TESOL QUARTERLY, Vol. 22, No. 2, June 1988

The Content of L1 and L2 Students' Lecture Notes and Its Relation to Test Performance

PATRICIA DUNKEL
The Pennsylvania State University

The lecture notes of 129 L1 and L2 students were examined in terms of five indexes for the content of notes: (a) the total number of words and notations, (b) the number of information units, (c) the number of test questions answerable from the notes, (d) the completeness of the notes, and (e) the efficiency of the notes. Three stepwise multiple regression analyses of the data were conducted to identify which of the indexes predicted achievement on the postlecture quiz covering lecture concepts and details for (a) the L1 note takers and the L2 note takers as a group, (b) the L1 note takers, and (c) the L2 note takers. Results suggest that L1/L2 note takers who scored high on the recognition measure compacted a large amount of the lecture material into propositional pieces of information and detected and recorded information that subsequently appeared on the postlecture quiz. Implications of the findings for pedagogy and note-taking research are suggested.

Taking lecture notes is widely accepted as a useful strategy for augmenting student attention and retention of academic discourse. Note taking is intuitively appealing to both learner and lecturer alike and is generally viewed as one class of mathemagenic activity (Rothkopf, 1970) that facilitates the process of learning and retaining lecture material.

The facilitative effect of note taking is thought to derive from one or both of two postulated functions: (a) the encoding function and (b) the external storage function (Di Vesta & Gray, 1972). Encoding (the process) supposedly aids learning and retention by activating attentional mechanisms and engaging the learner's cognitive processes of coding, integrating, synthesizing, and transforming the aurally received input into a personally meaningful form. The external storage function of note taking is seen as important because the notes (the product) serve as an external repository of information

that permits later revision and review to stimulate information recall.

The intuitive belief held by educators that note taking promotes learning has spawned numerous study skills programs to teach students how to improve note-taking skills. For those fluent in English, books, pamphlets, and "how to study" guides offer tips on taking notes during L1 lectures (e.g., Armstrong, 1956; Christiansen & Vargata, 1975; Laycock & Russell, 1941; Maddox, 1963; Pauk, 1974; Raygor, 1970; Sotiriou, 1984; Yarington, 1977). For those not yet proficient in English, ESL/EFL writers have authored programs to develop L2 listening and note-taking skills (e.g., Dunkel & Pialorsi, 1982; Mason, 1983; Otto, 1979; Plaister, 1976; Ruetten, 1986; Sims & Petersen, 1981; So, 1974; Young & Fitzgerald, 1982).

Although authors have attempted with various approaches to fashion better note takers out of L1 and L2 learners, little is in fact known about the auditory-cognitive and functional skills that comprise effective note taking. Furthermore, the empirical relations between the quantitative and qualitative aspects of students' lecture notes and the comprehension and retention of lecture information are not well known. Although several studies have examined the relation of the act of note taking to test performance (see Dunkel, 1985; Hartley & Davies, 1978), the majority have compared test scores of note takers with scores of those who do not take notes. These studies have largely ignored the relationship between the content of subjects' notes and their test performance.

To date, primarily L1 researchers have taken up the investigation of this relationship; the L2 literature has been silent concerning the connection between the composition of student notes and the comprehension and retention of L2 lecture information. As a result, L2 and L1 researchers and educators alike are unclear as to (a) what actually constitutes "good notes" and (b) whether certain aspects of the notes engender or indicate successful information processing. There is little L1 (and even less L2) empirical data that provide answers to the following questions concerning the content and utility of students' notes: Are quantity notes "quality" notes? Are terse notes "efficient" notes? Are those notes that contain symbols, abbreviations, and idiosyncratic mnemonics "ideal" notes? Are those notes that contain possible answers to test questions "effective" notes? Are there, in other words, identifiable, essential (perhaps universal) elements that comprise quality notes for all those who transcribe spoken academic discourse into note form?

Answers to these questions should furnish vital information about a skill that supports the academic endeavors of L1 and L2 college

students, a skill that is being increasingly taught in L2 classrooms in the attempt to help prepare nonnative speakers of English to succeed academically in English-speaking lecture halls. However, until we begin to probe the relationship between the content of L2 students' notes and lecture-information learning and recall, we will be unable to pinpoint the functional skills that comprise effective note taking in order to devise effective curricula to teach those skills to L2 learners. This article reports the findings of a study that examined the content of lecture notes taken by L1 and L2 students in order to probe the relationship between the content of the students' notes and their retention of concepts and details presented by the English-speaking lecturer.

EXAMINING THE CONTENT OF LECTURE NOTES TO FIND INDEXES OF QUALITY

More than 60 years ago, Crawford (1925a) attempted to relate the content of students' lecture notes to test performance. Finding a significant positive correlation between the number of lecture points recorded in students' notes and their recall of lecture points on an essay-type quiz, Crawford concluded that taking notes on a lecture point does not guarantee its being recalled at the time of the quiz but that failure to do so decreases the chances of recall. Aiken, Thomas, and Shennum (1975) also detected a positive relationship between the content of lecture notes and test performance. Lecture material recorded in students' notes was twice as likely to be recalled as material not taken down.

In a naturalistic investigation by Palkovitz and Lore (1980), 82% of the correct test responses were associated with the presence of relevant information in the notes. Similarly, an examination of the content of notes by Baker and Lombardi (1985) revealed that students rarely answered postlecture test questions incorrectly if the information was present in their notes.

