Literature on forgetting in prose learning was reviewed to establish the relative importance of five variables: (1) the definition of prose and the organizational structure of the material implied in the definition; (2) the time factor; (3) the mode of criteria measurement; (4) the similarity factor; and (5) strength of learning. Each of these was considered in turn and its relative contribution assessed. It was found that (1) all kinds of stimulus materials have been used in studying interference effects in prose learning—the most common type being constructed passages with identifiable elements that would be switched for interpolated learning and later directly tested for retention; (2) similarity seemed to be a determining variable as in paired-associate learning; (3) timing seemed to be an important factor; (4) the mode of testing influenced quantitatively the measured degree of learning and recall; and (5) strength of original and interpolated learning appeared to be a significant factor. A bibliography is included. (AW)
In the past few years there has been a great effort made to extend the interference theory of forgetting from the classical verbal learning laboratory into the area of prose learning. However, a clear-cut demonstration of the mechanism of interference in prose learning has proved to be more elusive than had been expected.

The Ausubelian challenge to researchers to perform educational research at the level of complexity that exists in the normal classroom was half-heartedly accepted and most prose learning researchers began to apply the basic retroactive design that had been proved successful in verbal learning laboratories, in attempts to tease out interference effects in prose learning. Numerous studies were carried out in the search for the evidence that was to conclusively prove the existence of interference in prose learning. The resulting cacophony of findings was predictable.

Unfortunately, prose learning is not paired-associate learning, and variables that played only a minor part and could easily be controlled for in the verbal learning laboratory were not so easily dealt with. For example: (1) What precisely is prose? Every type of written verbal material from poetry to approximations to English has been used; (2) How is similarity between original learning and interpolated learning defined?; (3) What effect does the strength of the learning have? Both the strength of the original learning (OL) and interpolated learning (IL) is fairly easy to control with
paired-associate studies, but with prose learning, strength of learning becomes contaminated by other variables such as the length, type, organization, and difficulty level of the material; (4) As far back as 1948, Underwood demonstrated that the temporal relationship between OL, IL, and Test of OL was a very important variable in determining the relative effects of retroactive and proactive interference; (5) Ebbinghaus and his followers knew what they were doing when they tried to isolate the effects of meaning from their studies. Unfortunately for prose learning investigators, it is meaning, and the structure of its organization that makes prose prose. Both variables have to be taken into consideration; (6) Which is more important to forgetting of prose, proactive or retroactive interference? Most verbal learning investigators lean toward proactive interference as lending the most long-lasting effects. Nevertheless, in prose learning, most studies thus far have attempted to look only at retroactive interference effects; (7) A final major contributing variable is the mode of criterion test measurement. There has been shown to be a difference in the sensitivity of recall and recognition test measures (Postman, 1952) and there is even some question as to whether or not the two kinds of tests measure the same thing.

It was the intent of this paper to review the literature on forgetting in prose learning, attempting to establish the relative importance of five of the foregoing variables: (1) the definition of prose and the organizational structure of the material implied in the definition; (2) the kind of similarity being dealt with in each study; (3) the time factors involved in each study; (4) the mode of criterion measurement; (5) the strength of learning. Each will be considered in turn and its relative contribution assessed.

**Type of Prose Materials Used**

All kinds of stimulus materials have been used in the study of interference
effects in prose learning: single sentences presented work-for-word in list fashion via memory drum; sentences presented on flash cards; sentences reproduced on slides and projected on a screen; textbook-type construction and presentation in mimeographed pamphlets; excerpts from novels; highly structured specially constructed paragraphs. The most common type was specially constructed passages with identifiable elements that could be switched for interpolated learning and later directly tested for retention.

The results of the studies reviewed seem to beg the obvious. If you want to find RI, you are best off using materials that are highly structured and have in them easily definable stimulus and response components. Investigators using unmodified materials directly extracted from books, rarely detected RI effects. There are some exceptions, but most of these seem to have other variables tied in the results. For example: Jensen and Anderson (1970) used extracts from books and still found RI, but attributed this to the fact that the material was extremely difficult and thus not well learned; Mills and Kessel (1965) found RI with short passages from novelist D. H. Lawrence but they did it using list presentation on a memory drum. Memory drum presentation seems generally to be an effective means of inducing enough control into the situation to produce RI effects. (Slamecka, 1960, 1962; Mills, et al, 1965, 1969). It might be argued though, that list presentation word-for-word probably inhibits the effect of integral aspects of prose such as organizational structure and inter-associations between elements.

