of the most reliable phenomena in human experimental psychology and refers to the fact that when we compare the effects of two presentations of the same verbal material that are separated (spaced) temporally to a case where these presentations are contiguous in time (massed), long-term recall of learned material is better if the learning episodes were spaced rather than massed. This effect can be marked and has been obtained across words, sets of words, sentences, and in

free recall, cued recall, and paired associates tasks (e.g. Melton, 1970; Rothkopf and Coke, 1963). This effect shows even in implicit memory tests (e.g. Perruchet, 1989). There have been several mechanisms suggested to account for the superiority of spaced presentation; these include rehearsal (e.g. Coccone and Brelsford, 1974), consolidation (Landauer, 1967), attention (e.g. Shaughnessy, Zimmerman, and Underwood, 1972), habituation (Hintzman, 1974), encoding effort (e.g. Johnston and Uhl, 1976), and contextual variability (e.g. Glenberg, 1979; Ross and Landauer, 1978).

The second principle, encoding variability, refers to the fact that information retrieval profits from variability in the encoding of the information. Thus, for example, when the materials to be memorized are presented each time in a different context (external or internal) they will be remembered better than when they are presented in the same context. This principle has been shown to hold across a variety of materials and contexts (e.g. words, Madigan, 1969; sentences, Thios, 1972). One explanation for this phenomenon is that varied encoding of information multiplies the retrieval routes (access paths) available at the time of later recall (Anderson and Bower, 1973).

These empirical generalizations, if applied, have important implications for education, suggesting that students should space their study episodes on a given subject rather than do all their studying in one sitting. Students should, in addition, tie the material to various contexts and not confine it to a specific context. This could be done by relating the materials to as many different everyday and intellectual contexts as possible. In designing their courses, teachers should try to go over new materials more than once and at different times. These repetitions should be tied, if possible, to different contexts (for details see Dempster, 1987).

Distribution of presentations¹ and encoding variability: a study by Smith and Rothkopf (1984)

Although, as mentioned above, these two empirical generalizations have been observed under various learning and test conditions, in most cases they were found in the laboratory (but see Reynolds and Glaser, 1964; Bahrick and Phelps, 1987; Bloom and Shuell, 1981; Dempster, 1987; Rea and Modigliani, 1985). To what degree, then, have we any indications that these variables can be applied to real-life learning situations?

A study by Smith and Rothkopf (1984) assesses their effects, using educationally realistic materials and procedures that in many ways mimic the regular classroom experience of students in both learning and testing situations. In this study a special 8-hour course in introductory statistics (both descriptive and inferential) was developed. It consisted of four lectures videotaped to provide the same instructional content for all conditions. The lectures represented a complete and integrated set of information on the topic. These four lectures were presented within a single day (massed presentation) or spread over a period of 4 days (spaced presentation). To investigate encoding variability the lectures occurred in a single classroom or in four different room settings. Five days following the last lesson, subjects were given a battery of memory tests on the material. These tests included free recall, cued

¹ It should be noted that whereas in the past, research on distribution of *practice*, when different stimuli are used in each presentation, and on distribution of *repetitions*, where the same stimuli are introduced in each presentation, was mostly conducted independently, in the current discussion both distribution effects are treated together. This is so because the manipulation by Smith and Rothkopf actually blends these effects. Furthermore, as will later be discussed, some of the mechanisms suggested to be involved in both phenomena, such as attention, fatigue, or context, are similar.

recall, matching, and problem-solving. The final test session occurred in a totally new room so that the test environment would not serve as a differentially useful retrieval cue. To minimize extra-experimental studying, subjects were not told about the test ahead of time.

Immediate tests after each unit did not reveal any differences between the groups. Results for the free recall test given 5 days after the last lecture, in which subjects were asked to outline the lectures thoroughly listing all the topics and symbols covered in the lectures, are shown in Figure 1. The results are quite clear. There was a strong, statistically significant, effect of days. The 4-day groups recalled about 13 per cent more information than the 1-day groups. There was also a 10 per cent advantage in recall for the four-room over the one-room groups, although this effect did not reach statistical significance.

Results of the cued-recall test showed a significant advantage of the spaced over massed presentation. There were no effects of rooms and days in the matching and problem-solving tests.

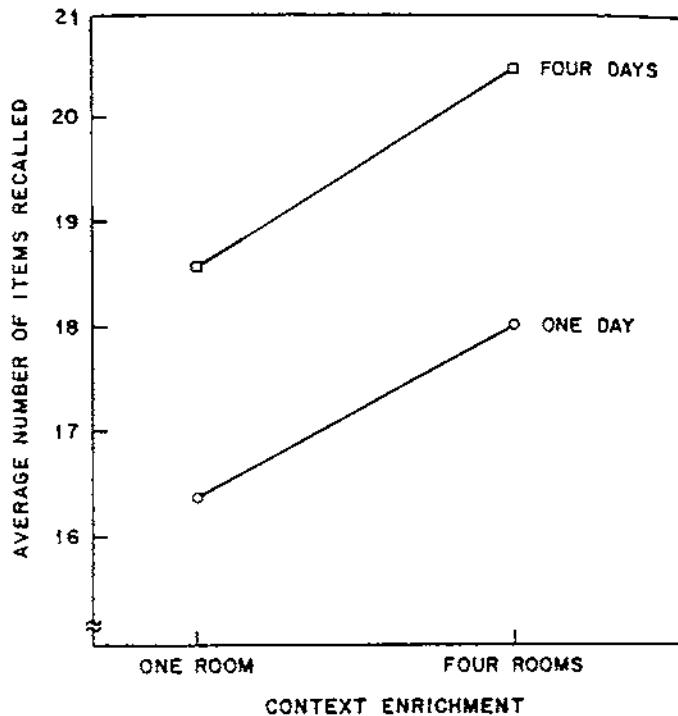


Figure 1. Effects of rooms and days upon number recalled on the general recall test (from Smith and Rothkopf, 1984). Copyright (1984) by Lawrence Erlbaum Associates, Inc.; reprinted by permission

This study represents an exemplary way in which research relevant to education should be conducted. There are several aspects of this study that could make it a model for conducting educationally oriented memory research.

Contributions to education

Demonstration of applicability. The first beneficial aspect of Smith and Rothkopf's study is the demonstration that generalizations (findings) established in the memory

laboratory are relevant to education. Using ecologically realistic materials and settings close to those found in formal education settings, this study confirmed two principles established in laboratory research. Both the spacing effect—temporally massed learning is less efficient than temporally spaced learning—and the encoding variability principle—variable encoding settings lead to better memory of the information learned—were obtained (although this latter effect did not reach statistical significance). Such a demonstration is an essential step in applying and using cognitive principles based on laboratory research. We need to demonstrate that these principles hold in natural settings, but this demonstration, nevertheless, has to be evaluated using the highest methodological standards and scrutiny. Such demonstrations, which indicate generalizability of laboratory findings under conditions high in ecological validity of the method, would probably be endorsed even by those who favour laboratory experiments (e.g. Banaji and Crowder, 1989).

Relevance to educational practices. Such results are extremely important for educational practice, since they imply that these two principles could be useful in real educational settings. Even more importantly, the results can guide us in developing educational programmes. For example, educational practice should take into account the fact that the findings were obtained only for the memory tests but not for problem-solving tests. This is important because teachers and educators are concerned not only with what students memorize in classes, but also with the degree to which the knowledge that students gain can be used to solve problems in other settings.

In addition, the current data, which suggest that the spacing effect is obtained even with non-verbatim repetition, will extend the conditions under which the results can be applied. This is because, both on a microscopic level (a lecture; a reading unit) and on a macroscopic level (a full course; a textbook), only some of the materials are repeated verbatim, while many of the consecutive presentations include new materials (some of which might be related to previously learned materials). The obtained results for non-verbatim information are relevant and could be applied and practised in class, where such non-verbatim repetition is the rule (see Dempster, 1988, for a detailed discussion of the mechanisms which mediate actual educational applications).

