R E P O R T R E S U M E S

ED 019 902 FL 000 785

THE KEATING REPORT--A SYMPOSIUM.

BY- STACK, EDWARD M. AND OTHERS

NATIONAL FED. OF MODERN LANGUAGE TEACHERS ASSN.

PUB DATE APR 64

EDRS PRICE MF-$0.25 HC-$0.96 22P.

DESCRIPTORS- *EVALUATION, *RESEARCH PROJECTS, *SECOND LANGUAGE LEARNING, *LANGUAGE LABORATORY USE, *PROGRAM EFFECTIVENESS, RESEARCH UTILIZATION, RESEARCH REVIEWS (PUBLICATIONS), LANGUAGE LABORATORIES, RESEARCH METHODOLOGY, EVALUATION TECHNIQUES, STATISTICAL ANALYSIS, KEATING REPORT,

THIS CRITIQUE OF RAYMOND F. KEATING'S 1963 "STUDY OF THE EFFECTIVENESS OF LANGUAGE LABORATORIES" CONSISTS OF FOUR ESSAYS DISCUSSING THE BIASED THINKING AND FAULTY RESEARCH PROCEDURES WHICH CHARACTERIZE THE REPORT. THESE ESSAYS EXAMINE SUCH FAULTS IN THE REPORT AS THE TOTAL LACK OF DEFINITION OF WHAT IS UNDER INVESTIGATION, THE USE OF INAPPROPRIATE AND OUTDATED TEST MATERIALS TO JUDGE THE EFFECTIVENESS OF LANGUAGE LABORATORIES, AND THE NEGATIVE PRESENTATION OF RESEARCH RESULTS. DESPITE ITS MANY SHORTCOMINGS, HOWEVER, THE KEATING REPORT YIELDS OBJECTIVE EVIDENCE THAT THE THE LANGUAGE LABORATORY IS IN EFFECT DOING WHAT IS EXPECTED OF IT. THIS ARTICLE IS A REPRINT FROM "THE MODERN LANGUAGE JOURNAL," VOLUME 48, NUMBER 4, APRIL 1964, PAGES 189-210. (SS)

ERIC Full Text Provided by ERIC

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED019902

FL 000 785

*The Keating Report—A Symposium**

Laboratories: The Effectiveness Controversy

EDWARD M. STACK, *University of North Carolina (Raleigh)*

DURING the current academic year there have been two reports concerning the effectiveness of the language laboratory. The very fact that the laboratory makes readily available a great supply of authentic materials in the field under study makes its value almost self-evident. Over the years we have devised methods of putting these materials into a form which requires the active participation of the student; and this practice, essential to language as to every other skill, could hardly be rejected. Yet it is necessary to produce "evidence" for administrators and purchasing agencies in the academic world to justify purchases of all kinds of equipment, and "facts and figures" are grasped for this purpose. One of the current reports, entitled *A Study of the Effectiveness of Language Laboratories* by Raymond F. Keating has been used to prove that language laboratories are not really worthwhile, and that expenditures should not be made for them. This is grist for administrators who prefer to divert language-teaching funds to other projects, and for teachers who are suspicious or fearful of new-fangled aids and ideas. The former have nothing to base judgments and decisions upon other than such reports; the latter need to be convinced through experience such as that supplied by NDEA Institutes, and their colleagues. The second report, *Foreign Language Laboratories in Secondary Schools* (in *A-V Learning*, Vol. VII, No. 1, Oct.–Nov. 1963; Board of Education of the City of New York) provides objective evidence that the language laboratory is of distinct advantage: it shows in general that students using the laboratory were equal in traditional "grammatical" skills to non-laboratory students, and at the same time superior in speaking skills—all in the same amount of class hours.

These two conflicting reports might be quite puzzling to a non-language administrative official. (Unfortunately the Keating Report was mailed, with a covering letter, to a large list of administrative officers. The NYC Report did not receive this specialized distribution.) Which report is to be credited by the academic profession? A careful reading of the experimental designs, and of the conclusions reached by the investigators, will quickly resolve this question.

THE KEATING REPORT

It would seem axiomatic that an investigator should begin by defining his terms. The title includes the words "language laboratory" as the main topic of investigation. Yet Keating does not know or say what this means: "Nor was any attempt made in the present study to define *a priori* what a language laboratory was" (page 22). Nor does Keating worry about what is *done* in a language laboratory: " . . . very little attempt has been made in this study to interpret the results in terms of specific features of the instructional program" (page 38). And again, on page 22, "No claim can be made that this study evaluated the use of the language laboratory as implemented according to any single instructional rationale." In other words, not having defined the laboratory, why bother to see *how* it was used? Yet it is crystal clear that the most important part about a language laboratory is the quality of the audio-lingual structured program used in it. Only the *program* teaches—not the machinery. In selecting his laboratory students, Keating evidently took someone's word that the 'laboratory' was used; he did not even ask in what *way* it was used. So

* The four papers which follow are rebuttals of Raymond F. Keating, "A Study of the Effectiveness of Language Laboratories," Institute of Administrative Research, Teachers College, Columbia University, 1963. Pp. x + 61. $1.75.

we now have a report on the efficacy of something or other, used somehow.

