SESSION IV

ON-LINE COMPUTERS IN TEACHING PSYCHOLOGY: I. TEACHING COGNITIVE PROCESSES

N. JOHN CASTELLAN, JR., *Indiana University, President*

Teaching cognitive processes: Software for laboratory instruction in memory and cognition

JANICE M. KEENAN and ROBERT A. KELLER
University of Denver, Denver, Colorado 80208

This paper describes a set of computer programs that allow students to participate as subjects in well-known experiments selected from the current literature on memory and cognition. The experiments address five major areas: levels of processing, encoding specificity, semantic memory, sentence-picture verification, and constructive processes in prose comprehension.

WHY DEVELOP SOFTWARE FOR LABORATORY INSTRUCTION?

High-quality instruction in science requires not only that students become acquainted with the facts but that they also be trained in the research methods by which the facts are obtained. Unfortunately, there are a number of factors, such as space, money, and large class enrollments, that limit the amount and type of laboratory training that can be provided. Many departments do not have the space to house student labs. Nor do they have equipment that can be given over solely to student labs or money to purchase such equipment. Finally, the large number of students enrolled in many psychology courses makes it virtually impossible to organize and supervise realistic laboratory experiences.

Because many psychology experiments require only a subject to emit behavior and an experimenter to present stimuli and record the subject's responses, the campus computer can be programmed to serve as the experimenter while the students serve as subjects. The computer can present stimuli such as prose passages and questions, lists of items, simple pictures, and so on,

We gratefully acknowledge the programming assistance of Susan D. Baillet for the constructive processes in prose comprehension package, Robert Guttentag for the encoding specificity package, and Derek Price for the levels-of-processing package. Development of the software was supported by Grant SER 77-02769 from the National Science Foundation's LOCI program. Preparation of this manuscript was supported by Grant SER 77-06935 from the National Science Foundation's RIAS program.

and collect response measures such as ratings, number of correct responses, response times, number of trials to criterion, or what have you. It can even give the subjects feedback on their performances.

This approach circumvents problems of space and equipment because it uses facilities that the school already owns and operates. The computer terminal becomes the student's laboratory. With a few simple instructions, a student can sit down at a terminal and run him/herself as a subject in an experiment.

There are several advantages to this approach. First, there is no need for supervision in the laboratory; the computer can be programmed to fully instruct the student in the experimental procedure. Second, because the students do not need supervision, they can run the experiments outside of class, thus freeing class time for more in-depth instruction in the issues under study. Third, because the experiments are not run in class, the number of terminals can be much smaller than the number of students in the class, so large enrollments can be accommodated fairly easily. Fourth, unlike traditional labs in which half the students might serve as subjects and half as experimenters, all students can serve as subjects and obtain first-hand experience with the experimental task. In fact, the programs can be set up so that the student can also participate in the design of the study and thus serve as the experimenter as well as the subject. Finally, the data from all subjects can be stored in the computer. This permits easy access to summary statistics and frees the lab instructor from the

tedious task of tabulating each individual's data and calculating class averages.

These are all logistical advantages of a computerized laboratory course. There are also, of course, instructional advantages. Simulations can be constructed based on current models of the psychological processes involved in the task. Computer simulations provide more theoretical training than traditional labs because the students can observe first-hand the predictions of a model in terms of the computer's performance as a subject. They can learn the consequences of various assumptions and theoretical statements by manipulating parameters in the simulation. And, because the students also run themselves as subjects, they can use their performances on the task to evaluate the model's performance.

All of these hypothetical advantages became realities when Bewley (1974) published software for implementing experiments and models in cognitive psychology. We began using Bewley's programs as soon as they became available, and student response to them was overwhelmingly positive. Learning was easier and more enjoyable when the issues being developed in the classroom could be readily and meaningfully experienced in the computer lab. But there was one problem with Bewley's programs: Many of the experiments and models quickly became outdated.

Several major theoretical and empirical advances occurred in cognitive psychology in the early 1970s, such as the notion of levels of processing in memory, the distinction between episodic and semantic memory, and the increasing emphasis on the constructive nature of human cognition. These advances led to new research issues, new theories and models, and even new experimental paradigms. Faced with a discrepancy between what we were discussing in the classroom and what Bewley's (1974) programs offered in the lab, we decided to build on Bewley's pioneering efforts and develop new software that would address more current issues. (Levy, Fischler, and Griggs, 1979, began developing software for a laboratory in cognition and perception at approximately the same time. Interestingly, there is virtually no overlap between our package and theirs.)