Contributions to memory theory and research

Increase in external validity. The fact that we find concepts and findings established in laboratory research confirmed in somewhat different settings outside the laboratory might contribute to the external validity of the phenomena by expanding the range of situations under which the phenomena occur (see also Ross and Landauer, 1978). It implies high generalizability of principles (see Banaji and Crowder, 1989).

Constraints on the scope of laboratory findings. Ecologically valid study can provide us differential knowledge by constraining the findings obtained in laboratory tasks. We expect students to gain in their university education both information and understanding of the materials, including knowledge of main concepts, principles, and definitions. In addition, we expect them to develop knowledge about the procedures involved in using this knowledge. These two types of knowledge were named by Anderson (1976) as declarative and procedural knowledge. For example, in the context of statistics taught in Smith and Rothkopf's study, we expect students to know

symbols (e.g. \bar{X} , H_0 , S^2), statistical terms (e.g. variance, median, confidence intervals)—declarative knowledge—as well as to calculate and compute various values (e.g. 95 per cent confidence interval; independent sample t -test)—procedural knowledge.

The results obtained in this study showed the effects of spacing and encoding variability on declarative knowledge (e.g. free recall). These variables, however, did not affect procedural knowledge (the problem-solving test). So Smith and Rothkopf's study, which utilizes educationally relevant assessment, provides an important constraint on these variables by showing that they do not affect procedural knowledge. Such a result, if confirmed and replicated in further additional studies, might differentiate the usefulness of these variables in educational settings. While original laboratory studies, which until recently employed, almost exclusively, tests tapping declarative knowledge, showed the spacing effect not to be too sensitive to the test used, we have here a problem-solving test, tapping procedural knowledge, in which subjects do not show the usual spaced-presentations advantage.

Extension of conditions under which original phenomenon occurred. This study could exemplify the possible contribution of educationally valid research in extending the conditions under which the findings might be obtained. Notice an important difference between traditional research on spacing and contextual variability and Smith and Rothkopf's research: traditional research, almost exclusively, considered these effects for verbatim-repeated events—that is, materials presented a second time or in a different context were usually the same nominal stimuli (but see some exceptions, Glenberg and Lehmann, 1980; Dellarosa and Bourne, 1980; Glover and Corkill, 1987; Ross and Landauer, 1978). However, in Smith and Rothkopf's study the spacing effect was obtained even when the materials presented in each day or in each environment were not *identical* to the ones presented earlier, although some concepts naturally were repeated in the multiple lectures. This change from the laboratory definition is a result of conducting educationally oriented research. Conditions in the educational settings are often different from those in the laboratory. In natural settings teachers usually do not repeat materials verbatim during the course (although some overlap occurs and might even be necessary). So we find that educationally oriented memory research, by its nature, might extend laboratory findings to conditions other than those originally established in the laboratory.

Relevance to theoretical controversies. From a theoretical point of view, the results obtained in this study, which show the principles to apply to non-verbatim repeated events, might shed some light on the theoretical controversies over the mechanisms underlying these effects. As mentioned earlier, several mechanisms have been suggested to explain the spacing effect. Among them are habituation, rehearsal, inattention, consolidation, and contextual variability (Hintzman, 1974, 1976). We might group these theories into three major categories. The first category attributes the disadvantage of massed repetition to interference of the second presentation with the ongoing processing of the first presentation (rehearsal and consolidation hypotheses). Theories in the second category attribute the disadvantage in performance under massed presentation to less processing of the second presentation (habituation, inattention). The third category, which includes the contextual-variability hypothesis, attributes the disadvantage of massed presentation to a lack of change in context.

According to such an explanation, if we assume that retrieval of different contexts might help in recovering the information, the advantage of the spaced repetition is due to the increased probability of encoding the two appearances of the same stimuli under different contexts. Note that all but the last theory are strictly encoding and storage theories; the contextual variability hypothesis has a substantial retrieval component.

The results of the Smith and Rothkopf study, at least on face value, suggest a mechanism responsible for the spacing effect. In particular, the fact that non-verbatim repetition resulted in replication of the standard spacing effect can be less easily explained by encoding theories because the materials presented, which are different at each time, should not cause any *specific* interference with the ongoing processing of the first stimuli (rehearsal or consolidation explanations). In addition, because the material is *different* at the second presentation, it should not cause less processing allotted to the second presentation (habituation, inattention). An explanation which involves a retrieval mechanism, however, is clearly viable in a non-verbatim repetition case. Materials in the spaced condition are coded under different contexts regardless of whether the materials on the second presentation are similar to or different from those used in the first presentation—this increases their chances of being retrieved later. So the contextual variability explanation in the current study is the most parsimonious explanation (see similar results by Glenberg and Lehmann, 1980; but different ones by Ross and Landauer, 1978).

Such a conclusion should be evaluated carefully, because the repetitions in this study were neither completely verbatim nor completely different, but rather probably partially redundant. This is because in each new lesson (of the four included in the course), while new materials were introduced, some of them were introduced in the context of the materials already learned, causing some relearning or at least reminding of the previously learned material. This partial redundancy characteristic of repetition is probably the most prevalent and typical way in which it occurs in everyday life, and particularly in educational settings. Repeated events usually tend to appear with new information, and new events usually appear in the context of familiar information. (In school, new materials learned are usually tied, at least partially, to previous information learned). It is important to note that such a natural parameter of the spacing effect, combining old and new information (partially redundant repetitions) has hardly been explicitly investigated in the laboratory. Illumination of relevant parameters which could be further analytically studied in the laboratory, might be one contribution to memory research of studies like the one by Smith and Rothkopf, characterized by high ecological validity of its method (Banaji and Crowder, 1989).

Contributions to both theory and practice: considerations of individual differences

An important aspect of the Smith and Rothkopf study that is usually missing from laboratory memory research is the use of individual-differences variables. Laboratory-based memory research has always been concerned with studying and establishing basic cognitive mechanisms. The assumption is that even if there are individual differences, the basic mechanisms of cognition are similar across individuals (although there are exceptions to this view; e.g. Cofer, 1967). As a result, empirical principles and generalizations are held to apply uniformly to all normal populations.

The study by Smith and Rothkopf provides a good example of the problematical

nature of such an approach. The individual differences variable used in their study was field dependence-independence, considered to be one aspect of a general personality dimension defined as an analytic (field-independent) versus a global (field-dependent) orientation to the world. Subjects in each of the experimental conditions were split according to their score on the Group Embedded Figures Test (GEFT—Witkin, Oltman, Raskin, and Karp, 1971).

Figure 2 presents the effects of spacing the materials and of encoding variability of learning on free recall for field-dependent and field-independent subjects. The expected effects of both variables were obtained for the field-dependent subjects only. There was neither an advantage of the 4-day nor of the four-room conditions for field-independent subjects. One possible explanation for this pattern of results related to field independence theories (Goodenough, 1976) is that field-independent subjects may have focused on the instructional stimuli alone, ignoring the environmental context in which the stimuli were presented. Field-dependent subjects could not ignore the context, and therefore contextual information was encoded, thereby improving free recall performance. (Another possibility raised by the researchers is that the differential use of contextual information occurs at the time of retrieval, resulting in the obtained pattern of results.) Such a result limits the spacing and encoding variability effects obtained in the laboratory to field-dependent subjects only.