The purpose of the study was to find out whether students who used the laboratory (however defined) were, or were not, superior to those who had no laboratory work. The predisposition of the perpetrators of this pamphlet is revealed in the Foreword by Mr. William S. Vincent, Executive Officer of the IAR, Teachers College, who implies a political boondoggle as he grimly says, "The availability of matching grants [for language laboratories], provided because someone in Washington—I do not know who—thought it might be a good idea, has suddenly escalated outlays for this item into nine digits." (The U. S. Office of Education points out that "nine digits" would be at least $100 million, whereas actually about $15 million has been made available under Title III.) This insinuation is belied by the fact that the language laboratory was invented and developed by language teachers; it was never a whim, advertising plan, nor political issue. The gross inaccuracy, or perhaps exaggeration, concerning the funds indicates some bias. Further negative predisposition may be reflected by the "null hypotheses" to be tested (page 3), in which the statements to be proved or disproved are:

A. Students who have used language laboratories are not superior in reading comprehension to students who have not.
B. Students who have used language laboratories are not superior in listening comprehension to students who have not.
C. Students who have used language laboratories are not superior in fidelity of pronunciation to students who have not.

Even allowing that some statistical quirk requires the adoption of a negative test question, it is to be noted (1) that neither the word "used" nor "language laboratories" is ever defined, as previously pointed out, and (2) that the obvious and stated objectives of the language laboratory are ignored, in that the problem inherent in A above is not one impinging upon the language laboratory at all. (Of course, some rudimentary techniques for reading training are available, but not in wide use.) This is only a hint of the misinformation or misapprehension about language laboratories and their purpose and use which pervades the pamphlet.

Given a misinformed and predisposed negative sendoff, what of the tests? Keating has skipped the intervening processes and gone directly to the results of a year of teaching. His samplings are comprised of the "laboratory groups" and the "non-laboratory groups." He says that "There was but one criterion for inclusion in the laboratory groups as referred to . . . , namely, use of the language laboratory during the year of study" (page 32). What kind of a laboratory? Just a tape recording in a class? A listening post? An audio-active or recording laboratory? With monitoring? Using four-phase pattern drills? Listening to the teacher read from the textbook? Talking English to another student through an intercom? To Keating it makes no difference. A lab is a lab is a lab. A minimal laboratory with an excellent program is preferable to a fully-equipped shiny one using poor programs, but there is no differentiation here. What *does* matter to Keating's study is how well-paid the teachers are (" . . . higher salaries . . . should result in superior teaching") and the I.Q. levels of the students (in spite of evidence that even people with low I.Q.'s can converse intelligibly and even loquaciously in their native language).

The laboratory groups and the control groups were given two sets of tests. For reading comprehension and listening comprehension, the ETS forms of 1940 and 1955 were used. These dates are worthy of note, since both antedate any significant dissemination of structural linguistic methods to the schools and universities, and may therefore be suspected of being strongly oriented towards students with traditional backgrounds in aims and methodology. Students trained under traditional grammar-translation methods would perform better in tests devised especially for them; the vocabulary of graphically-oriented textbooks is not that of the modern structural text. They would not be fair tests for students under the audio-lingual system.

The second test was specially prepared for this study. Called the *French Speech Production Test* and printed in full on the last page of the Keating Report, it demonstrates in content and presentation that the author was unaware of audio-lingual methods. It is a short test. Part

One consists of showing the student ten cards, each with a simple line drawing and *an English word* telling what the drawing represents! (Why the drawing? *or* Why the word?) The cards bore the drawings and words such as "head," "book," "tooth," and "one." The students were required to produce the French word represented. The official answer list would not pass a test in elementary French, since the nine nouns are not accompanied by the appropriate article. One wonders if a student were penalized for responding *le train* instead of just *train* as indicated. On the basis of a team of two raters who were in agreement most of the time, "It is concluded, therefore, that this test instrument as used in this investigation was highly reliable" (page 31).

Part Two of the *French Speech Production Test*, on which the laboratory group made their sole good showing, states that "For each test item, the student was required to read aloud the sentence in French indicated." This is followed by such sentences as:

> Ta carte est là.
> Donnez-moi la main.
> Il faut un peu d'eau.
> Ton père arrive.

Assuming that the samples given in the test cover the principal pronunciation difficulties of French, it is still dubious whether this is a test of speech production, or of converting graphic symbols into sound. A good laboratory student might be able to make a statement such as *Ta carte est là* in the course of a dialogue or upon an appropriate audio or graphic (pictorial) stimulus, but articulating written symbols is another matter. It is a test of reading aloud, and this skill is rare even among native speakers. It is a skill developed by radio announcers, commentators, and professional readers. It is *not* a test of what is being taught in the language laboratory directly. And this is the basis for Keating's conclusion that the laboratory is not too effective!