Several L1 note-taking researchers have attempted to determine how "quality of notes" relates to postlecture recognition or recall performance. Howe (1974) reported a significant positive correlation between "efficiency" of note taking (defined as communicating the maximum number of "meaningful content units" in the minimum number of words) and subsequent recall success. Quality of notes taken was defined by Fisher and Harris (1974b) as the number of ideas from the lecture included in the notes. For 18 sets of notes taken, a significant positive correlation for quality of notes and recognition scores was obtained.

Hartley and Marshall (1974) assessed the quality of students' notes on the basis of (a) the number of words that students noted down, (b) the number of test questions that appeared answerable from the students' notes alone (i.e., whether the answer appeared explicitly in the notes), and (c) the number of information units transcribed. The latter criterion, suggested by Hartley and Cameron (1967) and Howe (1970b), consisted of dividing up the lecture transcript into information units and awarding a mark for each unit of information contained in the student notes, even if there was only a word of reference to it. "Good" note takers (those who transcribed the greatest number of information units and test information) outperformed "poor" note takers (those who recorded the fewest units and test answers) on a test of general comprehension, recall of factual details, and recognition of conclusions that could be drawn from the lecture material.

Norton's (1981) study of the effects of note taking on long-term recall similarly relied on the criteria of (a) the total number of words written down by each student and (b) the amount of information in the notes that would enable the student to answer the test questions correctly. Norton uncovered a statistically significant correlation between the number of words written down and test scores made 3 weeks following the lecture. The correlation between the amount of test information in the notes and test scores was not significant. There also appeared to be at least a relationship in the Norton study between the "quantity" of notes taken and end-of-course grade.

Examining the notes taken by college students actually enrolled in courses, Locke (1977) found a significant positive correlation between completeness of lecture notes and course grades. Completeness was defined as the percentage of total thought units in the lecture appearing in each student's notes. A thought unit equaled the name of a person, place, or concept; the definition of a concept; or some other pertinent fact (e.g., a date, an example).

In an investigation of student variables related to lecture note taking, Nye (1978) analyzed the (a) content, (b) layout, and (c) legibility of students' notes. With respect to content, 48% of the main and minor points made by the lecturer were recorded, on the average. Nye found small but significant positive correlations between the final course grades and (a) the number of words in the notes and (b) the number of minor points written down. The correlation between the final grade and the main points in the notes was not significant. Little variation in writing neatness and legibility or layout of the notes was evident.

One surprising finding in Nye's (1978) study involved gender differences in the notes of students. Corroborating the findings of

Hartley (1976), Hartley and Cameron (1967), Hartley and Fuller (1971), and Maddox and Hoole (1975), Nye found that the female students took more complete notes than did the males; however, the female subjects did not achieve better test scores. In the Fisher and Harris (1974a, 1974b) studies, females scored significantly better than males on the recall test, but these researchers failed to report whether or not the female students took more notes than the males.

If gender differences between note takers exist, it is possible that there may be other differences attributable to ethno-cultural background or language proficiency among note takers from such diverse regions as the United States, the Middle East, the Far East, and Latin America. Such differences might provide important information concerning the note-taking strategies employed by these students, as well as the relationship between the content of L1 and L2 students' notes and their comprehension and retention of English-language lecture material.

PURPOSE OF THE STUDY

Almost a decade ago, Hartley and Davies (1978) remarked that "no investigator, to our knowledge, has commented on the differences (if any) between notes taken by students in different cultures" (p. 214). To date, this issue remains largely unexplored. Thus, in light of the dearth of research concerning cross-cultural differences in students' notes, the complete lack of research on the content of L2 students' notes, and the increased pedagogical focus placed on training L2 students to develop listening and note-taking skills in English, a study was conducted to (a) determine whether cross-cultural differences are evident in the notes taken by L1 and L2 students, (b) develop a set of indexes for content of notes that relate to (and predict) postlecture retention performance for L1 and L2 note takers, and (c) identify the relative weight of the indexes in accounting for the variance in performance on the postlecture retention measure for L1 and L2 note takers.

METHOD

Subjects

A total of 129 undergraduate students in intact classes of Freshman English Composition at the University of Arizona were randomly selected to participate in the study. Of these, 66 were native speakers of English, and 63 were nonnative speakers of English (TOEFL score > 450). The ethno-cultural composition of the international student sample is displayed in Table 1.

TABLE 1

Ethno-Cultural Background of the L2 Subjects

Country	Number	Country	Number
Far East Asia		South Asia	
Hong Kong/Taiwan/ China	6	Pakistan	2
		India	1
Indonesia	2		
Japan	3	Africa	
Malaysia	14	Botswana	2
Thailand	1	Lesotho	1
Vietnam	2		
Middle East		Europe	
Jordan	1	Italy	1
Iran	1	Norway	1
Kuwait	1	Yugoslavia	1
Lebanon	1		
Oman	3	Latin America	
Qatar	3	Brazil	1
Saudi Arabia	2	Mexico	6
Syria	2	Venezuela	2
United Arab Emirates	1	Caribbean	
Yemen	1	Trinidad	1

Materials

The instructional material consisted of a videotaped lecture on the evolution of the Egyptian pyramid structure. The length of the lecture was 22 minutes, 51.8 seconds, and it contained 2,672 words, including verbal fillers, false starts, and redundancies. The rate of presentation was 118.6 words per minute.

The postlecture retention test consisted of 15 multiple-choice items covering conceptual information in the lecture and 15 multiple-choice questions covering facts and details presented in the lecture.