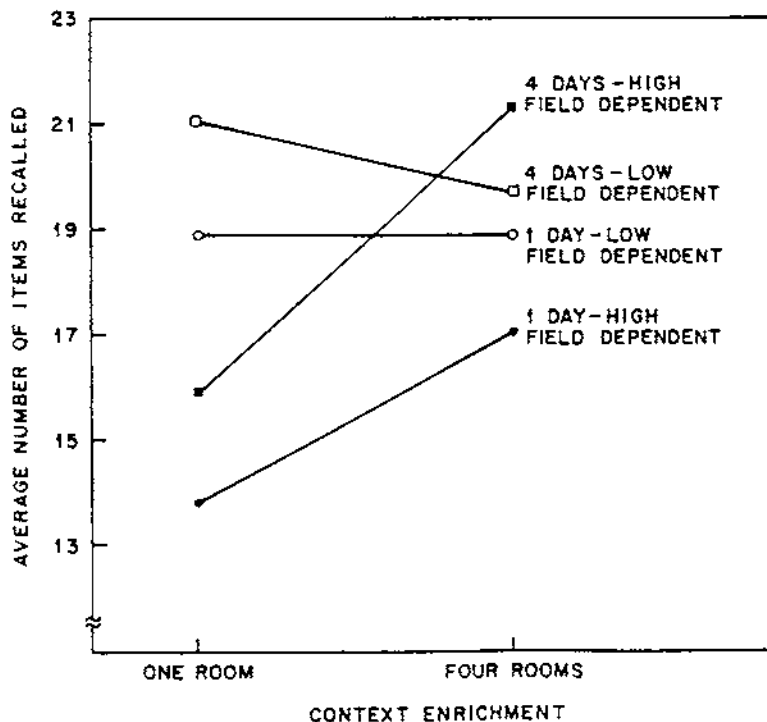


Figure 2. Effects of days, rooms, and field dependence on the general recall test (from Smith and Rothkopf, 1984). Copyright (1984) by Lawrence Erlbaum Associates, Inc.; reprinted by permission

The significance of these individual-differences interactions with major laboratory-based variables should not be underestimated. It is theoretically possible that many of the important principles established in laboratory-based research are qualified by such individual differences and might apply only to certain segments of the population. Because of the nature of basic research, which mostly ignores individual differences, these patterns of interactions have not been revealed.

Knowledge about interactions with individual differences might contribute to both practice and theory. From a practical point of view such interactions have important implications for education. For example, it appears that principles of both encoding variability and of spacing are relevant and should be applied for field-dependent students only. As a result, course designs and lecture structures that stress spaced repetitions of concepts in different contexts would be particularly helpful for these students but would not affect the learning of field-independent students. The same knowledge would be helpful for programmes on how to study, which should discourage massed cramming and organizing materials in only one manner, especially for field-dependent students. (The results of Smith and Rothkopf indicate that field-independent students will not gain by adopting these practices.)

From a theoretical point of view, considerations of individual differences might both constrain and support theoretical formulations. On the one hand, these results constrain the external validity of the spacing and encoding variability effects. On the other hand, such individual differences could provide a test of laboratory-based memory mechanisms. Such tests might refine our understanding of these mechanisms. In the current case the results support the contextual-variability explanation as the underlying mechanism for the spacing effect. This is because attention to context, as a major difference between field-dependent and field-independent students, was developed independent of the spacing effect, and the pattern of results obtained supports the contextual explanation of this effect. This is a clear demonstration of Melton's (1967) suggestion of studying individual differences in terms of the process constructs of contemporary theory.

To summarize, conducting ecologically valid research in which the materials to be studied, the settings, and the evaluation are as similar as possible to real educational settings, has several benefits. First, it can serve as a demonstration that principles established in the cognitive laboratory are relevant to education. Second, the results of such studies can increase our confidence in using the principles in several facets of educational settings, including learning, curriculum, course design, and instruction. Third, because ecologically oriented study by its nature uses some specific parameters not necessarily identical to those used to formulate the original theories, its results could expand the conditions under which some well-documented laboratory phenomena are obtained. Fourth, the results of such studies can have implications for theoretical formulations developed on the basis of laboratory research. Fifth, such studies might constrain the knowledge obtained from laboratory tasks. Finally, such studies might provide us with information about important variables (e.g. individual differences) that traditionally were not investigated in the laboratory but that are one of the hallmarks of ecologically oriented approach (see Bruce, 1985). Findings about such variables might have important implications for applied educational settings as well as for theoretical formulations related to these variables.

Two cautionary notes are needed. First, conducting a study in ecological settings might have its methodological drawbacks. For one thing, at the end of each session

of Smith and Rothkopf's study, students solved some problems related to the materials studied in the session. These could have served as additional uncontrolled repetitions. In addition, the multiplicity of the tests, always conducted in the same order, could have created some practice or fatigue effects affecting the later tests.

Second, this study still was not fully ecological since the unexpected final test is different from regular class settings where final exams are generally expected. In addition, subjects received money for their participation; hence their motivation might have been different from real college learning settings where students probably study for different reasons.

Maintenance of knowledge: a study by Bahrck (1984)

A second example of the benefits of educationally relevant research is Bahrck's (1984) study of the maintenance of knowledge of material learned in school. In this study 773 individuals were tested to see how well they remembered Spanish learned as long as 50 years ago. Tests of reading comprehension, recall and recognition vocabulary, and grammar were administered together with a questionnaire to determine the level of original training, the grades received, and rehearsals during the retention interval in the form of reading, writing, speaking, or listening to Spanish.

The study is especially relevant to educational practice, partially because the learning stage, in contrast to laboratory research, took place in natural settings during the course of learning a foreign language in high school and in college. In addition, the retention interval—up to 50 years—makes the study ecologically valid because use of a foreign language does not necessarily occur immediately after learning it, as is often the case in retention tests in the laboratory. Finally, although the tests took place in the laboratory and not in natural settings, the variety of the tests used increases the confidence that they represent tasks that might be faced by the learner in the real world.

Several aspects of the results of this study are relevant to our discussion. One is whether the principles established in the memory laboratory apply to education. There are two characteristics of the results that seem to support such a claim. The first addresses the importance of the initial level of training (measured here by the number of Spanish courses taken). Laboratory-based memory research has repeatedly shown the importance of level of initial training in determining memory performance (e.g. Cooper and Pantle, 1967; Underwood, 1964). Examples of this principle in Bahrck's research can be seen in Figure 3, which presents information about the *learning* curves across subjects for grammar, idioms, and word-order information in Spanish as a function of initial training level. It is clear that initial training has a significant effect on performance. Moreover, Figure 4 shows the *retention* curves for idiom recognition as a function of initial level of training. It demonstrates characteristics of the trends for all other measures, in showing the advantage in retention as a function of initial level of training, but noticeably, with no interaction with retention interval (see Underwood, 1964).