Similarity Between Original Learning and Interpolated Learning

The similarity dimension is an easy one to manipulate with nonsense syllables supposedly devoid of meaning --- one merely changes some of the letters. When a similar procedure was carried over to prose learning by duplicating only some of the sentences from OL to IL, Hall (1955) found no decrement in scores
Wallach (1958) discusses the problem of similarity with prose materials and points out that it can be broken down into two areas: potential and psychological. Psychological similarity is ultimately subjective to the organism being studied. Potential similarity is more a matter of experimenter apriori judgment on some rational or other grounds. Three investigators tried to come up with some way of getting at the psychological similarity inherent in the materials they were using. Paveman (1971) used a scaling procedure and Slamecka (1962) and Mills and Kessel (1965) had other Ss judge the similarity inherent in the passages. All of these investigators found statistically significant RI results.

Potential similarity is often referred to as "topical similarity," "confusably similar" or "potentially conflicting." These names usually imply some type of apriori judgment on the part of the investigator. The problem with apriori judgment is that similarity ultimately is a subjective judgment on the part of the subjects undergoing the treatments. It seems reasonable then to expect that sometimes the investigator will be in tune with his subjects and sometimes he won't. The results seem to reflect this. For example, all of the Ausubel, et al, studies defining similarity in his way come up with conflicting results --- sometimes finding facilitation, but more often finding no statistically significant differences. Gilman (1970) called similarity the fact that the material was all about some form of medical illness (anemia, hemorrhage, shock) and found RI, as did Jensen and Anderson (1970), who used especially hard subject matter. On the other hand, when the "general theme was the same," (McGeoch and McKinny, 1934) and when "topical similarity" was employed (Schuell and Hapkievicz, 1969), RI failed to appear.

One popular mode of determining similarity was for the investigator to
identify the S and R components in the materials apriori, and use these for later testing by presenting the stimulus and asking for the response in typical paired-associate fashion (Anderson and Lyrow, 1971; Andre, 1971; Crouse, 1970, 1971; Wong, 1970). This seems to help in the probability of detecting RI, as all except Wong found statistically significant decrements, but it remains a judgment on the part of the investigator and doesn't get us completely away from the nominal-functional stimulus problem pointed out by Underwood in 1963.

For example, one investigator (Peairs, 1958) rationally decided that the subject of the sentence should operate as the stimulus and the predicate the response. That sounded like a good idea, but Andre (1971) decided that it would be a better idea to turn it around and call the subject of the sentence the response, and the predicate the stimulus. Hmmm.

Other means of designating similarity have been to: use Venn diagrams and to construct structurally similar materials (Anderson, 1971); to construct syntactically similar sentences (Feather and Miller, 1964); and to use differing orders of approximation to English in either the original and/or the interpolated learning (Kini and Cofer, 1960). None of these investigations uncovered statistically significant evidence of RI.

In sum, it seems that similarity is a determining variable, just as it has been with paired-associate learning. If RI is to be consistently demonstrated with prose materials, one necessary requisite is to find a mode of determining similarity that agrees with the set for psychological similarity that the subjects bring with them.

Time Between OL, IL, and Criterion Test

In 1948, Underwood changed the emphasis in forgetting of verbal learning by showing that the effects of retroactive interference dissipate in the first
48 hours, while the effects of proactive interference remain. These results corroborated findings obtained earlier by McGeoch (1933) and fairly well established proactive interference as the more persuasive in the forgetting of verbal learning.

The results with prose learning studies do not seem to uphold these findings consistently. The standard design for many studies has been to conduct the original learning, interpolated learning, and test of original learning all in one sitting. Of the studies reported that were conducted in this manner, most found RI in some form or another. However, if the Underwood and McGeoch findings hold for prose, one would have to predict that the effects of PI would dissipate over time and less RI would be found in delayed retention tests held 48 hours or more after IL. This does not seem to be a consistent result. Some studies employing delayed testing for up to five weeks found RI. One found it at five weeks when it hadn't been in evidence at the time of the immediate criterion test (Anderson and Byers, 1971). Another investigation using poetry materials found increased RI after seven days as compared with an immediate test (McGeoch and McKinny, 1934).

Only one investigation was designed especially to look at these differing effects due to spaced interpolation of the IL and test. Stratil (1970) varied both the timing of the IL and the timing of the test and found RI at the immediate criterion test when all learning occurred at one sitting, and at a 7-day retention interval following the OL-IL sequence. He did not find RI, however, when the 7-day interval was interspersed between OL and IL with the criterion test coming directly after study of the IL.