Procedures

Subjects were asked to view the videotaped presentation and concurrently to take notes on the information contained in the lecture. During presentation of the lecture, the experimenter wrote any notations made by the lecturer on the classroom chalkboard so that all subjects could clearly see the notations. The lecturer wrote the dates (with the exception of one) and the names of the people

and places referred to as the lecture unfolded. Subjects were informed at the start of the experiment that a test would immediately follow presentation of the lecture and that they would not have the opportunity to review the notes taken. (This was done to probe the process—the act—of note taking on the comprehension and retention of lecture material.) When the lecture ended, the retention test was administered immediately following collection of the notes.

One point was awarded for a correct response to each multiple-choice question. A maximum score of 15 could be achieved on the test of concept information and 15 on the test of detail/factual information contained in the lecture. The students' notes were scored in the following manner:

1. The *total-number-of-words score* equaled the total number of words, symbols, abbreviations, and illustrations pertaining to the information presented by the lecturer (Pearson's product-moment correlation of interrater reliability for counting total number of words = .97).
2. The *information-units count* equaled the total number of information units contained in the student notes (interrater reliability = .83). The lecture contained 278 information units, which the experimenter predetermined by using Anderson's (1980) propositional definition of an information unit as equaling the smallest unit of knowledge that can stand as a separate assertion and that can be judged true or false.
3. The *test-answerability score* equaled the number of test questions answerable from the subject's notes (interrater reliability = .98).
4. The *completeness score* equaled the total number of all possible information units contained in the lecture (278), divided by the total number of all information units written in a student's notes.
5. The *efficiency ratio* was equivalent to the number of information units, divided by the total number of words (i.e., words, abbreviations, symbols, and illustrations) found in the notes. (The completeness score and the efficiency ratio were derived from the total-number-of-words and the information-units scores.)

The experimenter examined the notes and tabulated the data; a colleague at the Center for English as a Second Language at the University of Arizona served as the second rater of the notes. She was trained to count the total number of words, the information units, and the test-answerability score in 40 sets of L1 notes and 40

sets of L2 notes drawn at random from the entire sample. Scores obtained by the experimenter and the second rater on these sets of notes were then examined to obtain Pearson's product-moment correlation coefficients for the three indexes. The experimenter and her colleague worked independently of one another in scoring the notes.

Statistical Analysis

A .05 alpha level of significance was set for all statistical tests. Stepwise multiple regression analyses of the data were carried out to identify which of the five content indexes predicted achievement on the concept and detail measures for (a) L1 and L2 note-taking groups combined, (b) L1 note takers, and (c) L2 note takers. The data were analyzed with the main-frame SPSS^{x} Batch system.

The stepwise regression technique is typically used to develop a subset of independent variables (the content indexes) that may prove useful in predicting the dependent variable (information retention) and to eliminate those independent variables that do not provide predictive power relative to achievement on the dependent measure (Tabachnick & Fidell, 1983). One of the principal reasons for using the stepwise regression analysis was to determine the relative contribution of the various indexes of note content to performance on the postlecture retention measure.

A major limitation of regression analysis concerns inference of causal relationship. As Tabachnick and Fidell (1983) point out,

> Demonstration of causality is a logical and experimental rather than a statistical problem. Statistics are helpful only in demonstrating that relationships occur reliably. A high multiple correlation indicates that a lot of variability is shared between one variable and a set of others, but not that the variables are causally related. (p. 91)

Shared variability might stem from currently unmeasured variables, and so it is possible that other learner characteristics, situational factors, or task requirements not manipulated in the present study were influencing the correlations found between the note-taking indexes and lecture retention, factors such as language proficiency, organization of the notes, note-taking pretraining, and background knowledge of the lecture content. Further research on note-taking quality might attempt manipulation of some of these factors to investigate the differential influence of their manipulation on the content and quality of lecture notes.

RESULTS

L1 and L2 Subjects Combined

Table 2 displays the correlations between the variables (R and R^2), the standardized regression coefficients (*beta*), the standard errors of *beta*, the degrees of freedom, the significant F ratios, and the semipartial correlations (sr^2). Table 3 displays the means and standard deviations for the stepwise regression analyses.

For L2 and L1 subjects combined, test answerability proved to be a predictor of achievement for recognition of lecture concepts. A stepwise regression indicated that after Step 1, with test answerability in the equation, $R^2 = .26$, $F\ (1, 128) = 45.43$, $p < .05$. After Step 2, with total number of words in the equation, prediction of concept recall improved, $R^2 = .29$, $F\ (2, 127) = 26.76$, $p < .05$. Addition of total number of words resulted in a significant, albeit slight, increment of R^2; however, the efficiency ratio, the completeness score, and the information-units count did not improve prediction of performance. It is interesting to note that addition of total number of words was inversely related to the test-answerability score, as evidenced by the negative *beta* weight of $-.21$ attached to the total-number-of-words index. In all ensuing regression analyses, each time total number of words entered the prediction equation, it carried a negative *beta* weight for that variable.

Performance in recalling lecture details could be predicted from the test-answerability score for both American and international students combined, $R^2 = .06$, $F\ (1, 128) = 8.48$, $p < .05$, but no other index of quality significantly predicted recall of lecture details by the entire sample.

