RETROACTIVE INTERFERENCE IN MEANINGFUL VERBAL LEARNING

GEORGE T. MILONAS

International Christian University

Eighteen Japanese freshmen students were assigned to two experimental and one control group to test Ausubel's (1963) subsumption theory of meaningful verbal learning and retention in a retroactive interference paradigm. Between the original learning task and the test for retention, one experimental group learned a similar but conflicting passage while the other experimental group learned a completely dissimilar passage. The control group did not participate in any interpolated learning task. Contrary to the findings of Anderson and Myrow (1971), whose material and modified methods were used, no retroactive interference occurred. The results of this study support Ausubel's subsumption theory of meaningful verbal learning and retention.

Ever since the first systematic study of retroactive interference (inhibition) by Mueller and Pilzecker in Germany in 1900 (Hilgard & Bower, 1966), the entire topic of interference in learning has been recognized as vitally important because of its relation to retention and forgetting.

Retroactive interference is the decrement in retention attributable to interpolated learning (McGeoch & Irion, 1952), and the operations that define it require a comparison of the retention of some original learning between two groups which differ in some aspect of interpolated activity (Underwood, 1949). Better retention by members of the control group over members of the experimental group defines retroactive interference.

A review of the experimental literature revealed that the majority of studies dealing with retroactive interference utilized rote learning and verbatim recall of such items as nonsense syllables and paired associates. Almost all of the studies examined reported some degree of retroactive interference. In the few studies using meaningful material (as opposed to nonsense syllables and paired associates) (McGeoch & McKinney, 1934; Deese & Hardman, 1954; Hall, 1955; Ausubel, Robbins, & Blake, 1957; Ausubel, Stager, & Gaite, 1968, 1969) retroactive interference was not found.

Rote learning is the learning of discrete and relatively isolated entities which are relatable to cognitive structure in only an arbitrary, verbatim fashion (Ausubel, 1963). This type of learning is generally referred to as verbal learning in the tradition of the British associationists who held that the basic verbal unit was an association of ideas. A wide range of verbal units have been used in the study of verbal learning. American investigators have made extensive use of nonsense syllables, but digits, single words, word couplets, and prose have also been used. Regardless of the materials used, the learning was rote and recall was verbatim.
Anderson and Myrow (1971), speaking of rote tasks, noted:

... lists of nonsense syllables and discrete words have long comprised the chief experimental tasks of psychologists interested in memory. While research with rote tasks is perhaps a self-perpetuating tradition, most experimental psychologists undoubtedly operate on the faith that their research is generally applicable to memory on verbal-symbolic material. Whether this faith is well placed is open to serious question on the basis of data now available (p. 81).

Both Ausubel, et al., (1957) and Sla-mecka (1959) agree that nonsense syllables are used because (a) meaningful antecedent conditions, which vary greatly from one individual to another, are thereby controlled, and (b) nonsense syllables, being additive units of equal difficulty, can be readily quantified. However, Ausubel's (1963) position is that learning and remembering potentially meaningful verbal material entails entirely different processes than learning and remembering rote tasks. The very nature of nonsense syllables, the fact that they are discrete and relatively isolated entities which are relatable to cognitive structure in arbitrary, verbatim fashion, limits the applicability of research findings which uses such syllables. Findings in such research are applicable only to retention of similar discrete, isolated entities; that is, to other rote learning tasks—a type of learning that, although not completely absent in the classroom, comprises a very small percentage of classroom learning. Rote learning findings cannot and should not be extrapolated from the laboratory situation into the classroom, an extrapolation which has been "commonly and uncritically" made (Ausubel, et al., 1957).

Ausubel's subsumption theory of meaningful verbal learning and retention, by contrast, addresses itself to the type of learning that occupies the vast majority of a student's time in the classroom. The scope of the theory is limited to meaningful verbal learning by human beings. Subsumption theory does not embrace other types of learning such as motor skills, rote learning, perceptual learning, concept formation, thinking, or problem solving; nor is it concerned with animal learning. Rather, it is a theory of learning that deals with how individuals comprehend, learn, organize and remember the large volume of meaningful verbal learning presented to them in the classroom; what Ausubel calls "reception" learning. It should be evident that the subsumption theory covers the major cognitive learning activities that take place in the school environment.

