An Experimental Evaluation of an Audio-Tutorial System.

Quebec Dept. of Education, Quebec.

8 Aug 72

14p.; Paper presented at the Association for the Development of Instructional Systems Conference (Cap Rouge, Quebec, Canada, August 8, 1972)

MF-$0.65 HC-$3.29

Elementary School Students; Independent Study; *Information Retrieval; Information Seeking; Instructional Materials; Intermode Differences; Library Skills; *Research Skills; Search Strategies; *Tape Recordings

*Audiotutorial System

As educators become more interested in providing an environment and atmosphere which are conducive to independent learning and study activities, it is increasingly important that students be able to systematically coordinate their research skills in order to effectively use a centralized instructional materials center. To test the effectiveness of an audio-tutorial system designed to provide instruction in information retrieval and selection through actual search problems, 60 students in the 6th grade were divided into two groups; one group used the system, the other did not. The differences between these two groups were assessed. Results showed that those students who had participated in the system had more favorable attitudes toward the library and did better on a subsequent posttest of their library searching skills. (SH)
An Experimental Evaluation of an Audio-Tutorial System,

A Paper Presented to
the Association for the
Development of Instructional
Systems Conference, Ecole Normale de Ste.Foy
Cap Rouge, Quebec, Canada
August 8, 1972.

Kyran Kennedy
Lecturer
Memorial University of Nfld.

Acknowledgement: The research report in this paper was assisted by grants under the Quebec Minister of Education's "Programme de formation de chercheurs et d'action concertée."
INTRODUCTION

This paper is concerned with the development of an audio-tutorial approach to aid pre-adolescent school children retrieve information from an organized body of learning materials called a resource center or library. More specifically it is a brief report on research and experimental work* carried out in this area.

Prior to the experiment** several schools in the Montreal area were visited in order to investigate elementary school children's habits and methods of using resource centers.

As a member of a research group, I interviewed principals, librarians, teachers and students. Most were aware that problems existed and some were attempting to solve them in diverse ways. Often, however, it was my feeling that as Frederick Hartz, Assistant Professor of Library Science, Trenton State College, summarized,

*This research was aided through a grant received by Dr. Gary Boyd from the Department of Education, Quebec.

**This experiment was carried out to fulfill the thesis requirement for a Master of Arts in Educational Technology, Sir George Williams University, Montreal.
"Only one librarian in four is sufficiently familiar with classroom activities of her school to analyse the teaching methods employed"¹ (1966). Librarians generally were unable to integrate the library as an effective instructional tool. Neither were the teachers in a more advantageous position for carrying out library instruction, and I was inclined again to agree with Hartz when he wrote, "Since a majority of the faculty themselves have to be prodded into making better use of the library, it is doubtful whether they could stimulate better student use"² (1966).

Against this background arose my observations and hypotheses which formed the initiation stage of the experiment in the audio-tutorial library self-instructional system.

My major observation was that although the students possessed the basic skills necessary for using the library they were generally unable to organize these skills into an overall systematic way of retrieving and selecting materials and information from the library.


² Ibid
My hypotheses were:

1. A possible solution to some of the problems encountered by students would be the development of a self-instructional procedure using the audio-tutorial method to train children in developing an overall search and retrieval strategy in locating and selecting materials needed to solve informational problems.

2. It was further hypothesized that such a project would create positive attitudes towards library use on the part of the children involved.

Before going into the design of the experiment I would like to briefly state the need for library instruction and the possible long-range goals of the proposed self-instructional program.

Today, as educators become more and more interested in providing an environment and atmosphere which are conducive to independent learning and study activities, it is increasingly important that students be able to systematically coordinate their research skills in order to effectively use a centralized instructional materials center. This is more obvious when we observe that concepts such as non-graded, non-text, and individualized instruction are becoming practical realities.

It may be further observed that children who do not possess the necessary ability to retrieve information from resources often become frustrated when given assignments for which library use is required. They therefore, often with the encouragement of the teacher, tend to rely upon a limited number of readily available - although not necessarily the best - resources for their information.

Also, the more creative aspects of the learning process
such as evaluation, elimination, combination, cultivation, and sharing cannot be carried out in an efficient and less time-consuming manner unless students have developed an integrated, systematic and automatic procedure for exploring and using available resources.

