

## Study and test formats in learning factual information

DONALD F. PRATT and ALBERT E. GOSS  
*Douglass College, Rutgers University, New Brunswick, New Jersey 08903*

Factual information in 4 by 4 or 4 by 2 matrices was presented to 192 college undergraduates in prose, constrained prose, tabular, or paired associate format for study, and was tested in each of those formats by reproduction or reconstruction of missing items. There were two study-test acquisition trials and a later retention trial with each topic. Totals of correct reproductions or reconstructions did not differ among study formats and differed only slightly among test formats. The latter differences reflected slightly more correct reconstruction responses on the first trial for constrained prose and tabular formats than for prose and paired associates formats. The essential invariance of acquisition and retention among study and test formats suggests that the many manifest or conjectured differences among formats are not effective variables. Moreover, this invariance reorients theory and experimentation on meaningful vs. rote learning.

Students encounter factual information in many different formats: ordinary prose and specially constructed prose in workbooks and programmed texts, outlines and tables, graphs and diagrams, and amorphous notes. Their acquisition and retention of this information is then tested in equally diverse formats that may resemble or differ from the study formats.

Factual information on four different topics was presented for study in one of four different formats: natural prose, constrained prose, paired associates, and tables. Acquisition and retention of this information was tested by reproduction (fill-in) or reconstruction (matching) (Luh, 1922) within each of these formats.

Information about the influence of study format and test format on acquisition and retention with different response conditions is useful as a guide to the preparation of learning materials. If study and test formats make no difference, they can be used interchangeably. If some combination proves optimal for reproduction, reconstruction, or both, it may be used when optimization of acquisition and retention is desired.

Ausubel (1963) expresses a widely held view that concepts and principles of meaningful verbal learning differ from those of the rote learning of lists of verbal stimuli. For him, the critical difference is that "meaningful learning as a process presupposes, in turn, both that the learner employs a meaningful learning set and that the material he learns is potentially meaningful to him" (Ausubel, 1963, p. 22). However, the meaningful and rote-learning tasks that Ausubel

considered also differed in kinds of stimuli, patterns of stimulus-response relationships to be acquired, formats of presentation, and response conditions.

In order to establish the viability of Ausubel's contention, it is necessary to show that with the same stimuli, the same patterns of stimulus-response relationships to be acquired, and the same response conditions, acquisition and retention do not differ among study and test formats including prose and lists. If formats make little or no difference, the issue of meaningful vs. rote learning can be pursued in terms of meaningful learning as a process. Thus, information about the influence of study and test formats on acquisition and retention may also guide experiments on meaningful vs. rote learning.

Consideration of features of the study and test formats used here does not yield clear-cut expectations of the direction and degree of their influence on acquisition and retention of the factual information of the topics. For each topic, that information comprises four characteristics of each of four persons, places, or things. Table 1 simulates the tabular study format of the topic "first women in politics." The other topics were "carnivorous plants," "natural resources of central Africa," and "things invented twice."

The prose format presents the stimuli of this information, and their interrelationships, in the context of natural grammar. The constrained-prose format is similar, but extraneous words do not vary, and response terms are underlined for easier identification.

Extraneous words are eliminated in the paired associate format, in which patterns of stimulus-response relationships are realized by contiguous placement of stimulus and response terms. The tabular format uses row and column headings to eliminate repetition even of stimulus terms.

John L. Santa and William B. Gillooly contributed to the planning and the analyses of this study. Requests for reprints should be sent to Donald F. Pratt, Department of Psychology, Douglass College, New Brunswick, New Jersey 08903.

**Table 1**  
**Example of Topic in Tabular Study Format: First Women in Politics**

	Name	State	Date	Served for
First Woman Mayor	Susannah Salter	Kansas	1887	1 year
First Woman State Senator	Martha Cannon	Utah	1896	2 terms
First U.S. Congresswoman	Jeannette Rankin	Montana	1917	2 separate terms
First Woman U.S. Senator	Rebecca Felton	Georgia	1922	2 months

Other differences among formats may also influence acquisition and retention. Prose and tables may be more familiar than paired associates. The stimulus terms of the paired associates are long and redundant. However, in this and the tabular format, the patterns of stimulus and response relationships to be acquired are in more regular order and are more clearly identifiable than in constrained prose and, particularly, in prose. The tabular format clarifies the two or three dimensions of the information. These dimensions are also in the stimulus terms of the paired associates, although this may be less evident. Except for the object named, the initial part of each sentence of constrained prose is the same: This might facilitate elimination of extraneous words, or it might increase intersentence interference.

The results of the few previous investigations of format differences in acquisition and retention contribute little to the resolution of these uncertainties. Among Goss and Loigman's (1972) conditions were pairs of words in conventional paired associate and prose formats. Acquisition was faster with the prose format. On the basis of other findings, the reason for this superiority appeared to be that the sentential context of the prose format made it easier for subjects to identify and rehearse pairs of words of each unit.