Meaningful verbal material, when learned, is related to existing concepts in cognitive structure in such a manner as to make it possible to understand various kinds of significant relationships such as derivative, correlative, or qualifying. Most of the material a student encounters in the classroom is relatable to a previously learned background of meaningful ideas and information. In fact, curriculum is deliberately organized in this fashion to facilitate new learning. Embedding the new material within the hierarchy of cognitive structure protects it from the interfering effects of proactively or retroactively introduced material. Roteley learned material, by contrast, is made up of discrete and relatively isolated entities which are only relatable to cognitive structure in an arbitrary, verbatim fashion which precludes the establishment of the above-mentioned relationships. No protective anchorage in cognitive structure is provided for rote material.

Because it does not interact with the existing cognitive structure in a substantive, organic fashion, roteley learned material is much more vulnerable to forgetting unless greatly overlearned. The human mind is not efficiently designed for long-term verbatim storage of arbitrary associations and hence, the retention span for rote
learning is relatively brief (Ausubel, 1968, p. 109). Inasmuch as rote learned material is essentially isolated from cognitive structure, it is influenced primarily by the interfering effects of similar rote material learned immediately before and after the learning task. In this light, it is not unreasonable to explain the learning and retention of discrete rote units in such stimulus-response terms as intra- and inter-task similarity, response competition, associative strength, and stimulus or response generalization. The learning and retention of meaningful material, however, is primarily influenced by the attributes of relevant subsuming concepts in cognitive structure with which they interact. The introduction of similar material immediately before or after the learning task has relatively little influence in the retention of meaningful material.

What is meant by “meaningful” verbal learning? The neobehaviorist’s definition of “meaning” is an associationistic-based definition whereas the cognitive psychologist’s definition is a cognitive process-based one. As a consequence, they are not talking about the same thing.

Ebbinghouse (1913) assumed that the nonsense syllables he invented were relatively homogeneous in eliciting associations. Glaze’s (1928) work disproved this assumption. He found that different syllables were capable of evoking different numbers of associations and thus the neobehaviorist’s position regarding meaning was established. Subsequently, Hull (1933) and Krueger (1934) confirmed Glaze’s findings.

Noble’s (1952) original study used two-syllable nouns obtained from the Thorndike-Lorge (1944) word frequency count. The m index has since been extended to 100 nonsense syllables (Mandler, 1955), to 329 nouns (Spreen & Schulz, 1966), and to a variety of words when using children rather than adults as subjects (Shapiro, 1964; Michelson, 1969).

Thus, meaningfulness, as evolved by investigators of rote learning, is typically defined in terms of the percentage of subjects who are able to produce an association to a given nonsense syllable, noun, or other word, within a limited period of time, or in terms of the total number of associations evoked by a particular verbal stimulus.

From the cognitive viewpoint, meaning is a “phenomenological outcome of a meaningful learning process in which the potential meaning inherent in the external world becomes converted into an individualized psychological state or content of consciousness” (Ausubel, 1963, p. 35). Meaning always implies some sort or form of representational equivalence between language (or symbols) and mental content. For learning to be meaningful, the material to be learned must be potentially meaningful; it must be relatable in a substantive and nonarbitrary fashion to ideational systems already existing in the learner’s cognitive structure.

What does Ausubel mean by “substantive and nonarbitrary;” the two criteria for learning to be meaningful? The first criterion, nonarbitrary relatability, simply means that if the material itself exhibits sufficient nonarbitrariness (or nonrandomness), a sufficient and almost self-evident
basis exists for relating it in a nonarbitrary fashion to the kinds of correspondingly relevant ideas that human beings are capable of learning. Meaningful learning could thus be nonarbitrarily relatable to specifically relevant ideas in the form of examples, derivatives, special cases, elaborations, etc., of those specifically relevant ideas. Or it could be relatable to a wider array of relevant ideas in the sense of being generally congruent with them (Ausubel, 1968).