Our first step was to obtain a grant from the National Science Foundation's Local Course Improvement (LOCI) program to support the project. Under this grant, we have completed all of the software development and have just about completed the student's and instructor's manuals. When these manuals are completed, we expect to have CONDUIT distribute the package. The software has been in use for over a year, not only in our courses at the University of Denver, but also at a small number of other institutions throughout the nation.

DESCRIPTION OF THE SOFTWARE

The programs are all written in BASIC. They are designed to allow students to run themselves as subjects

in well-known experiments from the current literature on memory and cognition. In addition, some programs provide either a simulation of the task based on a well-known model or a demonstration of the model, or both. There are five different sets of programs; each covers a different principle of cognition. The five are: levels of processing, encoding specificity, semantic memory, sentence-picture verification, and constructive processes in prose comprehension. Within each topic area, there are one or more experiments, each of which takes less than 45 min.

Levels of Processing

The levels-of-processing package teaches students that the nature of the encoding process plays an important role in determining subsequent memory performance. The package consists of two programs that partially replicate experiments reported by Craik and Tulving (1975).

The first experiment addresses depth of processing. It employs three encoding tasks that are designed to induce different depths of processing on word items. The structural or form task is assumed to induce processing at a fairly shallow level. It requires subjects to evaluate the pattern of consonants (C) and vowels (V) in a word. A sample question is, "Is a CCVVCC?", where the word SPEECH would yield a positive response and the word COPPER a negative response. The rhyme task requires subjects to evaluate the sound pattern of the word (e.g., Rhymes with CLAIM?), and it is assumed to induce deeper processing than the form task. The semantic task is assumed to involve the deepest level of processing. It requires subjects to determine whether a word is an instance of a given conceptual category (e.g., Is a type of weapon?). The program presents 60 of these encoding trials, 20 of each type, with 10 items requiring a positive response and 10 a negative response. The data collected during this encoding phase consist of the subjects' responses and their response times.

The levels-of-processing theory assumes that the deeper the level of encoding on an item, the better it will be remembered. Memory in this experiment is tested by an unexpected yes-no recognition task.

At the end of each run through an experiment, our programs present a summary of the individual's data. In addition, the programs offer the options of saving the data in a common data file for that experiment and of summarizing this class data file. The option of summarizing the class data file is offered at both the beginning and the end of each program. It is offered at the beginning so that the instructor can examine the class data without having to go through the experimental session. It is offered at the end of the experimental session, immediately after the student sees the summary of his or her own data, so that the student can compare his or her own performance to that of the group.

Figure 1 is a screen from a run of the program, presenting the class summary data from seven subjects who participated in the first levels-of-processing experi-

ment. The data consist of the average number correct, the mean response time to the encoding questions, and the average number of words correctly recognized in the memory test for each of the types of questions. The data are in accord with the levels-of-processing theory in that recognition performance is highest on items encoded under the semantic task—the task that induces the deepest level of processing. Students are especially impressed with this result when it is viewed in light of the encoding times: Although category items receive the least amount of processing in terms of time, they are remembered the best. Thus, it is not the amount of study time that determines memory performance, but rather the qualitative nature of the studying time in terms of depth of processing.

The second levels-of-processing experiment is designed to show that retention depends not only on the depth of processing but also on the amount of elaboration involved in the initial encoding of the event. The encoding task requires the subject to determine whether or not a word fits in a sentence frame. Sentence frames are either simple or complex, as shown in Figure 2. The student can specify the number of encoding trials, up to a maximum of 60. Both the subject's response and the response times are recorded by the program. Response times are measured from the onset of the word to the subject's response. Again, the memory test is incidental, although in this case it consists of a free recall test, rather than a recognition test.

The levels-of-processing theory states that the more elaborate the initial encoding of an item, the better it will be remembered. For this experiment, it predicts that items encoded with complex sentence frames will be remembered better than those encoded with simple sentence frames. The theory also invokes the principle of congruity, stating that elaboration will occur only when the item is congruous with the encoding framework. In other words, the prediction that items encoded with complex sentence frames will be remembered better than those with simple frames only holds for cases involving a "true" response, that is, when the items actually fit in the sentence frame.