In comparisons of acquisition with hierarchical diagrams and with paired associates, Nelson and Smith (1972) used single letters and Deffenbacher, Masicik, and Jarombek (1974) used biological classification terms. Acquisition was faster with the diagrams than with the paired associates. However, the presentation of relationships among stimuli was more complex for the paired associates than for the hierarchical diagrams.

Bureaucratic prose, short sentences, algorithms (decision trees), and tables were the formats used by Wright and Reid (1973) to present fictional directions for choosing modes of space travel. Through successive trials with directions no longer present, correct choices of modes decreased when the directions had been algorithms or tables, improved when the directions had been prose, and were constant when the directions had been short sentences. Within this complex interaction, no format produced more correct choices consistently.

Yekovich and Kulhavy (1976) presented topics of "woods," "wines," and "stars" (Level 1) with names of constituent classes of objects at Level 2 and their attributes at Level 3 (names), with attributes at Level 2 and names at Level 3 (attributes), and with names,

attributes, and values of attributes (Level 4) assigned randomly (random). These names, attributes, and random organizational forms occurred in a prose format (read) or in an hierarchical list that resembled a table. The hierarchical list was read (list) or subjects constructed a narrative (generate). These read, list, and generate study conditions were followed by a test phase in which subjects first reproduced critical list words in free recall and then recognized them within a list that also included 34 semantically related words and 34 distractors.

Each subject had three such study-test trials with each topic within one of the nine combinations of organizational form and learning conditions. Two days later the subjects attempted to reproduce the critical list words and to recognize them.

Across other factors, means of correct critical words for learning or study conditions were in the order list > generate > read. In the interaction of form and condition, the overall order of conditions held for attribute and random organizational forms. For name organizational form, the order was list > read > generate. The overall order in free recall only was list > generate > read with no interaction that involved study condition and organizational form. The hierarchical task or tabular format, particularly when simply read, led to better recall and recognition than did the prose format.

## METHOD

### Design

Each subject studied all four topics in the same format and was always tested by reproduction or reconstruction. Test format, topic, and order of occurrence were combined in a Graeco-Latin square, with a different subject assigned to each row of the square. Thus, subjects were tested in each format with a different topic for each format. Columns of the square defined the order of occurrence of each combination of test format and topic. Study format was crossed with test format and topic but test format and topic were not orthogonal.

### Materials

The four topics were selected within three major criteria: (1) variety of content; (2) similarity of the content of each topic to that encountered in general reading, but novelty of that content; and (3) possible arrangement of the content in a 4 by 4 or 4 by 2 by 2 matrix with a unique element in each cell. These matrices correspond to the tabular format.

The paired associate format was derived from a matrix by forming a compound stimulus from row and column headings and pairing it with the appropriate cell entry. Constrained prose was derived by forming sentences from rows of a matrix

with provision for the row and column headings. For the prose format, cell entries of rows of the matrix were scrambled and presented in one or two sentences. The order of rows in the table and of corresponding parts of the other study sheets was determined randomly except for "first women," where a natural and logical order of date and rank of offices was preserved.

Test formats were derived from study formats by deleting the 16 response terms and replacing them with underscored spaces of constant length. Within each test format, three different random orders of the rows were constructed. A different order was used on each test trial to keep subjects from rehearsing items without regard to row heading. Also, the first row of each of these orders differed from that of the study format to assure that subjects detected the change in row order immediately.

For reproduction, the subject filled in the appropriate response in each of the 16 spaces. For reconstruction, cell items were shown below the test, grouped by type of item, and listed alphabetically within a group. Because pilot subjects took longer for reconstruction than reproduction, abbreviations of nonnumerical items were provided, and subjects were told to use them. Subjects matched an item or its abbreviation by writing it in the appropriate cell.

#### Procedure

**Preparation of materials.** The study and test materials were in purple ditto on letter-size sheets. The sheets were assembled and stapled to form two booklets, one for study-test trials, the other for retention trials. The first booklet began with sample study and test sheets on the topic "first men to fly." These exposed the subjects to the study format of the learning materials and to the corresponding test format. Study and test sheets for the four topics to be learned then occurred in the prearranged order for a particular subject, separated by blank sheets.

**Presentation of materials.** Subjects were run in groups of varying size. The experimenter read the initial instructions to the subjects as they looked at the sample study and test sheets. A cassette tape deck was then turned on and prerecorded instructions led the subjects through the study-test trials. During each trial there were 8 sec for study instructions, 80 sec for study, 8 sec for test instructions, and 144 sec for the test.