The second relevant aspect of this study is the effect of the type of retrieval task on performance. Figure 5 shows retention functions for recall and recognition vocabulary and for reading comprehension (adjusted for various independent measures, such as initial training level, amount of rehearsal, etc.). It is generally evident in this figure that performance on recall tasks was more sensitive to the retention inter-

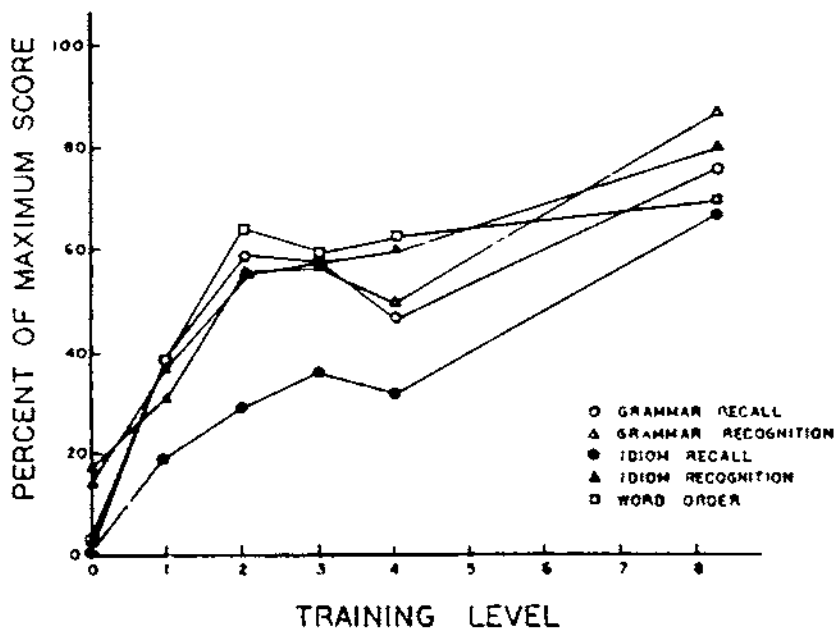


Figure 3. Learning curves for grammar, idioms, and word order (from Bahrick, 1984). Copyright (1984) by the American Psychological Association, Inc.; reprinted by permission

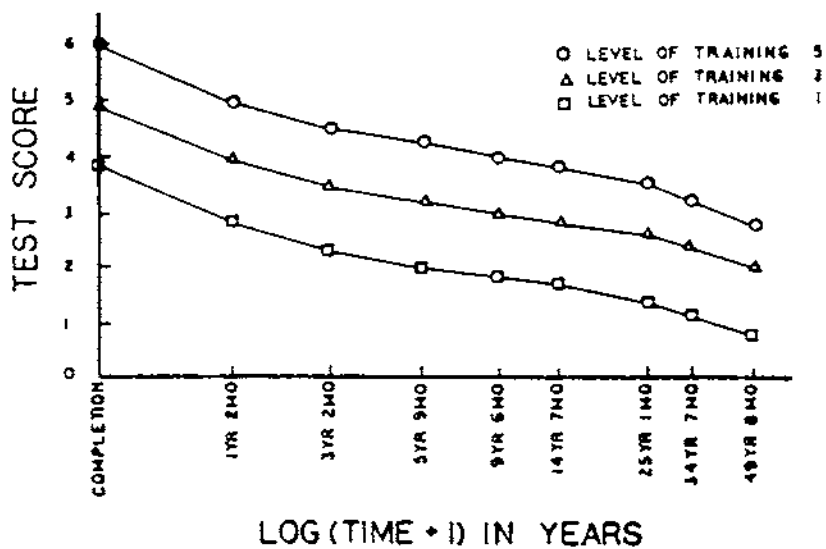


Figure 4. The effect of level of training on the retention for idioms (from Bahrick, 1984). Copyright (1984) by the American Psychological Association, Inc.; reprinted by permission

val, while performance on the recognition tasks showed only mild changes as a function of time. Such an advantage of recognition over recall performance is a well-established phenomenon in laboratory memory research (e.g. Postman and Rau, 1957).

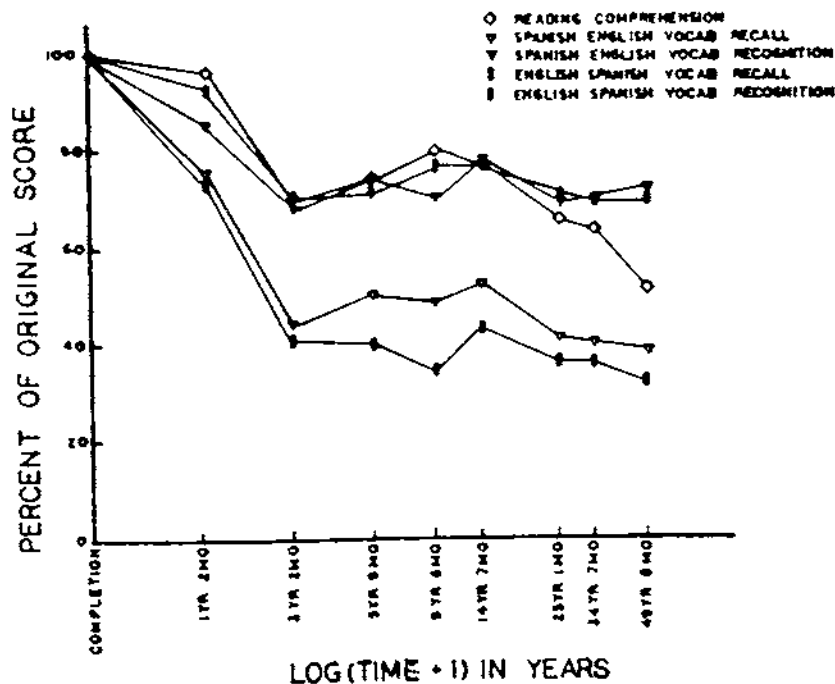


Figure 5. Adjusted retention functions for recall and recognition vocabulary and for reading comprehension (from Bahrck, 1984). Copyright (1984) by the American Psychological Association, Inc.; reprinted by permission

To summarize, this study demonstrates that principles established in the cognitive laboratory are relevant and can be extended to educational settings—in this case to the retention of knowledge of a foreign language studied in school. It is evident that the initial level of training and type of test have the same effect on educationally relevant material that they have in the laboratory. This is so in spite of the fact that we are dealing with semantic knowledge acquired gradually over time, and not with laboratory-based episodic knowledge. Such results should increase our confidence in using these principles in various facets of educational settings.

Some characteristics of the results of this study do not, however, reproduce laboratory research and might constrain its scope. One result relates to retention curves. If we look at Figure 6, which presents retention of Spanish-English recall vocabulary, it is evident that forgetting is not always of the exponential variety usually obtained in laboratory research of episodic memory (e.g. Ebbinghaus, 1885; Wickelgren, 1974). Under this study's conditions, forgetting educationally relevant material (a more semantic memory type of knowledge) is characterized by an initial exponential decline

for the first 3–6 years. After this period retention remains unchanged for a period of up to 30 years, before showing a final decline.

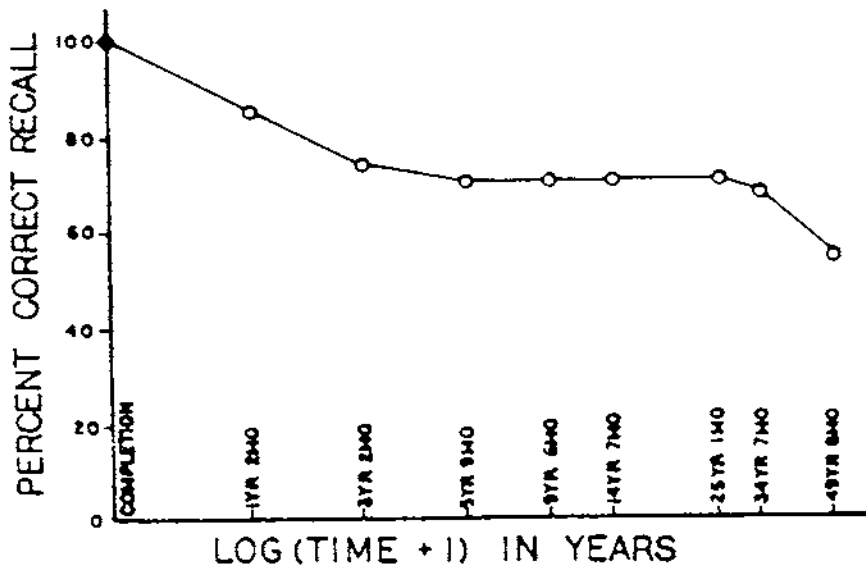


Figure 6. Adjusted retention function for Spanish-English recall vocabulary (from Bahrick, 1984). Copyright (1984) by the American Psychological Association, Inc.; reprinted by permission

Once again educationally oriented study using realistic materials might constrain knowledge obtained in laboratory research. This contribution to basic memory research was the result of using educationally relevant parameters in the research. Educators are hardly concerned about retention for 3–10 minutes, the favourites of memory researchers. The choice of long retention intervals, which are relevant and important for educational settings that seek to establish life-long learning, again resulted in new empirical generalizations different from those obtained in laboratory settings. The use of such parameters might extend and refine the scope and generalization of laboratory-based memory research.