The long-range goals of the proposed system are: To develop audio-tutorial programs concerned with providing instruction for children in information retrieval and selecting information by actually carrying out a research activity. For as Perkins emphasizes, "Students learn by doing and not by looking or by being lectured" (1970).

The long range instructional goals are that this program, having proved successful, will further enable pupils to carry out independent learning tasks with a minimum of frustration and at a higher level of confidence than previously possible.

It is hoped that the program, with detail changes, can be locally prepared and used in other elementary school libraries to achieve similar instructional objectives.

Before developing the design of the experiment such pre-design consideration as: analysis of retrieval tasks; a general description of entering behaviour of population; statement of cognitive and attitudinal behaviour which would have to be taken into consideration; statement of instructional objectives; development of appropriate

evaluation devices and obtaining necessary resources; had to be undertaken. As important as these are, it is impossible to discuss them in so limited a time—therefore, from here I would like to go on to describe the design of my experiment, the experiment as it was carried out, and to discuss some of the results and tentative conclusion—although the data have not as yet been fully analyzed.

EXPERIMENTAL DESIGN

My null hypotheses were: that there would be no significant difference between the experimental group receiving the taped library instruction and a control group not receiving this instruction in their ability to retrieve and select resources from the school library as based upon scores obtained from a locally devised performance evaluation. Furthermore, there would be no significant difference between the experimental and control groups in their attitude toward use of the library after the experimental group had experienced the instructional program. This was determined by comparing scores obtained by both groups on a locally developed library attitude scale.

If the null hypotheses were rejected, \( p=.05 \), the hypotheses stated above were to be accepted.

EXPERIMENTAL PROCEDURE:

Sixty grade six students out of a possible 86 were randomly assigned to two groups. Random sampling was done by matching student numbers with a table of computer generated random numbers. One group, the experimental group, received the instruction while the other group, the control group, did not.
Both groups were given the pre- and post-pencil and paper attitude test. Although it was my intention to also give pre- and post-performance tests, circumstances did not permit me to do so. However, analysis of the results and certain checks, to be discussed later, show that the two groups selected were unbiased as far as achievement, attitude, sex and age were concerned.

These are indicators that using random numbers is a reliable method of selecting unbiased sample numbers from a given population.

The instructional content of the program was developed by means of a task analysis, to which mention was made earlier. Instructional objectives were also formed and a performance test designed to measure whether subjects performed in accordance with them.

The program itself was written up in script form and recorded by a professional on a master tape which was then duplicated onto cassette tapes. The length of the script was 17 minutes. However, the actual performance time averaged approximately forty-five minutes.

Having made the appropriate arrangements with the teachers concerned the experimenter sent for each subject individually. When he arrived at the resource center he was greeted and presented with a search problem of the type, "Explain how volcanoes are formed". All instructions including those for the operation of the player, were
written down for him to read. If he had any questions concerning this matter the observer helped him with brief explanations. As instructed he took the cassette player equipped with a shoulder strap and headphones and performed the activities required by the "voice" on the tape. As he carried out these activities the observer noted his performance for difficulties. (These were few since most of them had been eliminated by a pilot study carried out on Grade Five students the previous week). However, the aim was that the observer was to interfere as little as possible, except where such a problem as operation of equipment or other similar difficulty arose. Having performed the retrieval activities the subjects were instructed by the tape to analyze the materials and select them for relevance. The following week both the experimental and control groups were given similar research problems. This was a performance test which was evaluated concurrently by two observers. First one subject from one group was asked to perform then another subject from the other group was asked to perform also, until fifty-nine of the sixty were completed. One of the control group members was not available for testing.

The instructional period ran for five days and the performer test period also was five days in length. At the end of the following week the post-test on attitudes was administered by the class teachers. This meant that there was an interim of twenty days between
the pre-attitude and post-attitude tests, Both tests were identical in content.