L2 Subjects

After Step 1, with the information-units count in the equation, $R^2 = .19$, $F\ (1, 62) = 14.70$, $p < .05$. After Step 2, with total number of words added to predictability of lecture concept recall, $R^2 = .28$, $F\ (2, 61) = 12.10$, $p < .05$. Thus, inclusion of total number of words with the information-units count resulted in a significant increment in R^2. Again, however, total number of words was inversely related to the information-units count of test performance for lecture concepts. No other indexes added significantly to the explainable variance in the criterion measure.

For the L2 performance on the detail measure, after Step 1, with information units in the equation, $R^2 = .13$, $F\ (1, 62) = 9.98$, $p < .05$. After Step 2, with the completeness score added to prediction of recall of lecture details, $R^2 = .21$, $F\ (2, 61) = 8.53$,

TABLE 2

Stepwise Regression of Quality Indexes

Group	*R*	R^2	*beta*	*SE B*	*df*	*F*	sr^2
L1 + L2 combined							
Concept							
TA	.51	.26	.51	.04	1, 128	45.43*	.26
TW	.54	.29	−.21	.04	2, 127	26.76*	.03
Detail							
TA	.24	.06	.24	.04	1, 128	8.48*	.06
L2							
Concept							
IU	.43	.19	.43	.01	1, 62	14.70*	.19
TW	.53	.28	−.34	.01	2, 61	12.10*	.09
Detail							
IU	.37	.13	.37	.01	1, 62	9.98*	.13
CS	.46	.21	−.31	.01	2, 61	8.53*	.08
L1							
Concept							
TA	.35	.12	.35	.06	1, 65	9.29*	.12
TW	.45	.20	−.40	.08	2, 64	8.46*	.08
Detail	—	—	—	—	—	—	—

Note: If no value displayed, no predictor proved significant. *beta* = standardized regression coefficient, *SE B* = standard error of *beta*; sr^2 = semipartial correlation; TA = test-answerability score; TW = total number of words in notes; IU = information-units count; CS = completeness score.

*$p < .05$.

TABLE 3

Means and Standard Deviations for Postlecture Recognition Performance for L1/L2 Note Takers

	Concept		Detail	
Group	*M*	*SD*	*M*	*SD*
L1 + L2	10.33	2.68	8.57	2.51
L2	8.81	2.64	8.02	2.52
L1	11.52	1.89	8.74	2.13

$p < .05$. A negative *beta* weight of $-.31$ was attached to the completeness score. The test-answerability score, the efficiency ratio, and total number of words did not increase the predictability of test performance for L2 note takers.

L1 Subjects

After Step 1, with test answerability in the equation, $R^2 = .12$, $F(1, 65) = 9.29$, $p < .05$. After Step 2, with total number of words added to predictability of lecture concept recall, $R^2 = .20$, $F(2, 64) = 8.46$, $p < .05$. Addition of this variable increased R^2 significantly, but the index carried a negative *beta* weight of $-.40$.

None of the five indexes significantly predicted the performance of L1 subjects on recall of lecture details.

DISCUSSION

The Issues of Quantity of Notes and Cultural Differences

The present study, a preliminary investigation of L1/L2 students' notes and their relationship to immediate retention of lecture concepts and details, was designed to determine whether a set of five note-taking criteria or individual components of the set were related to postlecture achievement. If so, it was thought that these criteria might comprise some general and essential elements of effective note taking, as previous research has suggested.

Although we must proceed cautiously in interpreting the results of this study due to the exploratory nature of the investigation, the data suggest that the tactic of "writing down as much as possible" during a lecture may not result in effective encoding of the lecture for L1 or L2 note takers. The negative *beta* weights attached to the total-number-of-words index indicate that test achievement was not directly related to the quantity of notes taken but rather to (a)

terseness of notations (embodied in the information units) and (b) inclusion of potential test information (embodied in the test-answerability index) for all the L1/L2 note takers as a group.

In general, the analysis of the data indicates that the effective L1 and L2 note takers were those who compacted large amounts of spoken discourse into propositional-type information units; transcribed content words (e.g., names, dates, statistics) using abbreviations, symbols, and a limited number of structure words; and detected and wrote down information that subsequently appeared on the postlecture quiz.

Terseness of note taking (involving the recording of lecture propositions) rather than mere quantity seems to be an essential ingredient of effective L1/L2 note taking as it relates to immediate retention of lecture information. Those L2 note takers who performed well on the postlecture quiz managed to compact lecture material into propositional information units by using abbreviations, symbols, and drawings, thereby limiting the total number of words in proportion to the number of propositions (or information units) transcribed. Those L2 note takers who did not perform as well on the quiz wrote down numerous structure words (e.g., articles and prepositions) so that their notes contained fewer information units overall but a larger quantity of words or notations. It appears that some students need practice in detecting and recording the information-carrying words while simultaneously ignoring (for purposes of note taking) structure words and other syntactic elements (e.g., past tense markers) that do not add to the informational load but that increase the total number of notations placed in the notes.

It may therefore aid those learners who are attempting to build L2 listening and note-taking skills and those nonnative-speaking students who face the task of taking notes on lecture material in English if their ESL instructors or their college professors were to provide them with advance organizers or "skeleton notes" containing the major points of information in the lecture. The skeleton notes should highlight those critical pieces of information that should be recorded in the notes and should free students from the need to scribble frantically to record the material on paper. If less swamped with the mechanics of concurrent note taking, students might be better able to concentrate on understanding the content of the lecture and to perceive the organization and structure of the information conveyed.