The second criterion, substantive relatability, implies that if the learning material is sufficiently nonarbitrary, an ideational equivalent symbol or group of symbols could be related to cognitive structure without any change in meaning. In other words, neither meaningful learning nor emergent meaning are dependent on the exclusive use of particular signs and no others. The same concept or proposition can be expressed in synonymous language and would convey precisely the same meaning (Ausubel, 1968).

It is abundantly clear that nonsense syllables and paired associates do not meet the criteria of meaningfulness discussed above; they are discrete and relatively isolated entities insofar as cognitive structure is concerned.

One further word regarding potentially meaningful material before leaving the subject of meaningfulness. Irrespective of how much potential meaning may inhere in a particular proposition, if the learner's intention (set) is to memorize it arbitrarily and verbatim (as a series of arbitrarily learned words), both the learning process and the learning outcome must be rote or meaningless. And, conversely, no matter how meaningful the learner's set may be, neither the process nor the outcome of learning can possibly be meaningful if the learning task is not potentially meaningful; if it is not substantively and nonarbitrarily relatable to his particular cognitive structure. Furthermore, the actual meaning derived from potentially meaningful material is the idiosyncratic, phenomenological meaning which the learner derives based on his own particular cognitive structure.

Anderson and Myrow (1971) analyzed the theoretical and methodological problems that may have prevented previous research from detecting retroactive interference with meaningful discourse. Contrary to the findings stated above, Anderson and Myrow did demonstrate a significant degree of retroactive interference while using connected discourse. Their research is the first clear-cut demonstration of retroactive interference in a meaningful verbal learning paradigm of which this experimenter has any knowledge. Other retroaction studies have been made using connected discourse but close scrutiny of these experiments revealed that retroactive interference was found only in those studies which required verbatim recall of the potentially meaningful material in the test for retention (Jenkins & Sparks, 1940; Jensen, Dibble, & Anderson, 1971; King & Cofer, 1960; Slamecka, 1959, 1960, 1962, to mention a few). Under such circumstances, the learning that took place was rote learning as distinguished from meaningful verbal learning.

The current study was conducted from a cognitive viewpoint and was based on Ausubel's (1963) subsumption theory of meaningful verbal learning and retention. Subsumption theory states that the existing cognitive structure of the learner is crucial in learning. This structure is hierarchically organized with the more inclusive concepts at the apex and with lesser inclusive concepts and information at the lower levels, all of which are linked to the next higher step by a process of subsumption. Meaningful verbal learning takes place when potentially meaningful material is perceived, interacts with, and is appropriately subsumed under a relevant and more inclusive conceptual system of the learner's cognitive structure. The discriminability of the new material from the established cognitive structure that
subsumes it is the second important factor in meaningful verbal learning and retention.

Subsumption theory predicts that when new learning is nonarbitrarily and substantively related to existing ideas in cognitive structure, the newly learned material is protected by such anchorage from the automatic interfering effects of retroactively introduced material. Whether or not interference occurs depends on the net interfering effects of such material on the stability, clarity and discriminability of the learning task.

Hypotheses

Based on subsumption theory, four hypotheses were derived and were tested in the current study. The hypotheses are divided among three areas; type of interpolated learning, response mode, and item type. The latter, item type, is made up of three related or substantially similar corollaries.

Type of Interpolated Learning

$H_1$ A statistically significant degree of retroactive interference will not be evident when the original learning task and the test for retention is interpolated by a highly similar and conflicting learning task.

$H_2$ A statistically significant degree of retroactive interference will not be observed when a substantively dissimilar learning task is introduced between the original learning and the test for retention.

Response Mode

$H_3$ The test for retention will yield higher scores when a recognition response mode is utilized rather than a recall response mode.

Item Type

$H_{4a}$ Facilitation type items on the test for retention will result in an increment in retention scores.

$H_{4b}$ Neutral items on the test for retention will result in neither an increment nor a decrement in retention scores.

$H_{4c}$ Interference items on the test for retention will not result in a statistically significant degree of retroactive interference.