The other tests (in which null hypotheses A and B were tested) were checks of skills not normally expected to be developed primarily in the language laboratory. This did not deter Keating, who had not bothered to find out what the laboratory was, or what its purposes were. He just went ahead and used tests dating from 1940 to 1955, although the first audio-lingual texts with fully programmed tapes did not appear on the market until about 1959. The reading comprehension test is based upon a graphic skill not normally provided intensively in laboratory courses until after the first year. Listening comprehension which is based upon literary texts, as is customary in traditionally-oriented grammar-translation courses of the 1940's and 1950's, is likely to contain a vocabulary and style unfamiliar to students trained in conversational French dialogues as presented in modern textbooks and on structural laboratory tapes. The tests were therefore loaded in favor of students using traditional, non-audio-lingual materials. Keating's results, predictably, showed the non-laboratory groups to be slightly better in these tests.

We in the language-teaching profession are aware by this time of certain facts concerning the laboratory. Among these are that (a) the laboratory provides intensive active *practice* of things presented in classes; it is not a self-instructional teaching machine; (b) one hour per week of laboratory use is quite general, however much we might like to add more time; (c) time used for laboratory is deducted from the total available language class time; (d) time spent in the laboratory doing four-phase pattern drills is *all* effective contact time, and it represents a notable advance in economic use of the student's time. Now Mr. Keating displays amazement at these things as he "discovers" them, and is able to make them sound like disadvantages by appropriate use of the word "only." For example: "The typical program entails only one period a week with the laboratory" (page 14). "Only one district reported that the time a student spent in the language laboratory contributed to an actual increase in total regularly scheduled class time spent with the target language" (page 13). Yet it is generally accepted that, however much we would like to add more contact time, curricula are not indefinitely expansible, and that the language laboratory must fit in during the number of hours regularly available. Why the "only" tags?

On pages 13–14 we read: " . . . it was reported that the time regularly spent in the language laboratory actually replaced regularly scheduled class time." This is normal. To con-

tinue with the quotation, " . . . What this all adds up to is something much less than foreign language instruction permeated by laboratory use." Thank goodness for that. I don't believe any language teacher or proponent of laboratory practice wants permeation. Just adequate, carefully planned, active, authentic practice of what has been presented and developed in class. At a maximum and under (unrealistically) ideal conditions, the laboratory would be used on alternate periods only. The class in between exploits and makes practical use of the material drilled in the laboratory, a d the class is used to prepare for the next practice. The use of a language laboratory as an auto-instructional teaching machine is not contemplated. (Experimentation is going on in certain universities, but it is still far from perfection or adoption.)

Finally, what are the conclusions constructed by the author of this report? First, the *laboratory groups were better* than non-laboratory groups *in speech production* (in spite of the quality of that test). As expressed in slanted English by this impartial researcher, it reads, "In only one instance, that of *speech production scores* at level I, was there found a significant difference that favored the language laboratory group." This is one-third of the three null hypotheses tested, and the only one of the three in which real competence is claimed for the laboratory, anyhow. Present laboratory technology and methodology is well-developed for the beginning courses, and particularly for the audio-lingual (in contradistinction to the literary) skills. This being the case, Keating has, in effect, confirmed that *the laboratory is doing what it is supposed to do;* that students who do use the laboratory are better than those who don't in the speaking skills. This conclusion is so cloaked in obscuring language that it may escape the attention of the lay reader.

A second result: As practice in the laboratory continues, the skill of speech production diminishes. The longer one practices, the worse one becomes. To quote Keating, "This single difference (preceding paragraph) in favor of the laboratory group on the speech production test disappears at Level II, and at Level III, the no-laboratory group is superior." We can conclude from this the absurdity that practice makes worse(!). The groups were all tested during the same year; who knows what kind of drills or laboratory work the Level II and III people had? By the second or third year it is quite conceivable that advanced students can pronounce simple nouns and even add the appropriate article. Yet one would think that an advantage gained in the first year would continue to be an advantage in higher levels, provided adequate practice was maintained. Keating's result defies common sense. Was the practice afforded in class superior to that used in the laboratory programs? If the materials in the laboratory were poor, this is quite possible. The researchers refuse to divulge the names of the participating schools, so nobody can find out what these schools meant by "laboratory," nor what kind of programs they used.

A third result: Intelligent students suffer from the use of the language laboratory. Quoth Keating, "High I.Q. students were found to be the most severely disadvantaged by the inclusion of the laboratory in the instructional program" (page 39). Keating thought it well to divide the tested groups according to their I.Q. scores, which more often than not reflect the quality of a subject's vocabulary and logic. Students with high I.Q.s are frequently graphically-oriented, and their affinity for logic would enable them to take good advantage of the traditional grammar-translation methods. These students would also naturally excel on the ETS tests dated 1940 and 1955, made especially for them—and for college-bound academic students. Now if some of these students had five hours in class, and others had only four hours in class (the other hour being given over to the laboratory), it follows that the latter would be set back by 20% of their class instruction. If the laboratory programs were not based on class work, or were not structured, or otherwise constituted a waste of time, it is easy to see how this would be a disadvantage to the high I.Q. students using the laboratory.