It also seems evident that although taking dictation may help students to develop listening comprehension at the beginning level of L2 learning, it does not seem appropriate to require advanced-level

L2 learners to attempt verbatim transcription of spoken discourse. Dictation at this level may result in the transcription of structure words in the discourse and may engender an attempt to note down as many words as possible in the flow of the lecturer's discourse. Since it was discerned that the quantity of notes was not related per se to retention achievement, it appears that greater emphasis should be placed on teaching advanced-level L2 note takers to synthesize segments of discourse into key phrases or sentences. Advanced-level ESL teachers may even be encouraged to provide opportunity for "spaced note taking" rather than "concurrent note taking"; they might, for example, have students listen to a half-minute segment of a minilecture, stop the lecture, and then ask the listeners to synthesize the information into key sentences and phrases or to outline the macrostructure of the message.

Since the indexes of note quality did not function in similar fashion across the L1/L2 groups to predict achievement (see Table 2), the notion that there is a single, unitary (or universal) note-taking method that is effective for all groups of students does not find support. It may be that particular groups of L2 learners need instructional curricula that tap lecture-information-processing strengths and address weaknesses related to the language, cultural, and educational background of those learners. It is therefore unlikely that any single note-taking program can address the note-taking needs of students from diverse ethno-cultural backgrounds. It may also be of questionable instructional value to present all L2 students (or for that matter, L1 students) with a general and single "model" of effective notes to emulate. Further research concerning the note-taking strategies employed by various ethno-cultural groups within the nonnative-speaker population is called for to test this assumption.

The finding that note-taking terseness was correlated with recall performance (as evidenced by the negative *beta* weights attached to the total-number-of-words score each time it entered the regression equation and by the repeated significance of the information-units count) goes counter to the findings of studies conducted by Hartley and Marshall (1974) and Kiewra and Fletcher (1984), who found that note-taking "quantity" equaled note-taking "quality" in terms of test performance. In the present study, the total number of notations found in the notes actually proved to be inversely related to the test-answerability score and the information-units count (two statistically significant, albeit weak, predictors of

concept and detail recognition performance) for both the L1 and L2 note takers. Additional research should probe the "quantity versus quality" issue of effective note taking.

The Issues of Test Wiseness and Note-Taking Efficiency

Test wiseness may have played a role in postlecture test achievement, since the ability to record the test answers and the critical information units in as few words as possible was correlated positively with test performance for the sample as a whole and for the L1 subjects as a group. However, it is interesting to note that the L2 note takers did not use a "test-wise" note-taking strategy; that is, they did not tend to record the information in their notes that subsequently appeared on the postlecture quiz.

L2 note takers may need to practice (a) taking the typical limited-response, multiple-choice test that is often used to assess the comprehension and retention of lecture material and (b) keying in on information that might be tested in a traditional limited-response format quiz. Test-taking as well as note-taking instruction may therefore need to be incorporated into instructional programs designed to prepare international students to achieve academically in U.S. universities.

A major concern of L1 note-taking researchers involves development of a quality-of-notes measure. Calculation of the ratio between the test-answerability score and the total-number-of-words score seems to be one potential index of note quality for L1 note takers. Using this measure, for example, a subject who records in 250 words the answers to 25 items that appear on a postlecture quiz would accrue a lower quality-of-notes index than would a subject who records the same amount of information in 100 words. This ratio may not, however, function as a revealing measure of the effectiveness or quality of L2 notes because the test-answerability score did not correlate highly with recognition achievement for the L2 note takers as a subgroup. Researchers should continue to seek an L2 quality-of-notes measure.

The efficiency-of-notes index (the number of information units divided by the total number of words) did not predict recall of either lecture concepts or details for any of the experimental groups. This finding does not corroborate the results of Howe's (1970a) analysis, which indicated a positive correlation between this index and recall; however, neither Fisher and Harris (1973) nor Kiewra (1984) documented a positive correlation between note-taking efficiency and achievement.

Although the notion of efficient notes is intuitively appealing, it must be recognized that the efficiency index, as it has been used in past research, may in actuality have little practical meaning. As Kiewra (1984) points out, one note taker may record only 2 important points in a total of 4 words, whereas another note taker may list 30 critical points in 120 words. The respective efficiency ratings of .50 and .25 are certainly unlikely predictors of these students' note-taking achievement. Investigators and note-taking teachers need to be cognizant of the kind of information recorded in notes, not just the amount of information recorded.

LIMITATIONS OF THE STUDY

The Omission of a Review-of-Notes Condition

Several researchers have concluded that the encoding benefit of note taking actually accrues from having the opportunity to review notes and not from the mere *act* of note taking itself (Annis & Davis, 1975; Crawford, 1925b; Freyberg, 1956). Incorporating a review of notes into the research design might have provided more, and perhaps different, information on the encoding effect of individual subjects' note-taking strategies as they relate to retention and test performance. In addition, testing subjects immediately following presentation of the lecture may have vitiated the effect of the encoding function of note taking; recall of the lecture material was doubtlessly high, since no additional learning experiences intervened between the learning of the lecture information and the testing of that information. Such a testing situation is not congruent with the actual university lecture environment. Replication of the study with a review-of-notes condition and delayed testing is desirable.

Sample Size

The American subsample was comprised of only 66 native speakers of English, whereas the international subsample contained only 63 nonnative speakers of English, so the experimental variance suggested by R^2 could be misleading. R^2 is subject to rather dramatic variation when sample size is small. Caution must therefore be exercised in interpreting the R^2 for the subsamples of the American and international note takers.