EXPERIMENTAL RESULTS:

Before going into the results of the experiment let us consider some further characteristics of the population sample. You may recall that the sample was selected by using random numbers. This process assures an unbiased sample in the experimental and control groups. As a verification of this, comparisons were made between the experimental and control groups to determine whether or not they differed significantly on any of the relevant variables. Sex-wise, randomization yielded almost perfectly split groups: the experimental group contained fifteen boys and fifteen girls while fourteen boys and sixteen girls made up the control group. The mean age of the experimental subjects was 11.57 years; that of the control 11.62 years. In order to find out if a significant difference existed in attitude or achievement between the two groups, t-tests were performed using the means obtained by both groups on the locally devised attitude pre-test and the Stanford Achievement Test, Intermediate II Form Y. Table 1 summarizes these comparisons. The observed t-ratios indicate clearly that the two groups did not differ significantly in either their attitude or their achievement measures.
TABLE 1  Attitude Pretest

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>df.</th>
<th>Difference</th>
<th>t</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>30</td>
<td>49.73</td>
<td>58</td>
<td>0.20</td>
<td>0.16</td>
<td>N.S.</td>
</tr>
<tr>
<td>Control</td>
<td>30</td>
<td>49.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Achievement

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>df.</th>
<th>Difference</th>
<th>t</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>29</td>
<td>7.30</td>
<td>54</td>
<td>0.02</td>
<td>0.05</td>
<td>N.S.</td>
</tr>
<tr>
<td>Control</td>
<td>27</td>
<td>7.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* At .05 level of confidence.

One of the main questions to be answered by the experimental results is: Did the members of the experimental group obtain significantly higher scores on the performance evaluation than the members of the control group? Table II summarizes statistical measures which allow an answer to this question.

TABLE II  Performance Test Evaluation Experimental and Control

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>df</th>
<th>Difference</th>
<th>t</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>30</td>
<td>20.17</td>
<td>57</td>
<td>4.65</td>
<td>3.74</td>
<td>.001</td>
</tr>
<tr>
<td>Control</td>
<td>29</td>
<td>15.52</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A two tailed t-test was performed to estimate the significance of the obtained difference between the experimental and control group means. Its outcome allows the rejection of the null hypothesis that claims no significant difference.

The performance test evaluations were made independently by two raters. The inter-rater reliability co-efficient (Pearson r) was .91.
In addition, since the performance time for individuals varied, a t-test was performed to detect any significant differences between the groups in amount of time spent on the performance task. The t-ratio obtained was 1.18 which with 57 degrees of freedom, did not prove significant at the .05 level.

Statistics were employed to decide objectively whether there had been any change in the attitude of experimental and control groups from the pre-experimental to the post-experimental conditions. Pre-attitude scores were compared with post-attitude scores for each group by means of a t-test. Table III summarizes these comparisons.

<table>
<thead>
<tr>
<th>TABLE III Attitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental Group</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td>Pre</td>
</tr>
<tr>
<td>Post</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td>Pre</td>
</tr>
<tr>
<td>Post</td>
</tr>
</tbody>
</table>

The t-ratios obtained inform us that there was a significant increase in positive attitude scores by the experimental group while there was no parallel increase noticed with the control group.

Test-retest reliability was obtained by correlating the combined experimental and control group scores received on the pre-attitude test with the scores obtained by the same individuals
on the same test after a time interval of three weeks. The reliability co-efficient was .68. The validity of the attitude scale was checked by comparing the attitude scores obtained by 116 pupils from grades VI and VIII with their teachers judgement of their attitude toward the library; teachers rated each pupil as having either a positive or negative attitude and a median split was performed on the attitude scores.

Table IV registers the frequencies obtained for the respective blocks. A 2 x 2 Chi-square test was run on the data and the resulting Chi-square value was 13.59. This exceeded the critical value with 1 degree of freedom at the .001 level of significance.

<table>
<thead>
<tr>
<th></th>
<th>High</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>56</td>
<td>40</td>
</tr>
<tr>
<td>Negative</td>
<td>2</td>
<td>18</td>
</tr>
</tbody>
</table>

**SUMMARY**

The foregoing Statistical Analysis indicates that the attitude test was reliable and valid and could measure change in attitude towards the library. In this case the post experimental group's attitude changed in a positive direction while no such change was measured between the pre- and post-control groups.

The performance test's validity depended upon the thorough analysis of the appropriate tasks to be performed. It's reliability was checked by comparing the two observers' ratings. A very high positive correlation was indicated.
Furthermore, statistical tests indicate that this audiotutorial instruction was responsible for the higher performance abilities of the experimental group.

Finally, it had been my intention to discuss what I think are some basic elements characteristic of this system which facilitate learning - from the student's point of view and the educator's. I would also like to discuss some of my less formal observations, but for obvious reasons this is impossible. Now let us look at some slides showing the experimental subjects carrying out retrieval tasks. At the same time let us listen to some excerpts from the instructional cassette tape.
REFERENCES