The impressive (though rather esoteric) mathematical tables appended to this report, it must always be remembered, are based on the *French Speech Production Test* and on two inappropriate tests antedating the advent of structural linguistics to the classroom. Yet in spite of all the deficiencies of this "research", and in spite of negativistic and obscurantistic language, the

inevitable conclusion one reaches in this report is this: *the language laboratory is most effective in beginning courses in teaching the speaking skills. Students that use it are superior to those that do not in this regard.* That's what we all thought to begin with.

The New York City Report

An experimental study of laboratory effectiveness was made by the New York City Board of Education, with the principal investigator being Mrs. Sarah W. Lorge. A very large staff was used, including mostly language personnel of the schools and technical personnel attached to the Bureau of Audio-Visual Instruction. There were, in fact, two studies. The first was to compare laboratory with non-laboratory students in French; the second was to determine the relative effectiveness of four different types of language laboratories (with a non-laboratory control group).

The existence of the second of these two studies is *prima facie* evidence that the investigators knew what a laboratory was, as did their official positions in the language branch of the New York City School System. Programmers, script writers, and tape production personnel formed a part of the experimental group.

The first study on the effectiveness of the language laboratory, carried on with good design and controls, was operated over a period of three years. It showed that beginning students using the laboratory were significantly better in speaking skills than non-laboratory students. Ninth-graders (in first-year French) profited most from the laboratory; they were the youngest and were just starting the language.

That is where the laboratory is most effective. Students at higher levels showed little or no significant difference between laboratory and control groups. " . . . yet with all classes using the same materials the laboratory students developed the skill [fluency] to a significantly greater degree." Failure of the third-year French students to profit from the laboratory was considered as follows: "It is possible that students had already developed fixed pronunciation habits in the two years prior to their entrance into the laboratory in 1961. Speech habits are hard to change." If students were not introduced to authentic speech habits in the laboratory during the first year, and enabled to perfect this skill in succeeding years, it seems that the advantage of the laboratory is lost to a large extent. We conclude that the advantage must be gained with laboratory practice at the outset and maintained throughout the course of study.

The Lorge Report states in its conclusions, "Gains which were made by laboratory groups in speech and listening skills were achieved without loss in traditional skills as measured by a standardized French test." Furthermore, students using the laboratory were motivated to continue voluntarily in language study by impressive percentages. In a group of third-year students, 74% of laboratory students elected the optional fourth year, while only 45% of non-laboratory students did so. The students recognize progress, which they enjoy; and they shun activities which are too often non-productive.

The first study bears this important conclusion:

> **Laboratory students gain just as much traditional skill as those in standard classes, plus added ability in the speaking skills.**

The Second Study was a test to compare four types of laboratory in competition with a non-laboratory control group. The types compared were:

(1) Fully-recording-playback, 20 minutes daily
(2) Audio-Active, 20 minutes daily
(3) Fully-recording-playback, once a week
(4) Audio-active, once a week
(5) Non-laboratory control group

Using a variety of good tests and objective administration procedures, results were tabulated at the end of each course. Unlike the Keating project, the method of conducting classes and the materials used were prescribed and monitored to ensure that all students received parallel instruction, with the exceptions being the *type of laboratory* used. Fourteen variables were rated: Pronunciation, intonation, fluency in reading, imitation, fluency in answers, appropriateness of answers, correctness of answers, global rating of speech, Coop I:

silent reading; Coop II: vocabulary; Coop III: grammar; Coop total; listening—fast; and listening—slow.

The results showed that Group (1), using fully-recording-playback equipment daily, rated either first or second in all the fourteen skills except 'slow listening.' Group (2), using audio-active equipment daily, rated generally second. It is clear that daily use, spread over a period of time, is preferable to the same amount of time just once a week. "The kind of equipment as well as the time spent in the lab made a difference in gains of speech." And all the laboratory groups were better than the control group in the global rating of speech.

The once-a-week laboratory groups fell behind the control groups in traditional skills, with the conclusion that " . . . unless a certain amount of time is spent in laboratory work, significant gains cannot be expected." This is important information for supporting our contention that additional time should be allotted to the language-teaching program for laboratory time. Many universities have already made this gain by increasing the traditional three-hour course to five or even eight hours a week. Language is a skill that takes long and intensive practice.

The big surprise (to me) was that the recording function in the laboratory made a real difference. Yet Group (1) with the recording laboratory surpassed the audio-active laboratory group (2) on the same time-schedule. "The Recording-Playback Daily Group achieved greater gains than any other in comprehension of rapidly-spoken French. Possibly the close attention required to compare model and imitation may have a favorable effect on comprehension." The psychological advantage is the increased interest of the student upon hearing his own voice; he pays closer attention to it under these circumstances.

The New York City report concludes that "Good results demand equipment of good quality with potential for a variety of learning experiences; teachers skilled in handling equipment; materials prepared specifically with regard to the goals of the course and techniques of laboratory learning; and careful allotment of laboratory time." This study shows the advisability of spreading practice in small amounts over a period of time, rather than lumping a large amount of practice at one time. The speaking and hearing skills of the laboratory groups as a whole were better than the control group, thanks to the intensive authentic practice afforded in any laboratory, regardless of equipment type. But the Lorge report shows that type of equipment also is influential in improving results.