Multicollinearity and Stepwise Multiple Regression Analysis

Interpretation of multiple regression is most accurate when the explanatory variables comprising the model are themselves

uncorrelated. An investigator wishes, in other words, to find a set of independent variables (the note-taking indexes) that are highly correlated with the dependent variable (lecture concepts and details recognized) but not correlated with each other. When strong intercorrelationships among the variables exist, it is difficult if not impossible to assess the unequal effects individual explanatory variables (the note-taking indexes) have on the response variable (performance on the concept and detail tests). As Berenson, Levine, and Goldstein (1983) note, strong intercorrelations among the explanatory variables "can result in such highly unstable regression coefficients that their values will be subject to dramatic alterations as a result of additions or deletions of variables or small changes in data points" (p. 414).

In the present study, multicollinearity, the term applied to strong interrelationships among predictor variables, was evidenced in the moderate to high Pearson product-moment correlation coefficients for some of the note-taking indexes (see Table 4). Hence, caution is advised when interpreting the unique effect of each note-taking index to lecture learning for the L1 and L2 note takers. Since several of the index variables proved to be moderately to highly intercorrelated (e.g., the information-units count with the test-answerability score; the test-answerability score with the completeness score), the *beta* weights may be unstable from sample to sample. It may be the case that certain indexes examined in the study (e.g., the completeness score) are in fact surrogates for others (e.g., the information-units count) and have little or no effect as predictors themselves. Because the indexes were multicollinear, identifying the measure of contribution of each of the note-taking indexes to achievement on the retention measure proves somewhat elusive; the exploratory nature of the investigation is thereby underscored.

TABLE 4

Intercorrelations Among Note-Taking Indexes

	TW	IU	TA	CS	ER
TW	1.000	.735*	.634*	.735*	− .404*
IU		1.000	.814*	1.000*	.272*
TA			1.000	.814*	.211*
CS				1.000	.227*
ER					1.000

Note: TW = total number of words; IU = information-units count; TA = test-answerability score; CS = completeness score; ER = efficiency ratio.

*$p \leq .01$.

It is anticipated that researchers taking note of the multicollinearity of the experimental variables will seek to identify and select variables other than those identified in the present study to analyze. The search for predictors of note-taking quality might, for example, include investigation of the relationship between memory performance and (a) the organization of the notes, (b) the use of idiosyncratic mnemonics in the notes, (c) the number of content versus structure words present in the notes, (d) the number of errors of omission and commission in the notes (for example, one L2 note taker from the Far East made an error of commission in transcribing the lecturer's reference to "grave robbers" as "great lovers"), (e) additional learner variables (e.g., gender, ethno-cultural affiliation, major field of study, or academic achievement). Continued investigation of the content of student notes may uncover additional indexes of note-taking quality that more accurately reflect the success or failure of lecture-information processing for both L1 and L2 note takers.

Considering Attitude, Motivation, and Background Knowledge

When queried about their note-taking practices, 25% of a sample of 64 University of Kiel students indicated that their note-taking practices varied according to the lecturer, the relevance of the subject matter, and their "mood" (Hartley & Davies, 1978). A subject's positive or negative attitude toward note taking, in general, also influences note taking. Norton (1981), for example, found significant intercorrelations between attitudes toward note taking and (a) the number of words noted down and (b) subsequent use of the notes.

Several attitudinal and motivational factors may have contributed to the small, albeit significant, correlations between note indexes and test performance in the present study. A perceived lack of content and task meaningfulness, as well as the simplicity of lecture information for L1 subjects, may have attenuated the strength of relationship between the note-taking indexes and test performance. Ladas (1980) points out that if the lecture material is so lacking in difficulty that students have the aptitude to master it completely without the help of any treatment, such as note taking (e.g., "Dick and Jane" stories for college freshmen), then no treatment effect will be discerned. Such a "Dick and Jane" phenomenon may have occurred in this experiment for the American students, since no attempt was made to determine the level of easiness of the lecture for the L1 learners (or difficulty for the L2 learners) and since the lecture was selected from a series of videotaped lectures designed

to give note-taking practice to preuniversity foreign students (Dunkel & Pialorsi, 1982).

In addition, motivation to take notes and learn the lecture information may have been low, since the students did not have to attend to the lecture or engage in the note-taking task to acquire information crucial to the content of their course or to success on an actual course examination. Care should be taken in future note-taking research using American and international student samples to ensure that the discourse and content of the lecture are of a sufficient level of sophistication and interest to challenge both groups of college students to learn the information and that there is a valid reason (other than willingness to cooperate with the experimenter) for them to listen and take notes on the lecture.

The topic of the lecture may have presented another problem in that some students (e.g., those from the Middle East) may have had an advantage over those from other ethno-cultural groups. They may, in other words, have possessed content schemata (Carrell, 1987) about the pyramids that aided and abetted their comprehension and recall of lecture material. Future investigations into the impact of lecture note taking across cultures should control for subjects' background knowledge of the lecture topic.

Edumetric Issues

Carver (cited in Ladas, 1980) claims that the ideal test for purposes of research about learning measures progress within individual gains. Ladas further points out that edumetric tests are designed to elicit maximum treatment differences and are maximally sensitive to the results of instruction (in this case, note-taking treatment). They are refined instruments that have been subjected to item analysis to identify and eliminate nonfunctional items with extreme difficulty or low item-discrimination indexes.

