

CHAPTER 7

FINDINGS

7.1 Quantitative Results

Of the four hypotheses tested in the present project, two were addressed quantitatively. These were the hypotheses that significant gains in learning would be shown in students' scores from pre to post tests as a result of students learning on computer about gerunds and infinitives with the matrix verbs in questions, and that the greatest gains would be realized when students used the Game Paddle (PDL) lessons. In both the HLC and LCC studies, the latter hypothesis was rejected. As for the first hypothesis, gains made by native speakers of English undergoing treatment at LCC were significant ($\alpha = .05$) as opposed to those made by a control group, for which results were not significant; however, the ESL students at HLC showed no significant gains in either the PDL or the REG groups.

Gains from pre to post tests for the PDL group at HLC appeared to be tending toward significance (and were significant at $\alpha = .10$), whereas those for the REG group were practically nil. This suggests that, with the ESL students at HLC, learning took place to a greater extent when students used the PDL lessons than when they used the REG ones. However, these results are not conclusive, since they did not meet the test for significance at $\alpha = .05$.

Quantitative results are summarized in Tables 1 to 7 (plus, there is a chi square analysis of retention included with the qualitative results, at the end of Tables 8, 9, and 10). These tables break the data down into that from the two research sites, into that from quantitative and qualitative studies, and further into that pertaining to PDL, REG, and CTL groups, with data pertaining to more than one group given in separate tables. This organization was deemed most suitable for quick reference, but not necessarily for presentation in the argument that follows; therefore tables may be referred to out of order in the following text. All results reported on the tables and in this chapter are for one-tailed tests. It is also noted where t values met tests at $p < .10$. No tests were made against critical values falling beyond $p > .10$ (for one-tailed tests).

7.1.1 Quantitative Results from HLC Data

In the experiment conducted at HLC, two groups, PDL and REG, were formed with N sizes of 13 and 11, respectively. The students were all non-native speakers of English. There had originally been 13 students in the REG group, but one did not complete and another compromised the pre test.

The two groups were shown (Table 3) to have been nearly equal before treatment through comparisons of their MTELP and pre test scores. The MTELP scores, which were known prior to the experiment, compared favorably ($t = .517$;

d.f. = 22; $p > .10$). Thus, it was shown that the two groups were originally similar in English language proficiency. Subsequent analysis of pre test scores ($t = .002$; d.f. = 22; $p > .10$) suggested that before treatment there had been essentially no difference in knowledge of use of gerunds and infinitives with the matrix verbs 'stop', 'remember', 'forget', and 'regret' for either group of students.

After treatment, students in the REG group showed a slight increase in mean pre to post test scores, from 22.91 to 23.18 (Table 2). A t value of 0.132 (d.f. = 20; $p > .10$) shows this to be a negligible difference. On the other hand, for students using the game paddles (Table 1), mean pre and post test scores increased from 22.92 to 26.23 ($t = 1.574$; d.f. = 24; $p < .10$; n.s.). Although short of significance at $\alpha = .05$, these data tend to suggest that learning was enhanced when students used the lessons with game paddles.

A study of retention of the material taught was also undertaken. On the follow-up questionnaire, there were three right/wrong questions whose purpose was to give an indication of retention. Twice as many PDL students answered these questions correctly as did REG students (Table 8), yielding $\chi^2 = 2.961$ (d.f. = 1; $p < .10$; n.s.). Again, although it was not shown conclusively (i.e. at $\alpha = .05$) that there was a greater degree of retention for students

working the PDL lesson, there seemed to be a tendency in that direction.

In an effort to gather more information about that tendency, the post test was given again to all students in the survey at HLC. However, the results (Table 1) were not particularly striking. In the first place, only 8 of the original 13 PDL group students sat the second post test. Although the PDL group still scored higher than did the REG students, the difference in the means of the two post tests was not significant ($t = .897$; d.f. = 17; $p > .10$). In addition, comparing the means of the two post test scores within the two groups, no significant differences were found. Values for t were $.545$ (d.f. = 14; $p > .10$) and $.518$ (d.f. = 20; $p > .10$) for the PDL and REG groups, respectively. So, scores on the two post tests were found to be essentially the same as before for both groups of students, and neither group was found to have retained the material better than the other.

7.1.2. Quantitative Results from LCC Data

Three experimental groups were formed at LCC: PDL, REG, and CTL. Each group had originally had ten students; however, one student in the CTL group invalidated the pretest by not completing it. A student selected at random to replace this student was put on the PDL lesson by mistake. Meanwhile, another student in the PDL group was considered not to be a

member of this group by virtue of her Nelson-Denny score being almost twice the next highest, and so this student was eliminated from the study. Thus, the resulting groups of PDL, REG, and CTL had N sizes of 10, 10, and 9, respectively.

These groups were shown to be approximately equal (Table 7) by comparison of means of their scores for the Nelson-Denny test of reading comprehension and vocabulary, and by a comparison of the pre test scores of students in each group. (Nelson-Denny scores were available for all subjects except for one student in the PDL group. Since her pre and post test scores were about normal for that group, that student was assigned the group mean as her Nelson-Denny score.) No t values for the N.D. scores showed significant differences between groups, indicating near equality.