Since the laboratory groups were on a par with the control groups in graphic skills, and superior to them in the speaking skills, this research indicates the positive contribution made by a well-operated language laboratory.

Conclusions

Both of these reports show that the language laboratory is making a positive contribution to language teaching. The Keating Report forces this conclusion despite tenuous premises, lack of definition of what is being investigated, and total ignorance of what went on in classrooms or laboratory. In spite of ill-correlated tests and misleading language in the presentation of the report, it develops that the language laboratory is doing precisely what is expected of it. The danger of the Keating Report is that it looks authentic and impressive, and that it is so worded as to make advantages seem to be catastrophes. It would appear from the prefatory remarks and from the negative kind of exposition that it was intentionally so designed. Fortunately the Lorge Report (available from A-V Learning, B.A.V.I., 131 Livingston Street, Brooklyn 1, N. Y., 15¢) is a well-documented and effective answer giving impressive evidence of the usefulness of the language laboratory.

A Critique of the Keating Report

JOHN J. PORTER AND SALLY F. PORTER, *University of Wisconsin-Milwaukee*

WELL designed and well analyzed research is desperately needed in foreign language learning and instruction, particularly on the effective use of the language laboratory. In light of the shortage of trained language teachers, there is little doubt that the laboratory, or some similar instructional device, is necessary if foreign languages are to be taught at all. In examining the study by Raymond F. Keating the pitfalls accompanying inadequate research on language instruction become quite apparent.

To analyze a study one must first decide what is being studied. Dr. Keating's book is concerned with the effectiveness of the language laboratory "as presently used." This last phrase is very important in the interpretation of Dr. Keating's results, as will be later pointed out. Dr. Keating has used Alfred S. Hayes' definition of a language laboratory as "a classroom or other area containing electronic and mechanical equipment designed and arranged to make foreign language instruction more effective." (page 1) Dr. Keating has accepted such a vague definition and has made no attempt at strict standardization of the equipment used for his study nor of the materials used in the language laboratories. He says, merely: "In almost every case it was reported that both commercial and teacher-produced materials were used for basic programming. In most cases no estimate of the time required by the teacher to prepare such materials was reported." (page 13) An additional point that needs clarifying concerns the amount of time possibly spent by the non-laboratory controls in listening to tape recordings, records, and other varieties of audio-visual stimuli.

The second question that arises involves the nature of language learning, defined by Dr. Keating as reading and listening skills and speech production. Before proceeding it should be noted that the majority of the schools studied removed one class period from the regular class time for practice in the language laboratory. Thus the laboratory subjects could be expected to perform less skillfully in reading tests than the non-laboratory controls since they spent $\frac{1}{4}$ to $\frac{1}{5}$ less class time on such materials. It should also be pointed out that reading proficiency is not an aim of the language laboratory, which stresses comprehension and production of speech. Proficiency in speaking a foreign language entails several skills, rhythm, syntax, vocabulary, and pronunciation, to say nothing of the art of conversation. Dr. Keating tests only pronunciation which he admits is an inadequate criterion. Despite this, he has given this test an emphasis equal to his tests of reading and listening skills in the interpretation of his results. His excuse that other writers have accepted this limitation in like manner, does not add weight to his conclusions.

Dr. Keating declares that "this study cannot be considered an experiment in any proper sense." This is a rather strange statement to make about a project involving more than 5000 students and the construction of a special test. If it is not an experiment, what is it? A preliminary investigation? With 5000 subjects? It would appear that Dr. Keating, by such statements, is trying to avoid the rigorous standards and critical evaluation which must be applied to research as extensive as his.

Dr. Keating's work utilizes levels of intelligence as the control variable. Since he admits using several different I.Q. tests to constitute his levels of intelligence, the effects of such a procedure on his variables must be examined. If these intelligence tests correlated perfectly so that a given score on one test would be exactly the same on another test, it would be legitimate to constitute levels of the I.Q. variable as Dr. Keating had done. Since he fails to identify the tests, it seems obvious that this is not the case, and we have no guarantee that a score in level 3 in one test would not fall in level 2 in another test. In fact the procedure used by Dr. Keating is statistically incorrect and serves both to confuse the results of his investigation and obscure any differences that did occur between his groups.

Another problem inherent in Dr. Keating's research concerns the method of testing the outcome of the various possible comparisons in

this non-experiment. As constituted by the author, his study is a three dimensional Factorial Design (Lindquist, 1953);[1] the factors being I.Q. Levels, Levels of French Attainment, and Laboratory or No-Laboratory procedures. Such a design is properly analyzed by the appropriate analysis of variance to determine the main effects of each of his three factors and their interaction. If these tests prove statistically significant it is then permissible to use *t* tests or some other two-group comparison as Dr. Keating has done. As the design stands, no attempt at the analysis of variance was made, and indeed, with the vast difference in size of the various sub-groups in his study, no such analysis seems possible on the present data.