An item analysis conducted during a pilot testing of the instrument used in the present study indicated a low reliability coefficient for the achievement measure (Kuder-Richardson Formula 20 = .31); however, since the data were collected on only 16 subjects and the overall reliability of the 30-item instrument was .71, it was decided not to eliminate or redesign any of the concept or detail items. An item analysis of subject responses in the present study revealed relatively low reliability indexes for the concept and detail measures: For L2 and L1 subjects, the Kuder-Richardson Formula 20 for the concept items equaled .59 and .42, respectively; for the detail items, it equaled .47 for L2 subjects and only .37 for L1 subjects.

It is quite possible that the criterion measures yielded spurious correlations between the note-taking indexes and the achievement measures because of the low reliability of the test instruments. Future investigations should include pilot testing of the criterion measure and item revision, if necessary, before the experiment is conducted.

The Laboratory-Like Conditions of the Experiment

A naturalistic investigation of the notes students take during actual classroom lectures and their relationship to performance on actual classroom exams is sorely needed. Some degree of control will be sacrificed in such a study; the researcher will not know, for instance, how much time students spend studying their notes before the exam. However, as Baker and Lombardi (1985) contend, the advantages of conducting an ecological study should outweigh the limitations.

Examination of notes gathered in a naturalistic experiment (such as those conducted by Baker & Lombardi, 1985; Locke, 1977; Norton, 1981; Nye, 1978) might reveal other indexes of note content that relate to, and test, lecture learning. Such an investigation might focus on the relationship between test performance and use of idiosyncratic mnemonics, organizational frameworks, errors of omission/commission, or postlecture revision of notes for purposes of adding to or clarifying points made during note taking. The area of L2 lecture-information processing and note taking is wide open to both experimentally minded and ecologically minded researchers; both types of research are needed.

The Exploratory Nature of the Analysis

As Gouran, Brown, and Henry (1978) note, stepwise regression analysis "is useful as an exploratory tool at preliminary stages of theory development. It can suggest a possible hierarchical structure but is generally insufficient to be confirmatory" (p. 62). Care should therefore be exercised in interpreting the findings of this study and relating them to pedagogical practices at this early stage of L2 note-taking research.

Future researchers interested in examining the process and product of lecture note taking might consider employment of hierarchical rather than stepwise regression analysis to establish the extent to which the set of predictor variables identified in this study conforms to achievement patterns found in additional studies. Use of both stepwise and hierarchical analyses may aid in the

construction of a set of criteria that factors into a "quality of notes" index for subjects from various ethno-cultural populations. Improving estimates of the relationship between content of notes and test performance for various groups of note takers could aid in the development of a model of effective lecture-information processing and provide information concerning the metacognitive strategies used by L1 and L2 students to facilitate such processing.

The Need for Continued Research Into the Process and Product of Lecture Note Taking

Ganske (1981) points out that although note taking is an important representation of the knowledge transition that takes place in university learning environments, it has largely been ignored by the L1 and L2 educational communities as a phenomenon of study. Until we begin examining lecture note taking by including a wide range of student attributes (e.g., ethno-cultural affiliation and language proficiency) and strategy employment (e.g., use of mnemonic or organizational devices) related to the note-taking process, we will remain, as Ganske contends, far from developing a functional theory of lecture-information processing and/or student note taking that can be translated into classroom practices and the design of instructional materials on study skills.
■

ACKNOWLEDGMENTS

I acknowledge the assistance of Professors Shitala Mishra, Hershel Thornberg, Roseann Gonzalez, and Frank Pialorsi (co-author of *Advanced Listening Comprehension*) during the formulation and implementation of this research study at the University of Arizona. I am especially indebted to Dr. David Berliner, whose early work in the area of note-taking research inspired me to begin investigating this phenomenon. I am also grateful to Professor Dennis Gouran, Head of the Department of Speech Communication at Penn State, and the anonymous *TESOL Quarterly* reviewers for their incisive and substantive critiques of an early version of the manuscript.

THE AUTHOR

Patricia Dunkel is Assistant Professor at The Pennsylvania State University, where she teaches graduate-level courses in the MATESL program and the doctoral program in speech communication. She is co-author of *Advanced Listening*

Comprehension (Newbury House, 1982), *Intermediate Listening Comprehension* (Newbury House, 1986), and *Start With Listening* (Newbury House, 1987). She also conducts research into the instructional applications of computer-generated speech.

REFERENCES

Aiken, E. G., Thomas, G. S., & Shennum, W. A. (1975). Memory for a lecture: Effects of notes, lecture rate, and information density. *Journal of Educational Psychology, 67,* 439-444.

Anderson, J. R. (1980). *Cognitive psychology and its implications.* San Francisco: W. H. Freeman.

Annis, L., & Davis, J. K. (1975). The effect of encoding and an external memory device on note taking. *Journal of Experimental Education, 44,* 44-46.

Armstrong, W. H. (1956). *Study is hard work.* New York: Harper & Row.

Baker, L., & Lombardi, B. R. (1985). Students' lecture notes and their relation to test performance. *Teaching of Psychology, 12,* 28-32.

Berenson, M. L., Levine, D. M., & Goldstein, M. (1983). *Intermediate statistical methods and applications: A computer package approach.* Englewood Cliffs, NJ: Prentice-Hall.

Carrell, P. (1987). The effects of rhetorical organization on ESL readers. *TESOL Quarterly, 18,* 441-469.

Christiansen, H. D., & Vargata, M. L. (1975). *Study power: Better study skills, greater success in college.* Tucson, AZ: Peter Juul Press.

Crawford, C. C. (1925a). The correlation between college lecture notes and quiz papers. *Journal of Educational Research, 12,* 282-291.