Likewise, when the pre test scores were examined, no significant differences were shown to have existed between the three groups. However, for the difference in means between the PDL and REG groups, a t value of 1.220 (d.f. = 18; $p > .10$) was obtained. This is a large, if not significant, difference, especially considering that the critical value of t at $\alpha = .10$ is 1.330 (d.f. = 18). Thus, a more than negligible difference in prior knowledge of gerund and infinitive use may have existed between students in the PDL and REG groups. If in fact such a difference did exist,

it may help to account for the apparently poorer showing, as measured by t values, of the PDL group when compared to REG group students in the LCC experiment.

Means of pre and post test scores showed significant differences for both the PDL and REG groups at LCC, yielding t values of 2.053 (d.f. = 18; $p < .05$) and 4.568 (d.f. = 18; $p < .0005$), respectively (Tables 4 and 5). These values are significant; thus, the hypothesis that both lessons would result in learning was accepted. However, the results do not support the hypothesis that the PDL group would exhibit greater learning. This might have been due in part to a ceiling effect, which existed with all three groups, but which was more apparent with the PDL group. The CTL group (Table 6) exhibited a negligible change from pre to post test scores ($t = .322$, d.f. = 16; $p > .10$).

7.1.3. Discussion of the Quantitative Results at HLC

The experiments were designed so as to reveal whether or not students working the lessons with game paddles would be more cognizant of what they were doing than were those who interacted with the lesson by pressing buttons. In analyzing the results of these experiments, it is important to pay particular attention to the language levels of the students and to the amount of time they were actually able to derive benefits from the lessons they worked.

The students at HLC were ESL students with little prior knowledge of the subject matter being taught and tested. On average, they were able to answer correctly just 23 out of a possible 40 items on the pre tests they were given. It is not unreasonable to assume that some had never had prior instruction in the linguistic features they were dealing with, and that they were thus being exposed for the first time to a subtle distinction in gerund and infinitive use in English.

Another major factor influencing the results was the time allowed for the experimental treatments, and the time within that allowance that the students were actually able to concentrate on the linguistic features being taught. At HLC, the time allowed, 35 minutes, was possibly not long enough. One question on the follow up questionnaire (#4) addressed this possibility, and ten out of the 24 respondents at HLC reported not having a satisfactory amount of time for working the lessons (Table 8). Of these, only 3 reported that they actually did not finish the lesson, but it is not clear that the respondents were all working from the same definition of the term "finish".

In the case of the REG group, to finish the lesson was probably understood to mean 'to work the lesson from beginning to end, completing the problem and recapitulation sections'. For these students, there was a clear-cut beginning and end

to the lesson. In the case of the PDL group, the term "finish" would have been less clearly defined. To the students who answered the questionnaire, it apparently meant 'to work the lesson until satisfied that learning had taken place'. To the researcher, who had designed the lesson, it meant 'to work until satisfied that learning had taken place, and then to work through the recapitulation and rules section'.

According to the information they gave on the follow up questionnaires, two REG students reported not finishing their lesson in 35 minutes (Table 8, Question #4); however, the researcher observed three who did not reach the rules portion of the lesson. (An additional five students in this group mentioned that they would have liked to have had more time to work the lesson.) Only one student reported not "finishing" the PDL lesson, but data on whether the remaining students had worked the recapitulation were not collected (because it was not possible to watch all students as they worked the lessons). Few students in the PDL group were observed to have worked on the recapitulation portion, and many were observed to have worked nothing but problems. These students, by the researcher's definition, did not finish the lesson.

Most of the PDL students exhibited a hesitancy to explore options on the computer (11 of these students had never used a computer before; only five had used one more

than twice -- see Table 8, Question #1). Therefore, few students attempted typing "enough", which would have put them into the lesson's recapitulation mode (see Chapter 4, Section 4.2.7 for a description of this feature). Most preferred remaining with the predictable pattern of making choices with the game paddles, pressing a button, working a problem, and then manipulating the game paddles again. It is likely that PDL students who had done this a number of times before being asked to stop said that they had "finished" the lesson.

So, in evaluating the experimental results, we must first take into account whether the students finished the lesson (i.e. worked the rules and recapitulation). Also, we must consider what they actually did while they were on the computer. Students in the REG group usually required little or no assistance after having begun their lessons. Their lessons began with a brief explanation of what was to follow, but after about five minutes, they would usually be working problems. However, students in the PDL group commonly had difficulty learning how to use the game paddles. The more dexterous students were able to get through the instructions and to figure out the system in around ten out of their allotted 35 minutes, but some students needed 15 or even 20 minutes before they could generate problems comfortably. Even then, most of these students tended to

ponder their choices, studying the chart and taking several minutes to form sentences, so that they frequently worked only five or ten problems before being asked to stop. The REG students, on the other hand, would typically work their problems at the rate of nearly two per minute.