Again, a question arises when one considers the large sample size utilized by Dr. Keating in his research. When 500 to 1000 subjects are used for an experimental comparison it is time to consider what it means, in a practical sense, to have a D_m of 1.83 on tests which yield standard deviations of 10.03 and 8.94 (Keating's Table I—p. 50). It is well known that, given a sufficiently large sample, any difference, no matter how small, may be shown to be statistically significant. To quote McNemar (1962):[2] "a statistically significant difference doesn't necessarily mean a difference either of practical significance or of scientific import. Sometimes a 'what of it' is not an impertinence" and " . . . on the other hand if a real difference is so small that its statistical demonstration requires thousands of cases, we may question whether it has practical or scientific importance." This question needs to be asked of Dr. Keating's study.

Finally, an examination of the data in Appendix A of Keating's work yields some rather puzzling figures. If Table I (p. 50) is examined it will be noted that there are 5130 subjects tested. If tables III through VI are examined, and these apparently represent a breakdown of Table I, it is apparent that only 3987 subjects are present. Where did the remaining 1143 go? In level 5? It is not clear. Furthermore it is instructive to compare the total number of subjects (2240) of Level 1 in Table I with the total number (1743) of Table III, which represents a breakdown of level 1, Table I, into I.Q. classes. Where did the 497 subjects go? Similar exercises in elementary addition will yield similar questions about other Tables in Keating's presentation. As a suggestion, it would be helpful to other interested investigators if Dr. Keating would file his raw data with the American Documentation Institute of the Library of Congress, so that some of these questions could be answered.

Summary

Dr. Keating has presented a large investigation which purports to demonstrate that language laboratories, as used in the New York Public Schools, are less effective than non-laboratory instruction. It is the present authors' contention that, while Dr. Keating's results reflect on the inadequate use of language laboratories in the school systems tested, it is in no way clear that any of these results bear upon the general adequacy of the language laboratory as an instructional device. Indeed, with the numerous problems in the materials chosen, the small differences obtained in group performance, and the inadequate analytical techniques, it would be best to conclude that an experimental design which tests the effectiveness of language laboratories remains to be constructed. A design for such a test based on verbal fluency is suggested below:

1. Using a standardized set of verbal materials for language laboratory instruction, compare, after one year of instruction, the effect of the use of language laboratory materials with a standard control group. Rank order the speech recordings of each group, using independent judges who are not familiar with the test constructed or the group to which the subject is assigned.
2. Control for I.Q. differences by administering one standard test to all pupils and use *two* I.Q. levels, rather widely separated, for example, from plus one standard deviation to plus two, on the test for high I.Q. group, and from minus one standard deviation to minus two standard deviation for the other group. There should be an equal number of subjects in each of the ex-

[1] Lindquist, E. F., *Design and Analysis of Experiments in Psychology* (Boston: Houghton Mifflin Company, 1953).

[2] McNemar, Quinn, *Psychological Statistics* (New York: John Wiley and Sons, Inc., 1962), pp. 69 and 86.

perimental groups, preferably between 25 and 50. Using the analysis of variance analyze for the effects of I.Q. and the effectiveness of the laboratory as rated by the independent judges. Examine the interaction between these variables.

This is an initial and basic design from which further research may proceed. While it is not as comprehensive as Dr. Keating's design, at least this study would permit adequate evaluation of *some* of the effects of the language laboratory on learned behavior.

*Review and Criticism**

EUGENE W. ANDERSON, *South Carolina Department of Education*

IF MY remarks bear a closer resemblance to an "explication de textes" that Zola might have uttered on the legal proceedings of the Dreyfus case than they do to the dispassionate type of book review normally heard in learned conferences such as this, it is because Keating's report and even more particularly the prepublication research bulletin by which its conclusions were disseminated to the provinces represent, if unanswered, a very real challenge to a most profound commitment on the part of the language teaching profession, not to speak of the considerable financial investment already made by understanding school and college administrators. I shall not even object if one chooses to brand this *lettre provinciale* as SAMLA's *Défense et illustration du laboratoire de langue.*

What is rare or quite so traumatic as a day in June when a foreign language supervisor lets ham and eggs grow cold over a bombshell appearing in the editorial columns of South Carolina's largest and presumably most influential newspapers:[1]

In a "carefully designed and carefully controlled" experiment, according to the main burden of the editorial, a researcher by the name of William Vincent had made the startling discovery that non-laboratory groups of students outperformed their laboratory counterparts on all four levels of high school French in every instance.

The editor went on to conclude that now, tragically, the nation was finding out that hard-earned tax dollars were being worse than wasted, since Federal matching funds had encouraged local districts to adopt programs resulting in lower quality education.

Any fleeting disbelief I might have entertained about the authenticity of the editor's information was quickly dispelled by the funereal finality of his proclamation that the source was no less a citadel of educational investigation than the Institute of Administrative Research, Teachers College, Columbia University.

The frustration experienced in that dark moment brought back all too clearly the sophomoric inspiration behind the following poem which, though lacking in any real literary merit, had raised me to the status of poet of the day in our creative writing class:

Dr. Doe says it's so;
The truth it must be.
Though below we bid go
Joe B. Doe, Ph.D.,
Dr. Doe says it's so;
The truth it must be.