EVALUATING THE POTENTIAL APPLICABILITY AND RELEVANCE OF LABORATORY RESEARCH AND MEMORY PHENOMENA TO EDUCATION: SOME CONSIDERATIONS AND SUGGESTIONS

So far we have looked at some examples of educationally relevant memory studies and have shown such studies to result in patterns of findings that are both similar to and different from those obtained in the laboratory. Now we must consider whether there might be some rules or conditions that might allow us to predict when laboratory results will hold up in natural settings. We must ask how memory researchers can make their research more educationally relevant. In the following section several suggestions will be made for such guidelines.

Evaluation of differences in relevant parameters between laboratory and educational settings

How well laboratory research can be applied to natural settings depends to some degree on the extent of the differences between the conditions under which the experiments were held in the laboratory and the conditions in the field, where learning and instruction take place (school, college, etc.) The larger the differences, the slimmer the chance that results obtained in the laboratory will extend to the field.²

Let us look at some categories of differences between laboratory and the field.

Materials

A predominant portion of the memory laboratory research, until recent years, used very simple instructional materials such as nonsense syllables and single words. What students learn in school, however, is usually much more complex and structured (see Howe and Ceci, 1979). These differences are not only quantitative but also qualitative in nature. The importance of the types of materials chosen can be illustrated by Bartlett's work (1932). Using more ecologically valid materials (stories) he completely changed prevailing views about memory mechanisms, showing the constructive rather than the reproductive nature of memory.

Settings

The laboratory is perceived by subjects as an artificial environment. In addition to the lack of the group influence and interaction found in the classroom, subjects' motivation and affective reactions may change. As a result they do not behave in it as they do in real life. This can affect their motivation to perform. In real educational settings where their performance might determine their future, they are generally more highly motivated than they are for laboratory research. These motivational differences can put laboratory subjects at a completely different point on the Yerkes-Dodson arousal-performance curve (Yerkes and Dodson, 1908). In real life, higher motivation can help students' attention and learning efficiency, or it might cause them to be too aroused and anxious, resulting in lowered performance.

Similarly, laboratory settings may minimize affective reactions. This might be beneficial for laboratory memory researchers in helping them to establish a clearer picture of the relationships among variables studied. As a result, however, such studies might miss important components that affect students' learning in real life—the affective and motivational components related to learning. So while memory researchers limit themselves to investigating apparent cognitive mechanisms, educational researchers are interested in other facets of student learning, including affect, motivation, and personality characteristics. This is important because it gets at the heart of educational researchers' cautiousness about laboratory memory research. It also touches the question of the interaction of cognitive and non-cognitive variables.

Let me elaborate on this point. A major purpose of people involved in various aspects of higher education is to learn about patterns and mechanisms of student learning and instruction. If you look at possible determinants of successful learning

² Such an approach is not incompatible with Mook's (1983) claims against 'count-'em mechanics: the more differences there are, the greater the external invalidity'. As Mook himself states: 'Of course there are also those cases in which one does want to predict real-life behavior directly from research findings. Survey research, and most experiments in applied settings such as factory or classroom, have that end in view. Predicting real-life behavior is perfectly legitimate and honorable way to use research.'

in higher education (and in education in general), it is clear that pure cognitive mechanisms, such as those investigated in the cognitive laboratory, are just one determinant, and probably not the most important one. The outcome of learning at the college level is related to knowledge and intellectual abilities, domains of cognitive research. This outcome, however, is also greatly affected by personality characteristics, individual cognitive style and motivation (see Entwistle, 1987). For example, individual learning style is extremely important in determining success in higher education. Learning style can be considered to be an expression, in the academic context, of more fundamental, and relatively stable, components of cognitive style and personality. Personal motives are extremely important and influence both the direction and the quality of the efforts exerted in studying (Taylor, 1983). One of the main conclusions derived from the factor analyses in a study of student approaches to learning by Entwistle and Ramsden (1983) was that underlying personality traits associated with the tendency to prefer specific styles of learning are very important for academic success.

Subject population

While this parameter might be relevant to considerations of ecological research in general, in the case of education, especially higher education, it is less crucial. This is because the natural population from which cognitive researchers draw their subjects consists of university students, exactly the population we wish to generalize to. We should not, however, forget that most participants in psychological research are introductory psychology students, who are not necessarily similar to other students in their learning skills or other cognitive characteristics. Such differences among subject populations can clearly have a strong effect on performance in educational settings.

Tests

Most laboratory memory research uses prototypical memory measures, such as recall or recognition. In education, however, the variety of measures is larger and includes measures of problem-solving. In addition, school curriculum, characterized by the partial overlap in material in different courses, makes savings from one course to another an important determinant underlying students' performance. However, measures of savings (or relearning) mostly have been given attention by memory researchers either a long time ago (e.g. Ebbinghaus) or only very recently (implicit memory measures).

Criteria for evaluating the applicability of memory research to educational settings

Criterion 1: Choice of relevant parameters

As a rule of thumb we should expect that the smaller the differences in the above parameters between laboratory research and the field, the higher the chances that laboratory research will be relevant to learning and instruction in school. In this sense, memory researchers interested in education should try to decrease the differences in the above-mentioned characteristics between the settings in which they conduct their research and those to which they are interested in applying their research—school or college. They can do this by choosing parameters as close to those in the actual educational settings as possible.

To increase the ecological validity of laboratory experiments we should carefully evaluate the components of learning in higher education and then devise research to study these components. There are many aspects relevant to student learning settings that laboratory research fails to address. For example, very few studies have investigated the characteristics of learning and retention of material over relatively long periods of time (e.g. a semester). We also know very little about learning that combines different sources (lecture, textbook, discussion section, examinations). Similarly, while in many of our laboratory experiments we run subjects individually, university learning is often characterized by active group studying. In addition, we know little about the effects of studying various topics in parallel as happens in a regular semester where the student takes several courses simultaneously.

The importance of real-life characteristics of the laboratory should not be underestimated. As mentioned above, Bartlett's theoretical revolution, one of the leading forces in the shift from behavioural verbal learning to cognitivism, was based, on a change in one of these parameters—the materials. His conclusions were based on observations done when meaningful, ecologically valid materials were used instead of the mostly artificial materials used earlier. We might find as dramatic findings as he did when changing other aspects of the laboratory situation discussed earlier—the settings, the subject population, and the tests (see also Bugelski, 1971).

A related formulation in the memory domain was presented by Jenkins (1979) under the title of the 'problem pyramid' or the 'theorist tetrahedron'. In this model Jenkins offers a classification scheme for memory experiments. According to this classification the phenomena of learning and memory that we observe depend on the nature of the subjects, the orienting tasks, the kinds of materials used, and the criterial tasks employed. Even more important, according to Jenkins, the results of any particular experiment are context-sensitive in the sense that they will reflect a *certain* interaction of these factors. Changing the parameters of any of the factors might change the pattern of results.