Therefore, even though all students in the experiment spent equal amounts of time in front of computers, it was apparent that the PDL group students were spending much less time than the REG students working the problems in the lessons. Furthermore, the PDL students were observed to have worked not even half as many problems as did the REG students. It is therefore interesting that students working in this slow and plodding way made nearly significant gains in post test scores, whereas the students who worked two or three times as many problems on the REG lessons made only negligible gains.

It is equally interesting to consider the strategies with which individual students approached the different lessons, particularly in conjunction with the results of these strategies. For example, no relationship was found between degree of completion of the REG lesson and performance on the post test. S9-REG/HLC, a student in the REG group, completed the lesson, restarted, worked to problem 5 again before being asked to stop, and then dropped 4 points on his post test. Another student, S5-REG/HLC, managed to work through

the entire REG lesson twice in 35 minutes, but still went down 5 points from his pre test score. Apparently, he had settled on a strategy for successfully getting through the lesson which carried over little to the real world.

Conversely, neither S10-REG/HLC nor S4-REG/HLC had finished the lesson after 35 minutes, yet they increased 2 and 5 points respectively from pre to post test (Table 2).

Before beginning their post tests, 8 of the 11 REG students had worked all 32 problems and gone through a recapitulation of the rules governing the behavior of the linguistic items studied. On the other hand, students working the PDL lessons had for the most part worked only a few problems, and most of them had never entered the quiz, rules, and recapitulation sections. In spite of this, they made, on average, greater gains from pre to post tests than did the REG students. Two reasons postulated for these gains are that the students were more actively involved in what they were doing and that they had the option of altering the lesson according to their own learning styles and strategies.

Although not many students became comfortable enough with the computer to exercise their options, at least one did, and with astonishing results. When S10-PDL/HLC experienced frustration with the game paddles, rather than attempt to work with them, she avoided them by typing

"enough" and working through the quiz, rules, and recapitulation. When she finished this, she had ten or fifteen minutes remaining in the treatment, so the researcher suggested she try working with the game paddles again. She selected only one problem from the chart before typing "enough" again and working completely through the quiz, rules, and recapitulation a second time. Working in this way, S10-PDL/HLC improved from 15 to 30 on the pre and post tests, the greatest leap by any subject in the entire experiment (Table 1). It may be that S10-PDL/HLC's having been afforded the opportunity to follow the learning strategy that most appealed to her was primarily responsible for her remarkable performance; had she used the REG lesson, she wouldn't have been able to exercise such an option.

There might also be implications here regarding the success of deductive, as opposed to inductive, learning. S10-PDL/HLC obviously employed a deductive strategy, with some success. Further evidence of the relative efficacy of deductive and inductive learning was provided when S3-PDL/HLC, a PDL student who concentrated on forming sentences during his 35 minute allotment, asked the researcher to explain the grammatical points discussed in the computer lessons. He had just taken the post test, in which he dropped 4 points from his score on the pre test (Table 1). But after having had the grammatical rules explained to him face to face, he

seemed finally to understand what the lessons had been about, and he confirmed this by verbally paraphrasing the rules in the presence of the researcher. In his case, the inductive approach had not worked, yet he appeared to understand the principles involved when given the rules. Perhaps the inductive process had prepared him to fully comprehend the researcher's explanation of the rules.

One question that arises in conjunction with S10-PDL/HLC's success at the computer is whether the fact that she doubled her score when the next highest increase was by a factor of 1/3 skewed the results unfairly. If she had not been in the survey, the mean increase in pre to post test scores for the PDL group would have been from 23.58 ($s = 5.14$) to 25.92 ($s = 5.35$). This would have resulted in a t value of 1.089 ($d.f. = 22$; $p > .10$), which would have fallen farther short of significance than did the result obtained with her included in the study (Table 1).

This is in any case moot, since she should be included in the survey. Her MTELP score of 57 is just over the mean for all subjects surveyed. Although her pre test score of 15 was at the bottom of the range for pre test scores, it was matched by a student in the REG group. There follow one 18, two 19's, four 20's, and so on (Tables 1 and 2). Also, S10-PDL/HLC's score was not necessarily an aberration, since she obtained it by selecting her learning strategy from the

available options and then utilizing the computer to pursue that strategy. This was what she was supposed to do, and what it had been hoped that other students would do. Thus, there are compelling reasons for including her results in the PDL group data.

7.1.4 Discussion of the Quantitative Results at LCC

Although evidence of enhanced cognition with the PDL lessons was not found in the data from LCC, the CALL lessons were shown to be effective. In the whole group of 20 subjects receiving the experimental treatments at LCC, all but one showed increases in scores from pre to post tests, and both kinds of lessons were shown to be effective statistically (Tables 4 and 5). The control group, on the other hand, which worked CALL lessons unrelated to gerunds and infinitives, showed a negligible mean increase in scores from pre to post tests. Furthermore, three of the nine students went down in their scores, suggesting that improvement and reduction of score from pre to post test had nothing to do with the control treatment (Table 6). Thus, it was established that not only did the experimental CALL lessons teach what they were supposed to, but that similar results were not obtainable using a form of computer treatment not specifically dealing with the subject being taught.