Method

Subjects

The experiment sample consisted of 18 freshmen students of International Christian University (7 males and 11 females) who were paid volunteers drawn from the total membership of 308 freshmen who entered the university in April 1972.

Learning Material

The learning materials were the same three passages utilized by Anderson and Myrow (1971) in conducting their Experiment I (two experiments were reported in their study). The three passages were translated into Japanese and are described below.

The original learning (OL) passage was an anthropological discussion of a fictitious tribe in Africa, the Himoots. Being fictitious, it was totally unfamiliar to the Ss; being anthropological in nature, it was reasonable to assume that the cognitive structure of the Ss contained relevant and more inclusive concepts under which the learning passage could be subsumed. The passage was potentially meaningful in the sense in which the term is used by Ausubel (1963).

The related interpolated learning (RIL) passage was a parallel anthropological discussion of another fictitious African tribe called the Gruanda. An extremely high order of similarity, and therefore confusability, was incorporated into the Himoot and Gruanda passages by Anderson and Myrow. The previous OL comments concerning subsumability, meaningfulness and unfamiliarity apply equally to the RIL passage.

The unrelated interpolated learning (UIL) passage was the drug addiction article used by Ausubel, et al., (1968) and subsequently by Anderson and Myrow (1971) in their Experi-
The UIL passage dealt with the causes and types of drug addiction, a subject which had proved unfamiliar to the Ss in the Ausubel as well as the Anderson and Myrow study. The passage was potentially meaningful and anchoring concepts such as drugs, medicine, addiction and related relevant concepts were presumed to be an integral part of the cognitive structure of the Ss.

Tests

The criterion tests constructed for the experiment utilized Anderson and Myrow’s original 30 multiple-choice items. However, each of the tests constructed for the current study contained not 30 multiple-choice (MC) items, but 15 multiple-choice plus 15 short answer (SA) items; a total of 30 items on each test. Form A of the test was used as a pretest (pre-interpolated learning) and Form B as a posttest. The difference in the tests was that those items which appeared as MC items on Form A were shown on Form B as SA items, and those items which appeared as SA items on Form A appeared as MC items on Form B. In listing the items on both Form A and Form B, the first 15 items appear as MC items and the last 15 items as SA items. With this arrangement, each S answered each question once in the MC mode and once in the SA mode.

There were three types of test items: (1) Facilitation items were items whose stems contained information found in both the OL and the RIL passage. These items could be answered in the same way on the basis of both passages. For example, one question asks, “In what part of Africa do the Himoots live?” They live in the east central part of Africa which is also where the Gruanda tribe of the RIL passage lives. (2) Neutral items were items entailing material which was not discussed in the RIL passage. (3) Interference items, like the facilitation items, had a stem which was represented in both the OL and the RIL passages, however the RIL passage would indicate an incorrect response. For example, one question asks, “What must a Himootian male do before he is permitted to marry?”

Both passages discuss what a male member of the tribe must do before he is permitted to marry. In the case of the Himoot, he must demonstrate satisfactory progress with his work whereas the Gruanda youth in the RIL passage must complete a feat of daring, such as stalking a wild boar.

Design

The design of the experiment was a $3 \times 2 \times 3$ factorial design which included three main factors: (1) Type of interpolated learning identified by group (RIL by experimental group I, UIL by experimental group II and no interpolated learning by the control group); (2) response modes (recognition (MC) and recall (SA)); and (3) item types (facilitation, neutral and interference). The latter two factors are within subject factors; i.e., all Ss answered both recognition and recall questions and all three item types.

Procedure

On Day 1 Ss were randomly assigned to one of the three groups and all learned the OL passage concerning the Himoots. At the end of 30 min, the OL passage was collected and the Ss were administered the pretest (Form A). On Day 3 only the two experimental groups were present. The RIL group learned the similar and conflicting Gruanda passage while the UIL group studied the drug addiction passage. On Day 5 all Ss were administered the posttest (Form B) for retention of the original learning passage.