The data always provide strong support for the theory, as can be seen in Figure 3. This figure presents the class data summary from seven subjects. It shows

```
.... IS WHEN THE LEAVES FALL.
AUTUMN? Y
THE CHILD PLAYED WITH A YELLOW ..... IN HER WICKER CRIB.
CLOSED? N
THAT PARTICULAR ..... MINE HAD BEEN CLOSED FOR NEARLY TWENTY YEARS.
SILVER? Y
TOSS THE ..... TO ME
GARDEN? N
```

Figure 2. Sample trials from the encoding phase of the second levels-of-processing experiment.

LOP2 DATA SUMMARY		30-APR-79		7 SUBJECTS	
POSSIBLE 5 CORRECT PER ITEM TYPE					
SENTENCE	FRAME TYPE	ENCODING DECISIONS		MEAN RT	RECALL
		AV. # COR.			AV. # COR.
SIMPLE	TRUE	5		1291 (7)	1
SIMPLE	FALSE	5		1598 (7)	1
COMPLEX	TRUE	5		1459 (7)	4
COMPLEX	FALSE	5		1441 (7)	1

Figure 3. Summary of data from the second levels-of-processing experiment.

that items encoded with respect to complex sentence frames are in fact remembered quite a bit better than those encoded with simple sentence frames, but only when the words actually fit the sentence frames (complex trues).

Encoding Specificity

The encoding specificity package is concerned with retrieval operations. It consists of three experiments. The first allows the student to compare performance under three types of memory tests: recall, recognition, and cued recall, where the cues are strongly associated extralist cues. The student has the option of running a between- or within-subjects design. In a within-subjects design, the subject receives all three memory tests, each on a different list of 24 items. In a between-subjects design, the subject receives only one list and one test. Subjects specify the type of test in a between-subjects design or the ordering of tests in a within-subjects design by entering a number that is assigned to them by the instructor. In this way the instructor can counterbalance the design, and the subjects do not know which type of memory test they will receive until after they have studied the list.

Once the design is selected, a list of 24 words is presented at a 3-sec rate, and memory is tested by recall, recognition, or cued recall. When the test is cued recall, the cues are presented one at a time, and a response is required before the next cue is presented. This is shown in the top part of Figure 4. Then, in order to simulate the subject looking back over a piece of paper with all the cues, the program presents the cues again all at once, allowing the subject to look over them and input any more targets that can be remembered, as shown in Figure 4. The results of this experiment, shown in Figure 5, always conform to the student's intuitions: recognition > cued recall > recall. The next two experiments in this package, however, challenge the principles

LOP1 DATA SUMMARY		30-APR-79		7 SUBJECTS	
POSSIBLE 16 CORRECT PER ITEM TYPE					
QUESTION TYPE		ENCODING DECISIONS		MEAN RT	RECOGNITION
		AV. # COR.			AV. # COR.
TRUE FORM		8		2502 (7)	7
FALSE FORM		9		2403 (7)	6
TRUE RHYME		9		1747 (7)	7
FALSE RHYME		9		1743 (7)	5
TRUE CATEGORY		9		1624 (7)	9
FALSE CATEGORY		9		1719 (7)	8

MEAN NUMBER OF FALSE RECOGNITIONS = 0

Figure 1. Summary of data from the first levels-of-processing experiment. (Note—All figures are screens from actual runs of the program.)

```

WOOL          ? SHIRT
SEA           ? BAIT
PEACE         ? BRANCH

PRESS RETURN TO RECEIVE RECALL CUES AGAIN?

JUMP    PEACE    ITCH    BRANCH    ERROR    MIND
MOTOR   ANIMAL   WOOL    SEA      ATHLETE  WAR
MILLION ANCHOR   OLIVE   SPOIL   LIFT     DIE
JUDGE   LEMON    HAPPY   CONSULT HEAT     RIBBON

NOW TYPE IN ANY MORE TARGETS YOU CAN RECALL, AND
TYPE IN 'END' WHEN YOU ARE DONE.

? ALLOW
? UNIFORM
? END

```

Figure 4. A sample of the cued recall memory test in the encoding specificity package.

```

ENTER 1 FOR WITHIN SUBJECT DATA, 2 FOR BETWEEN? 1

DATA SUMMARY FOR ESP1      6 SUBJECTS      07/24/79

MEAN % CORRECT      N
RECALL:             24.3      6
RECOGNITION:        88.2      6
CUED RECALL:        66        6

PROGRAM ESP1 COMPLETE

```

Figure 5. Summary of data from the first encoding specificity experiment.

underlying the student's predictions for this first experiment. This technique of convincing students that their intuitions are correct and then showing them a series of counterintuitive results really has an impact; they long remember the encoding specificity principle.