Crawford, C. C. (1925b). Some experimental studies of the results of college note-taking. *Journal of Educational Research, 12,* 379-386.

Di Vesta, F. J., & Gray, G. S. (1972). Listening and note taking. *Journal of Educational Psychology, 63,* 8-14.

Dunkel, P. (1985). *The immediate recall of English lecture information by native and nonnative speakers of English as a function of note taking.* Unpublished doctoral dissertation, University of Arizona, Tucson.

Dunkel, P., & Pialorsi, F. (1982). *Advanced listening comprehension: Developing aural and note taking skills.* Rowley, MA: Newbury House.

Fisher, J. L., & Harris, M. B. (1973). Effect of note taking and review on recall. *Journal of Educational Psychology, 65,* 321-325.

Fisher, J. L., & Harris, M. B. (1974a). Effect of note-taking preference and type of notes taken on memory. *Psychological Reports, 35,* 384-386.

Fisher, J. L., & Harris, M. B. (1974b). Note taking and recall. *Journal of Educational Research, 67,* 291-292.

Freyberg, P. S. (1956). The effectiveness of note-taking. *Education for Teaching, 39,* 17-24.

Ganske, L. (1981). Note-taking: A significant and integral part of learning environments. *Educational Communication and Technology, 29,* 155-175.

Gouran, D. S., Brown, C., & Henry, D. R. (1978). Behavioral correlates of perceptions of quality in decision-making discussions. *Communication Monographs, 45,* 51-63.
Hartley, J. (1976). Lecture handouts and student note taking. *Programmed Learning & Educational Technology, 13,* 58-64.
Hartley, J., & Cameron, A. (1967). Some observations on the efficiency of lecturing. *Educational Review, 20,* 30-37.
Hartley, J., & Davies, I. K. (1978). Note-taking: A critical review. *Programmed Learning & Educational Technology, 15,* 207-224.
Hartley, J., & Fuller, H. (1971, August/September). The use of slides in lectures: An exploratory study. *Visual Education,* pp. 39, 41.
Hartley, J., & Marshall, S. (1974). On notes and note taking. *Universities Quarterly, 28,* 225-235.
Howe, M. J. A. (1970a). Note-taking strategy, review, and long-term retention of verbal information. *Journal of Educational Research, 63,* 285.
Howe, M. J. A. (1970b). Using students' notes to examine the role of the individual learner in acquiring meaningful subject matter. *Journal of Educational Research, 64,* 61-63.
Howe, M. J. A. (1974). The utility of taking notes as an aid to learning. *Educational Research, 16,* 222-227.
Kiewra, K. A. (1984). Implications for note taking based on relationships between note taking variables and achievement measures. *Reading Improvement, 21,* 145-149.
Kiewra, K. A., & Fletcher, H. J. (1984). The relationship between levels of note-taking and achievement. *Human Learning, 3,* 273-280.
Ladas, H. (1980). Summarizing research: A case study. *Review of Educational Research, 50,* 597-624.
Laycock, S. R., & Russell, D. H. (1941). An analysis of thirty-eight how-to-study manuals. *School Review, 49,* 370-379.
Locke, E. A. (1977). An empirical study of lecture note taking among college students. *School Review Research, 77,* 93-99.
Maddox, H. (1963). *How to study.* London: Pan Books.
Maddox, H., & Hoole, E. (1975). Performance decrement in the lecture. *Educational Review, 28,* 17-30.
Mason, A. (1983). *Understanding academic lectures.* Englewood Cliffs, NJ: Prentice-Hall.
Norton, L. S. (1981). The effects of note taking and subsequent use on long-term recall. *Programmed Learning & Educational Technology, 18,* 16-22.
Nye, P. A. (1978). Student variables in relation to note-taking during a lecture. *Programmed Learning & Educational Technology, 15,* 196-200.
Otto, S. A. (1979). Listening for note-taking in EST. *TESOL Quarterly, 13,* 319-328.
Palkovitz, R. J., & Lore, R. K. (1980). Note taking and note review: Why students fail questions based on lecture material. *Teaching of Psychology, 7,* 159-161.

Pauk, W. (1974). *How to study in college* (2nd ed.). Boston: Houghton Mifflin.
Plaister, T. (1976). *Developing listening comprehension for ESL students: The kingdom of Kochen.* Englewood Cliffs, NJ: Prentice-Hall.
Raygor, A. (1970). *Listening and note taking: A program for self instruction.* New York: McGraw Hill.
Rothkopf, E. Z. (1970). The concept of mathemagenic activities. *Review of Educational Research, 40,* 325-336.
Ruetten, M. K. (1986). *Comprehending academic lectures.* New York: Macmillan.
Sims, J., & Petersen, P. W. (1981). *Better listening skills.* Englewood Cliffs, NJ: Prentice-Hall.
So, W. Y. (1974). A new language laboratory program for advanced students. *TESOL Quarterly, 8,* 293-304.
Sotiriou, P. (1984). *Integrating college study skills: Reasoning in reading, listening, and writing.* Belmont, CA: Wadsworth.
Tabachnick, B. G., & Fidell, L. S. (1983). *Using multivariate statistics.* New York: Harper & Row.
Yarington, D. J. (1977). *Surviving in college.* Indianapolis: Bobbs-Merrill Educational Publishing.
Young, L., & Fitzgerald, B. (1982). *Listening and learning lectures.* Rowley, MA: Newbury House.