Results

Raw and corrected score means (corrected for guessing using the formula $R - W/3$) are shown in Table 1, as well as the percentage of original learning retained. The standard deviations were 2.20 and 2.48 for the pretest and posttest respectively. F tests for homogeneity of variance between the groups indicated that they were comparable; i.e., drawn from the same population.

The high level of original learning, coupled with the extremely high retention
Retroactive Interference in Meaningful Verbal Learning

TABLE 1
Original learning pretest and posttest means and percentage retained

<table>
<thead>
<tr>
<th>Group</th>
<th>Interpolated material</th>
<th>Pretest mean</th>
<th>Posttest mean</th>
<th>Percent retained†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Raw mean</td>
<td>Corr. mean</td>
<td>Raw mean</td>
</tr>
<tr>
<td>Ex Gp I</td>
<td>Related</td>
<td>22.3</td>
<td>21.8</td>
<td>22.2</td>
</tr>
<tr>
<td>Ex Gp II</td>
<td>Unrelated</td>
<td>25.3</td>
<td>25.1</td>
<td>24.7</td>
</tr>
<tr>
<td>Control Gp</td>
<td>None</td>
<td>23.3</td>
<td>22.9</td>
<td>23.3</td>
</tr>
</tbody>
</table>

† Percent retained = Posttest Mean/Pretest Mean × 100

TABLE 2
Summary of analysis of variance

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of Interpolated Learning</td>
<td>3.12</td>
<td>2</td>
<td>1.56</td>
<td>2.14</td>
</tr>
<tr>
<td>B (Response Mode)</td>
<td>5.76</td>
<td>1</td>
<td>5.79</td>
<td>7.93*</td>
</tr>
<tr>
<td>C (Item Type)</td>
<td>7.46</td>
<td>2</td>
<td>3.73</td>
<td>5.11*</td>
</tr>
<tr>
<td>A × B</td>
<td>0.13</td>
<td>2</td>
<td>0.07</td>
<td>0.01</td>
</tr>
<tr>
<td>A × C</td>
<td>4.70</td>
<td>4</td>
<td>1.18</td>
<td>1.62</td>
</tr>
<tr>
<td>B × C</td>
<td>7.35</td>
<td>2</td>
<td>3.68</td>
<td>5.04*</td>
</tr>
<tr>
<td>A × B × C</td>
<td>2.88</td>
<td>4</td>
<td>0.72</td>
<td>0.99</td>
</tr>
<tr>
<td>Within Treatments</td>
<td>65.45</td>
<td>90</td>
<td>0.73</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>96.88</td>
<td>107</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < .01

TABLE 3
Summary of analysis of covariance

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatments</td>
<td>3.66</td>
<td>2</td>
<td>1.83</td>
<td>0.362</td>
</tr>
<tr>
<td>Error</td>
<td>70.72</td>
<td>14</td>
<td>5.05</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>74.38</td>
<td>16</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

rate led to the decision to perform the analysis of variance using the uncorrected scores. Table 2 contains the analysis of variance summary. All analyses and computations were according to Edwards (1960).

Type of Interpolated Learning

The analysis of the posttest data indicated that retention by the three groups did not differ significantly. To further substantiate that the treatment each group received did not affect retention significantly, an analysis of covariance was also made. The results of this analysis are shown in Table 3. Again, an F value which was not significant was obtained indicating that the dependent variable, retention, was not significantly affected regardless of the type of interpolated activity.

There was no significant degree of retroactive interference regardless of whether a similar or a dissimilar learning task was
interpolated between OL and the posttest for retention. The findings of the analysis of variance and covariance confirmed Hypotheses H₁ and H₂.

Response Mode

The data comparing the two response modes indicated that the performance was much higher on the recognition (MC) portion than on the recall (SA) portion of the test. Figure 1 is a graph of the response mode effect.

The recognition mode yielded significantly superior retention scores compared to the recall mode thereby confirming Hypothesis H₃.