In applying Jenkins's model to education, before worrying about the various possible interactions between the components that determine the research plan, we should worry about the values of the parameters chosen in the laboratory and their applicability to everyday life (also see Brown, Bransford, Ferrara, and Campione, 1983). For example: were the materials chosen of any relevance to what students learn in universities? The materials studied in higher education are much more complex, and on a much larger scale, than the ones used in laboratory research. Similarly, are the tests employed of any relevance to educational settings? What about the settings? Are they relevant to the intentional nature of student learning? The greater the deviations in the parameters from the educational settings to the laboratory, the smaller the chances that we will be able to generalize from the laboratory to the field. This point relates to the one made by Jenkins while discussing the possible paths available for memory researchers. According to him, one of the paths is 'concentrate on the range of variables involved in particular problems of importance to us for ecological practical reasons' (Jenkins, 1979, p. 444).

Criterion 2: Considerations of relevant interactions

Beyond the deviation of each dimension of our studies from educational parameters, we must consider the interaction of these dimensions, as Jenkins (1979) suggested. Some interactions in education might be built in to the situation. For example,

students are different from each other in terms of their abilities, interests, etc. Tests, too, vary depending on the situation. In some educational settings evaluation is done using simple memory indices, while in other settings evaluation may require higher-level problem-solving. Let us look at a sample of these interactions.

Interactions between subjects and materials. The interaction of students' cognitive characteristics with the types of materials presented is part of the aptitude-treatment interactions framework recognized some time ago in the educational literature (e.g. Cronbach and Snow, 1977). Different students possess different aptitudes, and each of them might benefit from a different type of treatment (structure of materials). To choose intervention programmes designed to help students (e.g. how to structure the information presented), we need to know something about their aptitude. In addition to different levels of cognitive aptitude, students show different conative and affective characteristics (Snow, 1989). For example, the anxiety level of students affects the conditions under which they can learn. While low test-anxious students benefit more from less structured materials, high test-anxious students perform best when the materials are tightly structured (Smith, Wood, Downer, and Raygor, 1956). Such interactions must be studied and understood.

Interactions between materials and tests. An elegant study showing the importance of the interaction between the organization of the material with the type of test used was conducted by Mannes and Kintsch (1987). These researchers evaluated two theories on ways to improve retention of prose. One theory claims that advance organizers make comprehension easier and this helps retention. This theory draws on work by Ausubel (1968), and claims that providing an advance organizer (structure) to the materials to be learned gives students an appropriate schema to facilitate learning. The second theory holds that making comprehension harder by creating intra-task interference will help in later retention. This theory draws on levels-of-processing-memory research. This research shows that encouraging deep, multiple encoding will help retention (Craik and Tulving, 1975).

In Mannes and Kintsch's study, subjects examined an outline that gave them relevant background knowledge before they read an experimental text. This outline was organized either consistently or inconsistently with the structure of the text. Subjects then performed a variety of tests, either immediately following or after 2 days. Results showed that students who received the consistent outline performed better on such memory tasks as cued recall, recognition, and summary writing. Students who received the inconsistent outline, however, performed better on inference verification and creative-problem-solving tests. So it appears that a consistent outline is important in remembering a text, but an inconsistent outline is better in learning from a text. It seems reasonable to infer that the way the instructors structure the materials interacts with the type of test involved. Such important interactions have recently received the attention of memory researchers as indicated by notions like transfer-appropriate processing (e.g. Morris, Bransford, and Franks, 1977).

Summary. The above examples provide us with two basic suggestions for the application of memory laboratory research to education. First, we have to examine the degree to which laboratory research uses parameters close to the natural ones in educational settings. The closer the studied parameters are to those found in the

field, the greater the chance that the results will be applicable. Second, we have to be aware not only of the correct parameters but also of the interactions between various categories of these parameters. In education, successful learning and instruction depend a great deal on the interaction of treatments with task demands, with materials, and with subjects' characteristics. While it might not be easy to conduct memory research that takes into account all the above considerations in designing studies, memory researchers interested in education should consider them as they design studies. Even more important, they must realize the possible limitations placed on their studies when some of these considerations are not addressed.

Criterion 3: Effect size in the laboratory

Another criterion that could predict the generalizability of laboratory research to educational settings is the size of the effect obtained in the laboratory (or the variance accounted for by the studied variable relative to other variables studied). As a rule, the stronger the effect in the laboratory the higher the chances that this effect will show up in educational settings. This happens because the addition of each variable in real-life settings increases the chances that the effects obtained in the laboratory will be weakened. That is, as more variables that are related to the dependent measures have their influence in the educational settings, the more variance they will explain—variance originally explained by the laboratory-manipulated variable. Other things being equal, the stronger the original effect, the better the chances that even under multiple variable educational settings it will survive.

Criterion 4: Variation of ecologically valid variables in laboratory research

At times it might be better to let some variables in the laboratory vary rather than control them. This makes the research more similar and relevant to natural settings. To exemplify this point let us refer back to the study by Mannes and Kintsch (1987), mentioned earlier.

This study deals with reading scientific texts, something often done in universities. The length of the text (4.5 pages) was comparable to that of a short scientific article. To what degree can we apply their results to educational settings? Notwithstanding the advantages mentioned earlier, this study has some problematic aspects related to the settings. Students read the passage on a computer monitor. A couple of sentences were presented each time in a consecutive manner with an exposure duration of 250 ms per word. This is hardly the way students read in school. Usually students read from books with a full page in front of them at a given moment, and with control of the time spent on each paragraph. In such settings they can go back and forth in the text to clarify difficult points. In other words, natural reading is much more flexible than the type used in this experiment.

What might be the consequences of such discrepancy between natural settings and those used here? To answer this question we must consider the possible outcomes of the interaction between settings and the tests used. Do we have any good reason to suspect that using normal reading habits will change the pattern of results obtained? To what degree could we generalize to school settings the interaction obtained here between materials and tests, interaction that showed the effects of different presentation modes to be dependent upon the type of test used—memory versus problem-solving?

One way to answer this last question is to look at the literature to see whether

'natural' and 'unnatural' reading styles result in different outcomes. The more evidence we find for differences between the two, the more we must be sceptical about generalizing from the study.

Why did the researchers use such artificial reading conditions? The authors mention providing experimental control over reading times as the main reason. Is such a rationale reasonable? To control the effects of an important variable, the authors use a different parameter from the one usually used by students. Is there a reason to control reading time? Had reading time not been controlled, there could have been differences between the conditions on this variable, differences which the Mannes and Kintsch study could not detect. It is important to know whether such differences actually exist because they might provide another index to evaluate the different treatments. One learning condition might result in a better memory performance but also in longer time to learn. The experiment by Mannes and Kintsch does not provide such information.

This example raises a more fundamental issue. Assume that we are manipulating one variable in the laboratory and are interested in observing its effects on a given measure. If there is another variable that might be important to the phenomenon studied, should we control it, manipulate it, or allow it to vary? Table 1 illustrates this question.

Table 1. Possible experimental designs

Design	Variables		
	A	B	C
I	Level 1	Constant	Measure
	Level 2	Constant	Measure
II	Level 1	Level 1	Measure
	Level 2	Level 2	Measure
III	Level 1	Measure	Measure
	Level 2	Measure	Measure

In design I we manipulate the values chosen for variable A and measure their effects on variable C. This is done while controlling the variability of variable B by choosing a particular single constant level of this variable. This is the standard experimental model used in the Mannes and Kintsch study. Under reasonable assumptions it allows us to infer causality in the relations between A and C. That is, if we find systematic differences in C as a function of the different levels of A, we attribute these differences to the different A treatments. The obvious disadvantages of such a design are two-fold. First, the choice of a particular value as a constant in variable B may be quite arbitrary. Second, the variability of this variable is limited. This variability might be important, especially in a natural setting. Such variability may have important effects on variable C, effects that will not be observed using such a design.