Subjects received two successive study-test trials on each topic. After these trials, they took the comprehension part of the Nelson-Denny reading test. This took about 25 min. The subjects next took up the booklet for the retention tests, which consisted of a test sheet for each topic in the format of the acquisition trials, with order of the topics reversed. The tape deck was turned on again, and the subjects were paced through the tests with 144 sec for each. The session ended with several questions about the subjects' manner of studying and of taking the tests.

#### Subjects

Of the 213 men and women students from the course in introductory psychology at Douglass College who participated, data on 192 were used. After most of the subjects had been run, the distribution of Nelson-Denny scores was examined and found to be distinctly bimodal with an apparently separate distribution of scores of 30 or below. The 17 subjects of this distribution were replaced by later subjects who scored 32 or above. One potential replacement failed to score 32, and three other potential replacements were not needed.

## RESULTS

#### Scoring

A correct reproduction consisted of at least half of the appropriate item for a particular cell. However, first names alone were not considered correct. A correct

reconstruction consisted of the appropriate item for a particular cell, the abbreviation provided or any other unambiguous abbreviation, or a simple misspelling.

The total of correct reproductions or reconstructions for each test trial on each topic was the score of the principal statistical analysis. Stricter scores that corrected for guessing and more lenient scores that allowed partial credit for row-confusion errors were also analyzed.

In all analyses, only subject was considered a random variable. Whether topic is a fixed or random variable is equivocal. With topic as a fixed variable, the power of the statistics increases.

#### Analyses

Table 2 shows the means of total correct reproductions and reconstructions for all combinations of study and test formats. Study format made no difference. Scores on constrained-prose (mean = 9.7) and tabular (mean = 9.7) test formats were slightly higher than scores on paired associate (mean = 9.3) and prose (mean = 9.1) formats [ $F(3,504) = 2.66$ ,  $p < .05$ ]. Study or test format entered no first-order interactions. Together they interacted with response condition [ $F(9,504) = 1.94$ ,  $p < .05$ ] and with trial [ $F(18,1008) = 1.64$ ,  $p < .05$ ].

More correct responses occurred with reconstruction (mean = 11.3) than with reproduction (mean = 8.6) [ $F(1,160) = 17.81$ ,  $p < .001$ ]. Correct responses varied among topics [ $F(3,504) = 112.02$ ,  $p < .0001$ ]. Means for reconstruction and reproduction for topics were, respectively, 11.8 and 10.8 for "carnivorous plants," 11.5 and 8.5 for "first women," 9.5 and 9.5 for "natural resources," and 8.3 and 5.7 for "things invented twice" [ $F(3,504) = 16.10$ ,  $p < .0001$ , for this interaction].

Correct responses increased with order of occurrence [ $F(3,504) = 61.83$ ,  $p < .0001$ ]. The negatively accelerated increase is typical of warm-up and learning to learn.

Further analyses on subsets of data show that test format made a difference only for reconstruction and only on Trial 1. The lower scores occurred when subjects were first tested in either the paired associate format or the prose format after studying with one of the other

Table 2  
Means of Correct Reproductions and Correct Reconstructions  
for Combinations of Study Formats and Test Formats  
Across Topics and Trials

Study Format	Correct Reproductions				Correct Reconstructions			
	Test Format				Test Format			
	P	CP	T	PA	P	CP	T	PA
P	9.2	9.2	8.9	8.1	10.6	10.6	10.7	10.2
CP	7.7	7.4	8.8	9.1	8.5	10.5	10.2	9.1
T	9.3	8.7	9.4	8.8	9.4	10.8	10.5	9.8
PA	7.8	9.4	8.1	8.3	10.5	10.6	10.9	11.1

Note—P = prose; CP = constrained prose; T = table; PA = paired associates

three formats. For stricter test scores, test format had a more pronounced effect; for more lenient scores, test format had a less pronounced effect. These shifts reflect a penalty on row-confusion errors with strict scoring, and partial credit for such errors with more lenient scoring.

The interaction of test format, study format, and response condition was limited to Trial 3 and, on this trial, to an interaction of study and test formats only for reproduction. The mean for reproduction for study and test with the constrained-prose format was anomalously low.

Different but comparable groups of students produced associations to topics, row and column headings, and cell items, and rated these stimuli for age, familiarity, imagery, pronounceability, and number of associations (Pratt & Goss, 1977). Correlations between values of the attributes of cell items and correct reconstruction of those items were uniformly low and nonsignificant. While correlations between values of attributes and correct reproductions were somewhat higher, this was an artifact of the difficulty of reproducing dates. Accordingly, no analysis of covariance to adjust for values of attributes of cell items was undertaken.