Item Type

The analysis of variance resulted in an item type effect which was significant. Although the finding was significant, it failed to prove or disprove Hypothesis 4, pointing out a deficiency in the design of the experiment. This inconclusiveness led to an analysis of the differences for paired observations. The pretest and posttest scores were dependent means and a comparison of these means would indicate the increment or decrement in retention that each group experienced on each of the three types of items. A significant decrement would indicate a case of retroactive interference.

The nine analyses which were required to make the comparisons all resulted in F values which did not exceed the tabled values of F at the 5 percent level. There was no significant difference between pretest and posttest scores for any of the three groups in any of the three item types. This finding confirmed Hypothesis H₄a and H₄c. Hypothesis H₄b was not confirmed.

Response Mode and Item Type Interaction

The only significant interaction effect revealed by the analysis of variance was between response mode and item type. Figure 2 graphically illustrates the interaction effect.

On the recognition portion of the posttest, the data indicated that the mean percentage of correctly answered items was highest for interference items and lowest for neutral items. However, on the recall portion of the test, Ss scored highest on the facilitation items and lowest on the interference items.

DISCUSSION

Anderson and Myrow (1971) reasoned that previous research using verbal material posed several difficulties which may have prevented detection of the retroactive interference phenomenon. Failure to consider in detail the similarities and differences between the OL and interpolated...
learning passages was considered the biggest problem. They refer to Osgood's (1949) article and note that "... it has been recognized that retroactive inhibition occurs with paired associates when the stimuli in successive lists are identical or very similar and the responses are antagonistic. When both stimuli and response are very similar, there is retroactive facilitation. Little effect is observed when both the stimuli and responses in successive lists are different." (p. 82)

With this problem in mind, they designed a study to test for retroactive interference which would parallel the research with paired associate lists. Such research, they noted, "... would (a) used a recall measure instead of, or in addition to, a recognition measure and (b) employ test items which would be answered differently on the basis of the first passage than on the basis of the second." (p. 83)

Acknowledging the superiority of the Anderson and Myrow learning passages and tests, the same material translated into Japanese, was used for the current study. Replication of the entire Anderson and Myrow study was not possible due, primarily to the population sample. Anderson and Myrow used 145 Ss compared to 18 for the current study. Also, rather than having the Ss answer both a 30-item test in the MC mode and then in the SA mode, a 30-item test was devised in which 15 items were in the MC mode and 15 in the SA mode. In this manner only 2 tests were administered; Form A of the test immediately after OL and Form B as the posttest. The sample size also precluded using two control groups as did Anderson and Myrow. In addition, the Anderson and Myrow study entailed 8 days between OL and the test for retention whereas the current study, due to circumstances over which there was no control, involved only 5 days. The Ss in the current study were paid volunteers and were college freshmen while Anderson and Myrow used unspecified grade high school students. It was not indicated whether they were volunteers or were simply assigned by the high school authorities as Ss in the study.

The analyses of variance and covariance made very clear that in this learning situation, regardless of whether the interpolated learning was (a) similar and conflicting or (b) dissimilar to the OL material, retroactive interference did not occur. This finding denies the applicability of interference theory and rote learning theory in meaningful verbal learning situations.

On the other hand, the subsumption theory is supported by the findings of the current study. The learning material was not discrete and relatively isolated entities which make up nonsense syllables or paired associates; it was the type of learning material that occupies the vast majority of a student's time in the classroom, i.e., meaningful verbal material. As such, it was potentially meaningful and found anchorage in the cognitive structure of the Ss. Subsumption theory predicts that regardless of the similarity or dissimilarity of the learning task to previous learning, if the material is potentially meaningful, and if the cognitive structure of the learner contains the relevant concepts and ideational systems under which the new material can be subsumed, there would be no retroactive interference.

Osgood (1953) points out that a major variable in determining the amount of forgetting is the similarity of responses (which Anderson and Myrow so carefully designed into their materials) in original and interpolated learning. He theorizes that increased similarity, short of identity (which is akin to practice or relearning) leads to increased interference. However, with the exception of the Anderson and Myrow (1971) study, this relationship has been experimentally verified only in studies of rote learning.