The second experiment in this package is a partial replication of Thomson and Tulving's (1970) study on the effectiveness of cues. The subject studies a list of 24 cue-target pairs in which the cues are weak associates of the target. Sample trials from an actual run of the program are shown in Figure 6. On the first trial, these same cues are presented again at the recall test. On the second trial, the subject again studies a list of 24 weak cue-target pairs, but this time the nature of the recall test depends on which condition the subject is in. In Condition 1 the subject receives the same weak cues at recall as in the study list. In Condition 2 the subject receives extralist strongly associated cues, and in Condition 3 the subject receives no cues. Although students generally expect performance with strong cues to exceed that of weak cues, the results are always the opposite, as shown in Figure 7. Weak cues are better than strong cues because they were involved in the initial encoding, thus demonstrating the encoding specificity principle.

The third experiment is a modified replication of Tulving and Thomson's (1973) encoding specificity experiment. On the first two trials, the subject studies and is tested on two lists of cue-target pairs in which the cues are again weak associates to the target. On the third list, the subject again studies a set of cue-target pairs. Before being tested with the cues, however, the subject is given a free association task designed to elicit the targets as responses. This is shown in Figure 8. Then the subject receives a three-choice recognition test on the targets, as shown in Figure 9. Finally, the study

cues are presented and the subject is asked to recall the targets. Often, as indicated in Figure 10, students recall targets that they fail to recognize. Not only do they become convinced of the encoding specificity principle, but they also begin to think more deeply about the relationship between recall and recognition.

Semantic Memory

The semantic memory package consists of a replication of Rips, Shoben, and Smith (1973) and two simulations—one based on Collins and Quillian's (1972) model of hierarchical representations and one based on Smith, Shoben, and Rips' (1974) semantic feature model. The

```

PRESS RETURN WHEN YOU ARE READY TO START.
REMEMBER, THE TO-BE-RECALLED ITEMS ARE ON THE RIGHT.
THE RECALL CUES ARE ON THE LEFT.?

```

CUES	TARGETS
****	*****
WOOL	SHIRT

SEA	BAIT

SPOIL	ALLOW

OLIVE	UNIFORM

Figure 6. Sample study trials from the second encoding specificity experiment.

ESP2 DATA SUMMARY		07/23/79		6 SUBJECTS			
		FIRST LIST		SECOND LIST			
	N	STUDY CUES	TEST CUES	PERCENT RECALL	STUDY CUES	TEST CUES	PERCENT RECALL
CONDITION 1:	2	WEAK	WEAK	48	WEAK	WEAK	60.4
CONDITION 2:	2	WEAK	WEAK	45.3	WEAK	STRONG	48.2
CONDITION 3:	2	WEAK	WEAK	41.7	WEAK	NONE	16.7

Figure 7. Summary of data from the second encoding specificity experiment.

```

BEFORE RECALLING THE MEMORIZED ITEMS, YOU ARE GOING TO
TAKE PART IN A FREE ASSOCIATION TASK. FOR EACH OF
THE FOLLOWING 24 WORDS, TYPE IN THE FIRST FOUR WORDS THAT
COME TO YOUR MIND. TYPE THE FOUR RESPONSES TO EACH
WORD ON THE SAME LINE AS THAT WORD, SEPARATING
THE DIFFERENT RESPONSES BY COMMAS.

```

```

PRESS RETURN?

LOW          ? HIGH, BLOW, CEILING, BELOW
LAKE         ? RIVER, WATER, FISH, POND
SQUARE       ? ROUND, BLOCK, CUBE, BOX
CLOSED       ? OPEN, DOOR, SHUT, SESSION

```

Figure 8. Sample trials from the free association task in the third encoding specificity experiment.

```

NOW INSTEAD OF RECALLING THE TARGET WORDS, YOUR ABILITY
TO RECOGNIZE THE TARGET WORDS WILL BE TESTED. ON EACH
LINE WILL BE PRINTED THREE WORDS, ONE OF WHICH IS FROM
THE TARGET LIST. TYPE IN THE WORD YOU THINK IS THE
TARGET IN EACH SET OF THREE.