To tidy up these results, timing seems to be an important factor. As has been shown in the verbal learning laboratory, forgetting is probably not just a matter of interference operating singularly in some linear fashion. If the
hypothesized mechanisms of unlearning and response competition change in relative strength over time, it remains to be seen how they interact in the forgetting of meaningful prose.

Mode of Criterion Testing

There is standing evidence in the verbal learning literature that the mode of testing, either recall or recognition, influences quantitatively the measured degree of learning and recall. According to Postman (1952), recognition tests are generally less sensitive to differences in retention than recall tests. This should be an important factor with studies using prose materials because of the relatively smaller degree of RI obtained in comparison with paired-associate studies.

This speculation seems to hold up. Those studies reviewed herein which used solely multiple choice questions in the measurement of retention failed [with one exception (Gilman, 1970)] to find statistically significant RI (Anderson and Byers, 1971; Ausubel, et al., 1958, 1957, 1966, 1969; Gaite, et al., 1969; Shuell and Hapkiewicz, 1969; Wong, 1969). Those employing some kind of recall procedures generally achieved results indicative of RI.

However, in addition to quantitative considerations, there is some speculation whether recall and recognition tests measure the same thing. If the hypothesized mechanisms of response competition and unlearning operate with prose learning, then the two types of tests should tap different factors during different time intervals between IL and test. This begins to be a little complicated, and it is going to take well-designed, tightly-controlled, factorial studies to ferret out these various effects.

Strength of Learning

Only one study has been reported that was geared to look specifically at
this variable. King (1966) used a memory drum in the presentation of two levels of learning contrasted with three levels of meaningfulness and found significant RI only when the level of OL learning was low, no matter how high the meaningfulness of the material. He goes on to comment that the low amounts of interference often reported with meaningful material are probably due to the high levels of learning of OL. Jensen and Anderson (1970) also attributed failure to find RI in some studies to excessive strength of the original learning. Hills and Kessel (1965) specifically selected difficult materials with the idea in mind that they would probably be favorable to the detection of RI. Both of these studies uncovered significant RI effects.

While no studies have been performed looking directly at the effects of strength of IL, there does seem to be some evidence that when IL is freshly learned and thus discriminable of OL, RI effects are attenuated (Stratil, 1970; Gillman, 1970).

As with paired-associate learning, strength of original and interpolated learning appears to be a significant factor that should be more thoroughly investigated.
Discussion

The initial rush for THE study to demonstrate NI with prose materials is over. There seems to be little doubt that interference mechanisms operate in some way in the learning of meaningful prose despite the demurr of Ausubel and his colleagues. There also seems to be no doubt that the forgetting of prose materials is a complex business affected by all of the variables discussed herein and perhaps others not specifically cited.

There seem to be two paths that further research in this area may take: (1) A search for and more exquisite delineation of the variables at work and their relative influence on inhibition in prose learning. Some of these variables and tentative comments about their importance are registered above; and (2) investigations into ways to mitigate the influence of these variables in facilitating learning and retention of prose materials in the normal learning situation.

Educational Psychology is an applied science. Abstract theory is a delightful luxury --- one that serves its purpose in organizing ideas and in directing research. However, the second of the paths delineated above --- the one leading to practical applications of the knowledge we gain --- is just as appropriate as research aimed at developing more elegant theories. It is important that we begin to test out interferences concerning how to make use of research-verified information in directing and facilitating learning.

Anderson and Myrow (1971) moved in this direction by gathering verbal report data about the strategies employed by subjects in learning their materials. They showed that some strategies seemed to be more effective than others.

Another suggested mode that should be more thoroughly examined involves the use of advance organizers and other integrative mechanisms suggested by Ausubel (1963, 1968). To date these have achieved only minimal success (Ausubel, 1960; Ausubel and Fitzgerald, 1962; Ausubel and Youseff, 1966; Ausubel and Fitzgerald, 1961; Wong, 1971, 1972).
Other methods involving the manipulation of the learning materials in order to enhance learning are under investigation by experimenters (Rothkopf, 1965; Frase, 1968; Anderson, 1970). Hopefully, these are just beginnings that, along with others, will point the way to school applications of prose learning information.
References


Hall, J. F. Retroactive inhibition in meaningful material. *Journal of Educational Psychology*, 1955, 46, 47-52.


King, D. J., and Cofer, C. N. Retroactive interference in meaningful material as a function of the percent of contextual contrasts in the original and interpolated learning. *Journal of General Psychology*, 1960, 63, 145-158.