It can be inferred from the extremely high retention rates (Table 1) that the Ss cognitive structure contained the necessary
relevant conceptual systems under which the Himoot passage was subsumed, and the new material's anchorage to those relevant concepts was apparently very firm. The dissociability strength of the OL passage, over the 4 days between OL and the posttest, remained substantially above the threshold of availability although the threshold level differed for recognition and recall items.

Anderson and Myrow (1971), in discussing studies of memory using connected discourse stated, “No retroactive inhibition, or retroactive facilitation, has appeared only when a multiple-choice test was used.” (p. 83) The results of the current study contradict that statement; both MC and SA questions were used in the test for retention. The data indicates that the Ss performed better on the recognition portion (MC) of the test than on the recall (SA), however that was anticipated and predicted. As Ausubel (1963) has pointed out, holding dissociability strength constant, the threshold of availability is higher for recall than for recognition. The two methods of measuring retention make different demands on the availability of the newly learned material. In the case of recognition, the Ss had only to identify the correct answer whereas for recall, they had to reconstruct the stimulus situation autonomously.

Insofar as type of items are concerned, the stem of the test items were considered by Anderson and Myrow (1971) as the stimulus as used in rote learning theory. The test items, then, could be categorized according to interference theory. Facilitation items were items in which the stimulus (the stem of the posttest item) appeared in both the OL and the RIL passages in substantially the same form, and the responses were identical in both passages. Osgood’s (1949) second empirical law would predict that with identical stimuli and identical responses, there would be no interference and thus neither increment nor decrement in retention. This investigator reasoned that, with the item stem and the correct response appearing in both the OL and RIL passages, the group that was exposed to both of these passages (the RIL group) would, in effect, receive the benefit of a second learning of such items. The rehearsal, it was reasoned, would facilitate retention and result in an increment in retention scores. The data contradicted this reasoning. Table 4 compares the percent correct responses by item type and group.

The initial analysis of variance showed that a mean of 82.2 percent of the facilitation items were answered correctly—a higher mean than for the other two types of items—but it did not indicate whether facilitation resulted. A second analysis revealed that there was no significant difference between pretest and posttest scores for facilitation items. This latter finding meant that the hypothesis concerning facilitation items was not confirmed and, furthermore, it seemed to point toward support of Osgood’s (1949) second empirical law.

Neutral items entailed material that was not discussed in the RIL passage. Neither increment nor decrement in retention scores were found for neutral items. Both interference theory and subsumption theory would predict neither increment nor decrement for neutral items.

In the case of interference items, the stimulus (the stem of the posttest item) was represented in both the OL and the RIL passages; however the RIL passage indicated an incorrect response. This type of item represents the opposite of the
Retroactive Interference in Meaningful Verbal Learning

facilitation item, and again, Osgood's (1949) second empirical law was being tested. With identical stimuli and a competing response, maximum interference would be predicted by Osgood's law.

The results of the analysis indicated that there was no difference of significance between the pretest and posttest scores for interference items. To support interference theory there should have been a significant decrement in the posttest scores resulting in retroactive interference. Such was not the case.

The findings for interference items clearly do not conform to interference theory or to Osgood's second empirical law. A law, to be valid, must apply to all cases defined by that law, without exception. The second empirical law promulgated by Osgood encompasses the complete range of responses from similar (short of identity) through neutral to opposite, and the law limits the stimuli to those which are "functionally identical." With such boundaries, the law should apply to both facilitation and interference items as used in the current study. The fact that the law seemed to apply in the case of facilitation items but does not apply to interference items rules the law invalid in this case—the case of meaningful verbal learning. This does not mean to imply that it is likewise invalid in all cases; it may well be valid for rote learning. However, in this regard, Bugelski and Cadwallader (1956) designed an experiment to test all of the key points on Osgood's (1949) transfer and retroaction surface. Their findings confirmed Osgood's first and third empirical laws. These laws deal with the stimulus continuum. The second law, dealing with the response continuum was not confirmed and "... appears to be in serious error." (p. 364)