```

```

PRESS RETURN?

FEMALE  LITTLE  GIRL
? GIRL

ROCK    HARD    EASY
? ROCK

SPEED   SLOW    DRIVE
? SLOW

```

Figure 9. Sample trials from the recognition test in the third encoding specificity experiment.

```

ESP3 DATA SUMMARY          5 SUBJECTS  31-JULY-79
                             24 ITEMS PER LIST

                             FIRST    SECOND    THIRD
                             LIST     LIST     LIST

MEAN NUMBER OF ITEMS
    RECALLED:  17.8      16.2      18.8
    RECOGNIZED:      19.6
    RECALLED AND RECOGNIZED: 16.2
    RECALLED BUT NOT RECOGNIZED: 2.6
    RECOGNIZED BUT NOT RECALLED: 3.4
    NEITHER RECALLED OR RECOGNIZED: 1.8
    GIVEN AS FREE ASSOCIATES: 12.8
  
```

PRESS RETURN

Figure 10. Summary of data from the third encoding specificity experiment.

SIMULATION COMPLETE. PREPARE TO RECORD DATA.

PERCENT ERRORS = 0 %

CORRECT TRUE RESPONSE SUMMARY AS A FUNCTION OF THE
DIFFERENCE IN LEVELS BETWEEN TARGET & CATEGORY:

# LEVELS	# TRIALS	MEAN RT
0	18	1031
1	18	1334
2	18	1805
OVERALL	54	1390

PRESS RETURN WHEN YOU HAVE FINISHED RECORDING YOUR DATA?

PRESS 1 AND RETURN TO HAVE THE SAME DATA SUMMARIZED WITH
RESPECT TO THE SIMILARITY BETWEEN TARGET AND CATEGORY (AS
SMITH, SHOEN, & RIPS MIGHT DO) OR TYPE A 0 IF THIS IS
NOT NECESSARY: ? 1

Figure 11. Data output from the simulation of Collins and Quillian's (1972) model of semantic memory.

experiment requires subjects to verify simple facts such as "A cow is a mammal" and "A robin is a bird."

According to Collins and Quillian's (1972) model, verification proceeds by traversing a semantic hierarchy in memory. The greater the number of levels separating a subject and predicate in the hierarchy, the longer the verification times. The results of the Collins and Quillian simulation are shown in Figure 11.

The program allows the students to have their data summarized according to Collins and Quillian's (1972) levels. When the students execute this option, they often find that their data do not resemble the simulation data: Level 2 statements (e.g., A dog is an animal) are usually verified faster than Level 1 statements (e.g., A dog is a mammal). This is the result that led Smith et al. (1974) to develop their model in which verification time is a function of the similarity in features between the subject and the predicate. Our program provides a demonstration of the Smith et al. model; one of the trials is shown in Figure 12. According to the model, there are two possible stages of processing. In the first stage, the overall similarity of the subject and predicate is computed. If the similarity is below a criterion, C_0 , a "false" response is immediately executed; if it is above a criterion, C_1 , a "true" response is executed. If it is in between the two criteria, as is the case in the trial shown in Figure 12, a second stage of processing is required. Latency for this second stage is a function of category size.

The Smith et al. (1974) simulation allows students to manipulate the values for the upper and lower criteria. Students can determine approximately what values

would best simulate their data by having the program output their data according to similarity values. When they run the Smith et al. simulation, the data can then be output according to whether the items were above, below, or between the criterion levels specified by the student. This is shown in Figure 13. The data shown in this figure are from a run of the experiment. As the Smith et al. model predicts, latencies are longer for items having a middle level of similarity than for items having extreme levels of similarity or dissimilarity.

Sentence-Picture Verification

The sentence-picture verification package consists of a modification of Carpenter and Just's (1975) experiment and a demonstration of their chronometric model. The experimental task requires subjects to determine whether a sentence accurately describes a picture. An example of one of the trials is the following.

It is true that the star is above the plus.

*

+

This is an example of the true affirmative case—true because the sentence accurately describes the picture, and affirmative because the sentence contains no negative element. The other conditions in this experiment are false affirmative, true predicate negative (It is true that the plus is not above the star), false predicate negative, true denial (It isn't true that the plus is above the star), and false denial. The trials are presented in blocks of 12, two occurrences of each of the six types of trials. The subjects can specify the number of blocks of trials

TRIAL # 38

A CHICKEN IS AN ANIMAL

GENERATED SIMILARITY VALUE = 2.37

SINCE THE SIMILARITY VALUE FALLS BETWEEN THE SPECIFIED
CRITERIA, STAGE 2 PROCESSING WILL BE REQUIRED & ITS LATENCY
WILL BE A FUNCTION OF CATEGORY SIZE.
FOR THIS TRIAL THE CATEGORY ANIMAL IS LARGE
AND 107.292 MSEC OF PROCESSING TIME IS REQUIRED BY STAGE 2.