The fact that the May research bulletin, together with the July appearance of the report itself made clear that Keating, not Vincent, had carried out the language laboratory study; that the study, in the words of the researcher, "cannot be considered an experiment in any proper sense;" and that the laboratory students were significantly superior in at least one important instance, speech production at the first level—all of these later learned facts were of little comfort on that anniversary of D-Day plus one.

To derive any astrological meaning from the unexpected assault, I had to reverse roles a little and compare the language teaching profession to General Rommel, who had just nine infantry

* Presented in large part at the South Atlantic Modern Language Association Annual Meeting, Atlanta, November 16, 1963.

[1] The editorial from *The State*, (Columbia, South Carolina) is reproduced as an appendix to illustrate the effect the "Keating Report" has had on the general public.

divisions on hand in the Normandy sector and was caught with his panzers down to one.

You may think I am exaggerating the impact of this report on the educational world, but I have here a monthly bulletin placed just last Wednesday in the hands of the approximately 1,500 teachers of the Columbia, South Carolina, schools, in which doubts regarding the language laboratory have reached the grass roots of public instruction—hence the urgency of some rebuttal.

Any analysis of the Keating study would be incomplete without placing it in proper perspective in at least two ways: (1) the setting in which the piece of research was inspired; and (2) its chronological relationship to recent developments in language teaching.

Launched by the Institute of Administrative Research, Mr. Keating's study is one of two "probes" relative to federal aid and its effects on instructional programs, within the over-all context of "general" (and presumably unrestricted) aid versus the type of "selective" or "special" aid as exemplified by such programs as the National Defense Education Act.

The report summarizes the results of tests administered to 5,000 high school French students in 21 districts of the Metropolitan School Study Council. All four levels of French were tested in May, 1962, for achievement in reading and speech production, and Levels II, III, and IV were also given listening comprehension tests.

Principally the study was designed to compare the achievement of students whose French instruction included the laboratory experience during the school year 1961–62. The researcher also drew inferences from selective data which took into account I.Q. groupings, "high" versus "low" expenditure schools, and advanced students whose laboratory experience extended beyond the one year.

Results showed that in only one instance, speech production at Level I, was there found a significant difference favoring the laboratory group. Significant differences favoring the non-laboratory group predominated and appeared with each language skill tested.

Let us assume that no significant flaws are to be found in Dr. Keating's research design, the testing instruments, or his conclusions. What does it show when silhouetted against the main lines of recent instructional development as generally established by school administrators and the foreign language teaching profession?

The following statement appears in the 1959 Council of Chief State Officers *Purchase Guide*, which has served as something of a reference Bible to schools in their connection with NDEA:

"There is an almost total lack of adequate teaching material suited to the new methods being adopted in modern foreign language teaching. Traditional 'book centered' materials, even when accompanied by recordings, are never more than half satisfactory. We may as well face it, there will be several years of trial and experimentation before adequate material will be generally available. Meanwhile, teachers and administrators should be patient, reserve judgment on the effectiveness of electronic installations, and not expect them to produce as good results as they will certainly produce when equipped with good materials."

In an article entitled "1959 Revisited," appearing in the November, 1962, issue of *Audio-visual Instruction*, a publication of the National Education Association, Professor Frederick Eddy looks back upon the above appraisal as follows:

"It is a source of great satisfaction to say that in 1962 the entire 1959 answer to this question must be re-written. Well-led and powerfully stimulated and financed by Titles VI and VII of NDEA, enriched by the knowledge and efforts of many individuals and institutions, the profession has changed the whole materials climate in three short years. Now on the market and in wide use are the beginning levels of several sets of sound, effective materials for the languages commonly taught in American secondary schools. Not the least of their many qualities are, first, that they make the close integration of class and lab work inevitable, and second, that their aim is the efficiently ordered development of all four language skills: understanding, speaking, reading, writing. The succeeding levels of these sets are appearing regularly and competing sets have been announced. It is a very encouraging picture."

Chronologically speaking, Dr. Keating's report serves to document with remarkable

precision this 1961–62 materials break-through and its impact upon the beginning level. (Professor Eddy's expression "beginning levels" assumes the singular form when we slide the scale of time back a year to the period covered by the survey.) The laboratory students do, in fact, show significant superiority in the audio-lingual portion of the evaluation, such as it is.

What about the rest of the evaluation as it touched the first level, that is, the reading comprehension test? Here, too, through the Keating study, the language teaching profession adds the gift of prophecy to its gift of tongues.

In showing a 10 per cent lag in reading comprehension on the part of Level I laboratory students, Keating confirms the results of an experiment reported back in 1959 by Johnston and Seerley in the U. S. Office of Education bulletin, *Foreign Language Laboratories in Schools and Colleges*.

In that study, covering the first semester of beginning French, the experimental classes represented four high schools employing audio-visual—now commonly known as audio-lingual or audio-visual-lingual—techniques and materials. The control classes were in four other high schools using text materials considered standard at that time. The experimental group proved to be even more markedly inferior to the control group in reading comprehension than was the case of Keating's study.

But the overwhelming superiority of that pre-NDEA group in the listening and speaking skills, together with a modest gain in writing effectiveness, apparently offered a most formidable argument in favor of the newer total approach to foreign language learning. The promise offered by such experimentation in the late fifties was so great that educators decided to discount the relatively small loss anticipated for reading effectiveness in beginning courses.