Osgood's surface shows a gradually increasing negative effect from the stage of similarity, through neutral, to opposed responses. The Bugelski and Cadwallader findings, on the other hand, has the most significant departure from the control base line occurring with similar responses. The similar responses appeared to cause the greatest degree of retroaction (negative transfer) and the effect decreased with less similarity instead of increasing as Osgood would have it do. Actually, the curves Bugelski and Cadwallader obtained appear to conform to the Skaggs-Robinson hypothesis (McGeoch & Irion, 1952) more closely than to the Osgood surface. Learning opposed responses did have a negative effect but it was of the least degree and there appeared to be no distinction among the stimulus relations involving "opposed" responses. About the same degree of effect was found with identical, similar and neutral stimuli. They also found that neutral stimuli in interpolated paired associates, in general, had no significant effect on recall of earlier learning.

At least two other experiments have attempted to replicate Osgood's transfer surface. These studies, too, have not provided consistent findings. A study by Dallet (1962) is in essential agreement with Osgood's surface; the experiments of Wimer (1964) are not. Wimer's findings indicate "... that only relatively small amounts of transfer could be produced by the very numerous and sometimes extreme combinations of interlist similarity relations used." (p. 278)

The ambiguity of the Osgood transfer and retroaction surface which has resulted from the findings of Bugelski and Cadwallader (1956) and Wimer (1964) makes suspect the findings of Anderson and Myrow (1971) who have attempted to extrapolate rote learning findings to a meaningful verbal learning situation in the experiments they reported in the Monograph.

A significant response mode X item type interaction was revealed by the analysis of variance. Figure 2 shows that the Ss found the interference items least difficult when the MA response mode was
used (92.2% correct) and, contrariwise, most difficult when the SA mode was used (70.0% correct).

Although facilitation items show a similar effect (86.0 and 77.6 percent), neutral items showed a small increment in percent responses correct between the recognition and recall modes (68.9 and 72.2 percent respectively).

The low percent correct responses for the recall mode is attributed to the lower threshold of availability for recall as compared to recognition. In addition, verbal ability of the Ss is a factor. A verbal ability test was not administered in this study but Anderson and Myrow (1971) found that verbal ability was significantly related to recall ($F=12.19$, $df=3/84$, $p<.01$).

**Conclusions**

The current study to test Ausubel's (1963) subsumption theory of meaningful verbal learning and retention as it applies to retroactive interference confirms that theory and denies the applicability of interference theory insofar as meaningful verbal learning is concerned. Results of this study do not conform to the findings of Anderson and Myrow (1971) although identical learning materials and similar tests were utilized.

It should be pointed out that a direct comparison of the two studies is not possible for the following reasons:

1. The size of the population sample; i.e., 145 for Anderson and Myrow vs. 18 for the current study. Although, based on the estimates of population parameters, the procedures used in this experiment were correct for the small sample tested, it would be desirable to use a larger sample in order to detect more significant effects if such effects exist.

2. The experimental designs of the two studies were not similar due mainly to the size of the population sample.

3. Tests and testing procedures differed although the same basic test was utilized.

4. This study used paid volunteers, which may have affected motivation, whereas Anderson and Myrow did not pay their Ss.

5. The current study utilized university freshmen while Anderson and Myrow used unspecified grade high school students. The maturity of the cognitive structures of the Ss may have differed considerably.

6. The retention span from OL to posttest differed from 8 days in the Anderson and Myrow study to 5 days for the current study.

7. The English version of the OL and RIL passages were designed to control word association values, and a readability analysis was made of them as well as of the UIL passage. Whether that control was incorporated into the translated passages is unknown. Although a readability analysis was not made for the Japanese versions of the passages, it was assumed that the uncorrected reading levels (i.e., American school grade level of reading ability) were at least as high as the levels reported by Anderson and Myrow; grade 6.9 for the OL passage, grade 6.6 for the RIL passage and grade 9.1 for the UIL passage.

8. A verbal ability test was not administered in this study although it was known that verbal ability is significantly related to recall.

The disparate findings of this and the Anderson and Myrow (1971) studies would indicate further research is needed to resolve the question of retroactive interference in meaningful verbal learning.

**References**


(Received March 11, 1976)