If we want to replicate educational settings more closely, we should not arbitrarily choose a constant value for variable B. We might decide to vary (manipulate) this variable by choosing several levels for it. This case, the standard factorial design with two variables, is presented in design II. In this design we manipulate both

variables A and B and observe their effects on variable C. The advantage of this design over design I is in that it allows for more variability in variable B, making the settings more similar to the educational ones. In addition, such a design enables us to assess the effects of variable B and its interaction with variable A on variable C, something we could not do in design I. However, one critical disadvantage of such a design is the arbitrary choice of the levels of variable B. We may not choose the 'correct' or 'natural' levels of this variable as they occur in everyday life.

Design III corrects this problem. In this design we also allow variable B to vary, while measuring its values. This design can allow us (for example, by using multiple regression techniques) to assess the effects of both variables A and B on variable C, while leaving variable B to take values freely according to its distribution in educational settings. Clearly, we can still evaluate the effects of A alone on C, as in designs I or II, using statistical techniques like regression analyses and causal modelling (e.g. Kenny, 1979).³

To make the study of Mannes and Kintsch (1987) fit design III, we could allow reading time to vary while measuring it. This would make the settings much more realistic because the different ways the materials are organized could actually yield different reading times in the different conditions. Such a design would allow us to measure the effect of type of organization on reading time as well as on memory and problem-solving measures. In addition, such a design would allow us to assess the role of reading time as a mediating variable between type of organization and memory performance. The contribution of a given variable in educational settings is dependent not only on its potential ability to influence (as measured in the laboratory), but on its actual contribution to the variance in real settings above and beyond the contribution of other relevant variables.

In contrast to most current laboratory procedures, design III should generally be used for memory research that intends to have useful implications for educational settings. Such a design would allow us to observe natural variability like that found in the classroom, and at the same time give us the opportunity to evaluate causality. We may resort more to the use of sophisticated data analyses techniques available that can allow us today, under reasonable assumptions, to understand structural relations in a given research area when pure controlled laboratory research is not used.

This does not necessarily mean that memory research should be done without any controls, but it suggests that whenever possible such consideration should be taken seriously. Memory researchers, who historically have tended to use sophisticated research designs and experimental controls, might benefit from using more statistical control that lets them study more educationally relevant variables.

To summarize, the above suggestions might serve as guidelines to predict the usefulness of laboratory research in educational settings. Memory researchers who are interested in making their research results applicable educationally should consider these guidelines when planning their research. Clearly, memory researchers

³ More sophisticated designs than design III might be needed under some circumstances. For example, in cases where we are not confident that random assignment of subjects to the experimental groups resulted in equally balanced groups with respect to reading time, normal reading speed could be measured prior to assigning subjects to the different groups, and this measurement could be used later as a covariate. Another possibility would be to use causal models to generate specific hypotheses that can then be tested.

cannot take into consideration all these factors, but they should try to incorporate into their studies those factors that are relevant to the questions being asked. At the very least they should be aware of and consider these factors while conducting their research.

SUMMARY

There is a broad spectrum of educational problems to which memory researchers could contribute. Such contributions could benefit not only educational research and educational practice, but memory research itself. It appears that guidelines that integrate the rigorosity of the methods used in the laboratory by memory researchers with those variables found by educational researchers to be important (including individual differences variables and affective and motivational variables) will let memory researchers conduct well-thought-out and well-run laboratory and field experiments. These experiments would produce results and conceptions that could be applied in educational settings.

In addition, considering educational applications might serve as a catalyst for interdisciplinary laboratory research in psychology. As mentioned, student learning depends on many variables, among them cognitive, motivational, and emotional ones. An interdisciplinary approach in which various researchers in psychology work together in laboratory studies to study the interactions of cognition, motivation, and emotion will be better rather than each investigator being isolated in his or her own narrow area of interest. Here education might serve as a catalyst for cohesive theoretically oriented interdisciplinary research within psychology. This might offset some of recent trends of extreme specialization. The benefits then will be not only for understanding and improving student learning but also for better and more comprehensive theories in psychology. (See for example, Covington and Omelich, 1988; Snow, 1989; Tobias 1985, 1988.)

Educationally relevant memory research that considers real classroom context—where learning and retention involve motivation, affect, and individual differences—could benefit both educational practices and memory theory. Such considerations of relevant variables might strengthen existing models rather than changing them (e.g. see Bower and Cohen, 1982 for a way to incorporate emotions into a network-type model).

McKeachie (1974) claims that one of the major reasons for the failure of the laws of learning to provide an adequate explanation in educational situations was that they did not take into account important variables that are controlled in laboratory situations but that interact with other independent variables in natural educational settings. For the laws of memory not to repeat the same course, we have to consider these other important variables when conducting our research.

ACKNOWLEDGEMENT

I would like to thank Mary Jocelyn and Darryl Bruce for their comments on an earlier version of this manuscript. Parts of this paper were written while the author was a visiting scholar at the National Center for Research to Improve Postsecondary

Teaching and Learning (NCRIPTAL) at The University of Michigan. Requests for reprints should be addressed to Moshe Naveh-Benjamin, Department of Behavioral Sciences, Ben-Gurion University of the Negev, PO Box 653, Beer-Sheva 84120, Israel.

REFERENCES

- Anderson, J. R. (1976). *Language, Memory, and Thought*. Hillsdale, NJ: Erlbaum.
- Anderson, J. R. and Bower, G. H. (1973). *Human associative memory*. London: Wiley.
- Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. New York: Holt, Rinehart, & Winston.
- Bahrick, H. P. (1984). Semantic memory content in permastore: fifty years of memory for Spanish learned in school. *Journal of Experimental Psychology: General*, 113, 1-35.
- Bahrick, H. P. and Phelps, E. (1987). Retention of Spanish vocabulary over 8 years. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 13(2), 344-349.
- Banaji, M. R. and Crowder, R. G. (1989). The bankruptcy of everyday memory. *American Psychologist*, 44, 1185-1193.
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.
- Bjork, R. A. (1979). Information-processing analysis of college teaching. *Educational Psychologist*, 14, 15-23.
- Bloom, K. C. and Shuell, T. J. (1981). Effects of massed and distributed practice on the learning and retention of second-language vocabulary. *Journal of Educational Research*, 74(4), 245-248.
- Bower, G. H. (1972). Mental imagery and associative learning. In L. Gregg (ed.), *Cognition in learning and memory*. New York: Wiley.
- Bower, G. H. and Cohen, P. R. (1982). Emotional influences and thinking: data and theory. In M. S. Clark and S. T. Finke (eds), *Affect and cognition*, Hillsdale, NJ: Erlbaum.
- Bransford, J. D. and Johnson, M. K. (1972). Contextual prerequisites for understanding: some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 11, 717-726.
- Brown, A. L., Bransford, J. D., Ferrara, R. A. and Campione, J. C. (1983). Learning, remembering, and understanding. In J. H. Flavell and E. M. Markman (eds), *Handbook of child psychology*, 4th edn, vol. 1. New York: Wiley.
- Bruce, D. (1985). The how and why of ecological memory. *Journal of Experimental Psychology: General*, 114, 78-90.
- Bugelski, B. R. (1971). *The psychology of learning applied to teaching*. New York: Bobbs-Merrill.
- Coccone, D. S. and Brelsford, J. W. (1974). Inter-presentation lag and rehearsal mode in recognition memory. *Journal of Experimental Psychology*, 103, 900-906.
- Cofer, C. N. (1967). The specificity of individual differences in learning. In R. M. Gagne (ed.), *Learning and individual differences*. Columbus, OH: Merrill.
- Cooper, E. H., & Pantle, A. J. (1967). The total-time hypothesis in verbal learning. *Psychological Bulletin*, 68, 221-234.
- Covington, M. V. and Omelich, C. L. (1988). Achievement dynamics: the interaction of motives, cognitions, and emotions over time. *Anxiety Research*, 1, 165-183.
- Craik, F. I. M. and 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. and Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 1, 268-294.
- Cronbach, L. J. and Snow, R. E. (1977). *Aptitudes and instructional methods*. New York: Irvington.
- Dellarosa, D. and Bourne, L. E., Jr (1980). Surface form and the spacing effect. *Memory and Cognition*, 13(6), 529-537.