## DISCUSSION

For totals of correct responses and other scores, acquisition and retention did not vary among study formats. Acquisition and retention were better for constrained-prose and tabular test formats than for paired associate and prose test formats. But this advantage was limited to reconstruction on Trial 1, when a shift to the paired associate format or the prose format from any other format led to more row-confusion errors. There was no other evidence of an interaction of study and test formats in which a match of study and test format was superior to a mismatch.

The essential invariance of acquisition and retention among the study and test formats suggests that the many manifest or conjectured differences among the formats are not effective variables. Any differences that may exist between rote and meaningful learning of factual information of the kind used here apparently do not depend on format of presentation. For the formats most representative of the extremes of rote and meaningful presentation, paired associates and prose, there is even less evidence of a difference.

From subjects' answers to questions about their manner of studying and taking tests, there was no evidence that different formats encouraged different strategies, such as reliance on spatial memory for tables and verbal memory for prose.

The possibility remains that different principles do underlie acquisition and retention in the different formats, and that somehow these principles interact in some delicately balanced way among the different formats. Various explanations might be concocted; they appear unnecessary and would not be parsimonious.

There is always risk in defending a null hypothesis. However, the statistical tests for differences among study formats had enough power to detect any true difference large enough to be of any practical importance. Furthermore, Deffenbacher et al. (1974), Goss and Loigman (1972), Nelson and Smith (1972), and Wright and Reid (1973) did not obtain unequivocal differences between or among study formats.

Yekovich and Kulhavy's (1976) list and generate (tabular) formats of study led to more items correct across recall and

recognition than did the prose format. In the interaction of organizational form, study condition, and trials shown in their Figure 2, the mean numbers of items correct on Trials 1 and 3 are almost identical for the combinations of name organizational form with list and read study conditions. The difference over all three trials is negligible. These are the combinations that are most similar to the tabular and the constrained-prose/prose study formats of the present study.

There are limits to the generality of the results. The college students who served as subjects have studied information on many topics presented in a variety of formats, and have also been tested in a variety of formats. Their sophistication may have overridden or obscured differences among study and test formats that might occur with more naive individuals.

The particular patterns of the factual information of the four topics are another limitation. However, either implicitly or explicitly, these and similar patterns are common in factual information. Use of the same stimuli in the same patterns in each format corrected a defect of didactic comparisons of rote and meaningful learning such as Ausubel's (1963).

For the same stimuli and the same patterns of stimulus-response relationships, acquisition and retention were essentially invariant over study and test formats for reproduction and reconstruction. Those who share Ausubel's view can proceed to investigate meaningful vs. rote learning in terms of differences in process, specifically, of differences in meaningful learning set and potential meaningfulness of the material to the learner. Those who hold that an essentially common set of concepts and principles explains meaningful and rote learning (e.g., Goss, 1971; Musgrave & Cohen, 1971) can proceed to investigate possible effects of stimulus-attribute, frequency, temporal, and other variables across study and test formats.

## REFERENCES

- AUSUBEL, D. P. *The psychology of meaningful verbal learning*. New York: Grune & Stratton, 1963.
- DEFFENBACHER, K. A., MISCIK, J. G., & JAROMBEK, J. Acquisition and forgetting of information in long-term memory as a function of certain hierarchical structure variables. *Bulletin of the Psychonomic Society*, 1974, 4, 590-592.
- Goss, A. E. Paired associates and connected discourse in the acquisition of knowledge. In E. Z. Rothkopf & P. E. Johnson (Eds.), *Verbal learning research and the technology of written instruction*. New York: Teachers' College, 1971.
- Goss, A. E., & LOIGMAN, B. M. Reconciliation of opposites in verbal learning: Genotypical relationships, phenotypical situations. *Abstract Guide of XXth International Congress*. Tokyo: Science Council of Japan, 1972.
- LUH, C. W. The conditions of retention. *Psychological Monographs*, 1922, 31, No. 142.
- MUSGRAVE, B. J., & COHEN, J. C. Relationships between prose and list learning. In E. Z. Rothkopf & P. E. Johnson (Eds.). *Verbal learning research and the technology of written instruction*. New York: Teachers College, 1971.
- NELSON, T., & SMITH, E. E. Acquisition and forgetting of hierarchically organized information in long-term memory. *Journal of Experimental Psychology*, 1972, 95, 388-396.
- PRATT, D. F., & GOSS, A. E. Age, familiarity, imagery, pronounceability, and meaningfulness of verbal units of factual information. *Bulletin of the Psychonomic Society*, 1977, 9, 325-328.
- WRIGHT, P., & REID, F. Written information: Some alternatives to prose for expressing the outcomes of complex contingencies. *Journal of Applied Psychology*, 1973, 57, 160-166.
- YEKOVICH, F., & KULHAVY, R. W. Structural and contextual effects in the organization of prose. *Journal of Educational Psychology*, 1976, 68, 626-635.