STAGE 1 PROCESSING IS ONLY DEPENDENT ON THE TYPE OF RESPONSE
MADE. IN THIS CASE, THE RESPONSE = T AND THE TIME
REQUIRED IS 931.465 MSEC.

THE TOTAL LATENCY FOR THIS TRIAL, THEN, = 1038.757
AND THE RESPONSE IS CORRECT.

(PRESS RETURN TO CONTINUE)?

Figure 12. A trial from the demonstration of Smith et al.'s (1974) model of semantic memory.

PERCENT ERRORS = 7 %

CORRECT LATENCY SUMMARY AS A FUNCTION OF SPECIFIED
CRITERIA VALUES:

SIMILARITY	TRUE RESPONSES		FALSE RESPONSES	
	# TRIALS	MEAN RT	# TRIALS	MEAN RT
BELOW C_0	0	0	26	640
BETWEEN	9	851	15	1100
ABOVE C_1	45	634	0	0

Figure 13. Data from one run through the semantic memory experiment summarized according to the criterion ranges of the Smith et al. (1974) model.

PROGRAM VERIFY SUMMARY		6 SUBJECTS				05/29/79	
TRIAL TYPE	TA	MEAN FA	PERCENT TP	CORRECT FP	TD	FD	
BLOCK #							
1	100	83	83	100	50	83	
2	100	100	100	100	67	100	
3	100	83	83	83	83	83	
TRIAL TYPE	MEAN RT TA	FOR FA	CORRECT TP	RESPONSES FP	(MSEC) TD	FD	
BLOCK #							
1	1326	1521	2147	1906	4060	3085	
2	983	1454	2394	2064	3281	2608	
3	960	1070	1153	1630	2362	2144	

Figure 14. Summary of data from the sentence-picture verification experiment.

A TRUE DENIAL (TD) TRIAL WOULD LOOK LIKE THIS:

IT ISN'T TRUE THAT THE PLUS IS ABOVE THE STAR
*
+

AND THE ABSTRACT REPRESENTATION WOULD BE:

	NEG, [AFF, (PLUS, TOP)] (STAR, TOP)	FOR THE SENTENCE FOR THE PICTURE SET INDEX AT TRUE INDEX NOW = FALSE
BEGIN COMPARISON		
1ST PASS	-	
(START AGAIN)		
2ND PASS	- + +	INDEX NOW = TRUE
(START AGAIN)		
3RD PASS	+ + +	INDEX NOW = TRUE (ALL MATCH)

RESPONSE = FINAL INDEX = TRUE
TOTAL NUMBER OF COMPARISONS REQUIRED = 7

(+ = MATCH, - = MISMATCH. EACH PASS STARTS AT THE RIGHT
& MOVES LEFT. PRESS RETURN TO CONTINUE.)?

Figure 15. A trial from the demonstration of the Carpenter and Just (1975) model for the sentence-picture verification task.

they wish to receive at the beginning of the program. The data consist of percent correct and reaction times. They are output in a condition by block format, as shown in the class data summary of Figure 14.

As previously stated, the sentence-picture verification package also presents a demonstration of Carpenter and Just's (1975) processing model for the task. The demonstration shows how the sentence and picture are represented in each of the six conditions and how verification time is computed as a function of the number of comparisons required. A sample trial from the demonstration is shown in Figure 15. The demonstration also explains how the process of recoding can alter the sentence representation and thus yield somewhat different predictions for the verification times than the simple model.

Constructive Processes in Prose Comprehension

The software package for constructive processes in prose comprehension consists of three experiments. The first experiment is a replication of a study by Schallert (1976) on the effects of titles on biasing subjects' comprehension and memory for ambiguous prose. In the experimental session the subject reads a series of paragraphs containing ambiguous words. The following is an example of the paragraphs.

In the last days of August, we were all suffering from the unbearable heat. In a few short weeks, our daily job had turned from a game into hard labor. "All we need

now," said the manager in one of his discouraged moods, "is a strike." I listened to him silently but I could not help him. I hit a fly. "I suppose things could get even worse," he continued. "Our most valuable pitchers may crack in this heat. If only we had more fans, we would all feel better I'm sure. I wish our best man could come home. That certainly would improve everyone's morale, especially mine. O, well, I know a walk would cheer me up a little."