- Dempster, F. N. (1987). Effects of variable encoding and spaced presentations on vocabulary learning. *Journal of Educational Psychology*, 79(2), 162-170.
- Dempster, F. N. (1988). The spacing effect: a case study in the failure to apply the results of psychological research. *American Psychologist*, 43, 627-634.
- Ebbinghaus, H. (1885). *Memory: A contribution to experimental psychology* (republished 1963). New York: Dover.
- Entwistle, N. (1981). *Styles of learning and teaching*. London: Wiley.
- Entwistle, N. (1987). A model of the teaching-learning process. In J. T. E. Richardson, M. W. Eysenck, and D. Warren Piper (eds), *Student learning: research in education and cognitive psychology*. Milton Keynes: Open University Press.
- Entwistle, N. J. and Ramsden, P. (1983). *Understanding student learning*. London: Croom Helm.
- Gagne, R. M. (1970). *Learning and individual differences*. Columbus, OH: Merrill.
- Glaser, R. (1976). Components of a psychology of instruction: toward a science of design. *Review of Educational Research*, 46(1), 1-24.
- Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory and Cognition*, 7(2), 95-112.
- Glenberg, A. M. & Lehmann, T. S. (1980). Spacing repetitions over 1 week. *Memory and Cognition*, 8(6), 528-538.
- Glover, J. A. & Corkill, A. J. (1987). Influence of paraphrased repetitions on the spacing effect. *Journal of Educational Psychology*, 79, 198-199.
- Goodenough, D. R. (1976). The role of individual differences in field dependence as a factor in learning and memory. *Psychological Bulletin*, 83, 765-694.
- Hintzman, D. L. (1974). Theoretical implications of the spacing effect. In R. L. Solso (ed.), *Theories in cognitive psychology: the Loyola symposium*. Hillsdale, NJ: Erlbaum.
- Hintzman, D. L. (1976). Repetition and memory. In G. H. Bower (ed.), *The psychology of learning and motivation*, vol. 10. New York: Plenum.
- Howe, M. J. A. and Ceci, S. J. (1979). Educational implications of memory research. In M. M. Gruneberg and P. E. Morris (eds), *Applied problems in memory*. London: Academic Press.
- Jenkins, James J. (1979). Four points to remember: a tetrahedral model of memory experiments. In L. S. Cermak and F. I. M. Craik (eds), *Levels of processing in human memory*. Hillsdale, NJ: Erlbaum.
- Johnston, W. A. and Uhl, C. N. (1976). The contributions of encoding effort and variability to the spacing effect on free recall. *Journal of Experimental Psychology: Human Learning and Memory*, 2(2), 161-171.
- Landauer, T. K. (1967). Interval between item repetitions and free recall memory. *Psychonomic Science*, 8, 439-440.
- Madigan, S. A. (1969). Intraserial repetition and coding processes in free recall. *Journal of Verbal Learning and Verbal Behavior*, 8, 828-835.
- Mannes, S. M. and Kintsch, W. (1987). Knowledge organization and text organization. *Cognition and Instruction*, 4(2), 91-115.
- Marton, F. (1975). What does it take to learn? In N. J. Entwistle (ed.), *Strategies for research and development in higher education*. Amsterdam: Swets & Zeitlinger.
- McKeachie, W. J. (1974). The decline and fall of the laws of learning. *Educational Researcher*, 3, 7-11.
- McKeachie, W. J. (1986). *Teaching tips*, 8th edn. Lexington, MA: D. C. Heath.
- Melton, A. W. (1967). Individual differences and theoretical process variables: general comments on the conference. In R. M. Gagne (ed.), *Learning and Individual Differences*. Columbus, OH: Merrill.
- Melton, A. W. (1970). The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior*, 9, 596-607.
- Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- Mook, D. G. (1983). In defense of external invalidity. *American Psychologist*, 38, 379-387.
- Morris, C. D., Bransford, J. D. and Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, 16, 519-533.

- Neisser, U. (1978). Memory: What are the important questions? In M. M. Gruneberg, P. E. Morris, and R. N. Sykes (eds), *Practical aspects of memory*. London: Academic Press.
- Pask, G. (1976). Styles and strategies of learning. *British Journal of Educational Psychology*, 46, 128-148.
- Pask, G. and Scott, B. C. E. (1972). Learning strategies and individual competence. *International Journal of Man-Machine Studies*, 4, 217-253.
- Pauk W. (1984). *How to study in college*, 3rd edn. Boston, MA: Houghton Mifflin.
- Perruchet, P. (1989). The effect of spaced practice on explicit and implicit memory. *British Journal of Psychology*, 80, 113-130.
- Postman, L. and Rau, L. (1957). Retention as a function of the method of measurement. *University of California Publications in Psychology*, 8, 217-270.
- Rea, C. P. and Modigliani, V. (1985). The effect of expanded versus massed practice on the retention of multiplication facts and spelling lists. *Human Learning: Journal of Practical Research and Applications*, 4, 11-18.
- Reynolds, J. H. and Glaser, R. (1964). Effects of repetition and spaced review upon retention of a complex learning task. *Journal of Educational Psychology*, 55(5), 297-308.
- Ross, B. H. and Landauer, T. K. (1978). Memory for at least one of two items: test and failure of several theories of spacing effects. *Journal of Verbal Learning and Verbal Behavior*, 17, 669-680.
- Rothkopf, E. Z. and Coke, E. U. (1963). Repetition interval and rehearsal method in learning equivalences from written sentences. *Journal of Verbal Learning and Verbal Behavior*, 2, 406-416.
- Shaughnessy, J., Zimmerman, J. and Underwood, B. (1974). The spacing effect in the learning of word pairs and the components of word pairs. *Memory and Cognition*, 2, 742-748.
- Smith, S. M. and Rothkopf, E. Z. (1984). Contextual enrichment and distribution of practice in the classroom. *Cognition and Instruction*, 1(3), 341-358.
- Smith, D. E. P., Wood, R. L., Downer, J. W. and Raygor, A. L. (1956). Reading improvement as a function of student personality and teaching method. *Journal of Educational Psychology*, 47, 47-59.
- Snow, R. E. (1989). Aptitude-treatment interaction as a framework for research on individual differences in learning. In P. L. Ackerman, R. J. Sternberg, and R. Glaser (eds), *Learning and individual differences*, New York: W. H. Freeman.
- Taylor, E. (1983). Orientation to study: a longitudinal interview investigation of students on two human studies degree courses at Surrey University. Unpublished Ph.D. thesis, University of Surrey.
- Thios, S. J. (1972). Memory for words in repeated sentences. *Journal of Verbal Learning and Verbal Behavior*, 11, 789-793.
- Tobias, S. (1985). Test anxiety: interference, defective skills, and cognitive capacity. *Educational Psychologist*, 20, 135-142.
- Tobias, S. (1988). Adapting instruction to student characterization. Presidential address of Division of Educational Psychology, American Psychological Association, August.
- Underwood, B. J. (1964). A breakdown of the total-time law in free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, 9, 573-580.
- Wickelgren, W. A. (1974). Single-trace fragility theory of memory dynamics. *Memory and Cognition*, 2, 775-780.
- Witkin, Oltman, Raskin, & Karp (1971). *A manual for the embedded figures test*. Palo Alto, CA: Consulting Psychologists Press.
- Yerkes, R. M. and Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurological Psychology*, 18, 459-482.