Each paragraph either is preceded by a title biasing the subject toward its more common interpretation (strong meaning) or its less common interpretation (weak meaning) or has no title. In this example the strong-meaning title is "Worries of a Baseball Team Manager," and the weak meaning title is "Worries of a Glassware Factory Manager."

After reading the paragraphs, the subject is asked a series of multiple-choice questions like the following.

What did the manager say was all they needed?

- (1) A worker quitting during the summer rush.
- (2) A batter letting a good pitch go by him.
- (3) A player dropping a ball hit to him.
- (4) A general walkout called by the employees' union.

This question assesses the student's understanding of the word "strike." Alternative 2 is consistent with the strong meaning of the paragraph, and Alternative 4 is consistent with the weak meaning. The other two alternatives are false.

The data summary, presented in Figure 16, presents the percent of responses consistent with a strong or weak meaning as a function of paragraph title condition (strong, weak, or none). As the data indicate, titles have a strong biasing effect on the construction of word meanings.

The second experiment in this package is designed to study the processing and representation of information that is only implicitly represented in a text. It is based on a study by Keenan and Kintsch (1974). The materials consist of short paragraphs that have both an explicit and an implicit version, such as:

Explicit: Gas leaking from a butane tank touched off an explosion. The explosion leveled a service station and a new home.

Implicit: Gas leaked from a butane tank. The explosion leveled a service station and a new home.

A given subject reads only one of the versions. Either

CON1 DATA SUMMARY		05/16/79		6 SUBJECTS	
AVERAGE PERCENT OF RESPONSES AS A FUNCTION OF BIASING CONTEXT.					
CONTEXT: TITLE TYPE					
	STRONG MEANING	WEAK MEANING	NO TITLE	TOTAL	
STRONG RESPONSE	50	17	58	42	
WEAK RESPONSE	9	64	28	33	

Figure 16. Data summary from the first experiment on constructive processes in prose comprehension.

CON2 DATA SUMMARY		16-MAY-79		
IMMEDIATE GROUP: 3 SUBJECTS			DELAYED GROUP: 3 SUBJECTS	
STATEMENT FORM	MEAN % CORRECT	MEAN LATENCY	MEAN % CORRECT	MEAN LATENCY
EXPLICIT	91.7	1730.7	100	3113
IMPLICIT	83.3	2104.3	91.7	3201.7
FILLER	64	2020.7	61	3309.7

Figure 17. Data summary from the second experiment on constructive processes in prose comprehension.

CON3 DATA SUMMARY										05/23/79	
10 SUBJECTS TOTAL											
FOR EACH COMMERCIAL TYPE, THERE ARE 10 CRITICAL STATEMENTS AND 5 EACH FALSE AND INDETERMINATE.											
GROUP 1 - I		3 SUBJECTS									
RESPONSE	CRITICAL			TEST SENTENCE TYPE						FALSE	
	T	F	I	T	F	I	T	F	I		
ASSERTED CLAIM	6.33	0	3	0	1.67	3.67	0	3.33	1		
IMPLIED CLAIM	5.67	0.33	4	0	1	3.67	0	3.67	1.67		
FOR GROUP 2, PRESS RETURN?											

FOR GROUP 2, PRESS RETURN?

Figure 18. Data summary from the third experiment on constructive processes in prose comprehension.

immediately after reading the paragraph or after reading all of the paragraphs (approximately a 10-min delay), the subject is asked to verify a statement that concerns the information that is left implicit in the implicit versions. The verification statement corresponding to the example above is "The gas leak started the explosion."

As shown in Figure 17, questions concerning implicit information require more time to verify than questions concerning explicit information in the immediate condition, but this difference dissipates over a delay. This pattern of results is taken as evidence that implicit information is actually processed at the time of reading just like explicit information but that explicit information has a short-lived surface-form representation that facilitates question answering in the immediate condition.

The third experiment is also concerned with inferences. It is a replication of a study on advertising by Harris (1977). Specifically, it addresses the problem of instructing people to differentiate between the actual assertions and the mere implications of a commercial. The experiment manipulates the type of instruction the subject receives and the delay between the commercial and the presentation of questions about the commercial. The following is an example of one of the commercials and its test question.

Everyone occasionally has trouble falling asleep. When you need some help to make you feel drowsy, take Knockout sleep capsules. In a survey of 500 doctors, over half reported that they recommend the ingredients in Knockout capsules. Take their advice. Buy Knockout.

Question: Do most doctors recommend Knockout capsules to help you sleep?

The commercial states only that doctors "recommend the ingredients in Knockout capsules," not the capsules

themselves. Thus, the answer to the question should be "indeterminate," not "true." Nonetheless, subjects often respond "true" to this statement, even when they are given extensive instructions warning them of such pitfalls. The data summary presented in Figure 18 shows that subjects respond "true" to claims that were only implied about as often as they respond "true" to actual assertions.

IMPLEMENTING THE SYSTEM

Because none of the experiments require tachistoscopic presentation rates, single-user or real-time computers are not required. Rather, the programs are designed to operate on a campus timesharing system. We originally developed the programs for a Burroughs 6700; we now also have versions for a DEC PDP-11/70 with a UNIX operating system. The main requirement is that the computer should have fairly quick turn-around time. Otherwise, when the machine is slow to respond, experiments become lengthy, the students' attention lapses, and, consequently, the data can become unreliable.

The programs are written in a fairly low level of BASIC. Dialect-specific features were avoided as much as possible, but in those cases where they were used, they are noted as such.

There are no constraints on the type of display device required. The programs do not require upper- and lowercase symbols, cursor addressability, or screen erase. Nor does the CRT need to be able to receive information at high baud rates. The programs are written to run at 300 baud, but they can be easily modified to run at 110 baud.

Currently, all but two of the programs require less than 12K of memory (when the REMARK statements are deleted).

Five of the 10 programs measure reaction times: the two levels-of-processing experiments, semantic memory, sentence-picture verification, and the experiment on explicit and implicit versions of text. These programs were written so that response times are measured locally using a Polytronics universal response timer. This device is inserted in the line between the terminal and the modem for remote system operation. We use this local timing device because it provides more accurate timing than would calling the computer's clock when there are a large number of users on the system. If you are using a campus timesharing system, you may want to use such a local timing device; unfortunately, the Polytronics timers cost approximately \$650 apiece. If you have a dedicated machine, timing should be no problem, as long as the BASIC dialect has both a timing function and the ability to accept a single character from the keyboard without requiring a carriage return. This latter capability is not required when using the Polytronics timers because they automatically send a carriage return to the program along with the response time.

EVALUATION

Student response to this program is highly enthusiastic. The computerized laboratory course seems to increase both motivation and understanding. Evidence for increased motivation lies in the fact that students are willing to spend much more time on this course than on any of their others. In fact, enthusiasm is so high that students often run roommates, friends, and even parents in the experiments. Many go on to take computer programming courses. Evidence for increased understanding comes from our observations that students become very good at thinking at a theoretical level. As a result of this increase in understanding theories, the students have developed modified versions of the programs to test their own ideas.

Teachers' reactions are equally enthusiastic. They find the computerized lab to be more convenient than the standard lab, and more important, that they can cover more material and in much greater depth than before.

REFERENCES

- BEWLEY, W. L. *Cognitive psychology: A computer-oriented laboratory manual*. Hanover, N.H: Project COMPUTE, 1974.
- CARPENTER, P. A., & JUST, M. A. Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review*, 1975, **82**, 45-73.
- COLLINS, A. M., & QUILLIAN, M. R. How to make a language user. In E. Tulving and W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.
- CRAIK, F. I. M., & TULVING, E. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 1975, **1**, 268-294.
- HARRIS, R. J. Comprehension of pragmatic implications in advertising. *Journal of Applied Psychology*, 1977, **62**, 603-608.
- KEENAN, J. M., & KINTSCH, W. The identification of explicitly and implicitly presented information. In W. Kintsch (Ed.), *The representation of meaning in memory*. Hillsdale, N.J: Erlbaum, 1974.
- LEVY, C. M., FISCHLER, I., & GRIGGS, R. A. *Laboratory in cognition and perception*. Iowa City, Iowa: CONDUIT, 1979.
- RIPS, L. J., SHOBN, E. J., & SMITH, E. E. Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 1-20.
- SCHALLERT, D. L. Improving memory for prose: The relationship between depth of processing and context. *Journal of Verbal Learning and Verbal Behavior*, 1976, **15**, 621-632.
- SMITH, E. E., SHOBN, E. J., & RIPS, L. J. Structure and process in semantic memory: A feature model for semantic decisions. *Psychological Review*, 1974, **81**, 214-241.
- THOMSON, D. M., & TULVING, E. Associative encoding and retrieval: Weak and strong cues. *Journal of Experimental Psychology*, 1970, **86**, 255-262.
- TULVING, E., & THOMSON, D. M. Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 1973, **80**, 352-373.