High school language laboratories have been increased from a total of 64 as reported by Johnston and Seerley in 1959 to an estimated hundred times that number today—in quest of communication skills which will not only deepen the educational experience of millions of young Americans, but may well play a decisive role in the nation's free world leadership during the coming decades.

Since Keating's study, even without our examining the limitations he himself concedes, confirms so clearly what was anticipated for that time when integrated, well-programmed language laboratory materials would begin to appear in commercial quantity, further discussion might seem academic.

A closer look at the report, however, reveals not only the researcher's failure to see all of the limitations and to appreciate the full significance of what he perceives; it may bring into clearer perspective one or two major assumptions generally made by the foreign language teaching profession.

Thus, in a very real sense the Keating report can serve as a sort of Rosetta Stone through which linguists can not only come to understand the hieroglyphics of another culture; they can also do a little sharpening up on their Greek.

From the curricular viewpoint, the unpardonable sin of the researcher is his omission of a listening comprehension test at the beginning level. The full impact of this omission emerges when one examines New York State's official guide, *French for Secondary Schools*, published in 1960.

This nationally respected publication recommends that approximately 65% of class time be devoted to audio-lingual activity, while reading is allotted not more than 20% at the beginning level of a four-year sequence.

In order to dispel any uncertainty regarding the emphasis to be placed on listening and speaking, the New York guide further suggests that audio-lingual activity based on reading material be included in the class time devoted to reading (that is, not subtracted from the audio-lingual block of time).

The common assumption is that a language learner will have considerably more listening than speaking experience. With a recommended 65 per cent of class time spent in audio-lingual activity, it would be conservative to estimate that 45 per cent of the total time would be consumed by the listening activity.

It is not surprising, then, that howls of amazement resulted from the appearance of last May's issue of the *Institute of Administrative Research Bulletin*, which trumpeted the study and its imminent publication.

Since John Q. Public is very rarely lured from

his TV set to the spicy treasurers of the local library's curriculum guide shelf, and since the writing skill was not evaluated in the survey, the apparent promotion of the reading skill to two and a half times its recommended importance at the first level would not be noticed, much less questioned. And unless John Q. has seen a language laboratory in operation, he might easily overlook the fact that the equipment had not had the opportunity of showing what it could do in regard to that one skill with which first year students in a modern program have the largest number of contact hours.

Nor does there seem to be any rational basis for not constructing a listening comprehension test where there was no suitable one commercially available for the first level, since a four-level speech production test was specially devised.

The comparative difficulty of constructing, administering, and scoring objectively valid listening and speaking tests could well be the topic of a Ph.D. dissertation. But the usually dependable cost-quality correlations will help us more at this point. The Educational Testing Service in its individual break-down of prices for the seven pillars of teacher competence, lists the following:

Test and services for listening comprehension	$1.50
Test and services for speaking	$6.75

What about the important speech production evaluation as carried out in the study? I am sure most of you have first-hand knowledge of the various ingredients that can go to make up a comprehensive type of speech production test—such aspects as pronunciation, intonation, active vocabulary, mastery of a constantly growing number of grammatical structures, and unhesitating response to visual or spoken stimuli.

While Keating does not deny the possibility of a great many other instructional benefits to be gained through the use of the language laboratory, which were neither identified nor measured in his study, he apparently fails to see how severely the validity of his speech test is compromised when applied to four levels of French instruction. It hardly requires the training of a linguist to see how ludicrous the rather imposing title, *Metropolitan School Study Council French Speech Production Test*, becomes when the instrument is scrutinized in relation to each of the levels. In essence, the test accomplished the following evaluation:

Level I—Quality of utterance of 10 critical French sounds
Level II—Ditto.
Level III—Same thing.
Level IV—You guessed it.

In order to confine the speech production test as much as possible to the pronunciation aspect, only French words ranking high in frequency listings were used. Keating describes the total test as one which actually requires the student to produce each of ten critical sounds on two different occasions: once in isolation as part of a single word, and again as part of a word used in a brief French sentence *read* by the examinee.

Proponents of the language laboratory do not deny the importance of pronunciation of individual sounds in the development of audio-lingual skills. Indeed, they would maintain that completion of upper levels of programs now under development will bring some improvement in this skill, through both extensive practice and intensive drills.

What most of the nation's foreign language teachers will say, whether traditional or modern in their approach, is that a reasonably high pronunciation plateau should be reached by the end of the first year, and certainly no later than the end of the second year. Good teachers are practically unanimous in insisting that early training for pronunciation be intensive. If this training is intensive and good, the learner can expect to move up very gradually from the plateau he reaches during the first year; if it is intensive and poor, significant improvement beyond the first year seems out of the question unless instruction is drastically altered. In his widely used *Introduction to French Speech Habits*, Pierre Delattre warns the teacher with all Gallic clarity that good habits must be taught before the learner inadvertently falls into bad ones. Only drills of an intensive nature can accomplish this.

Any inferences educators may have drawn from the research bulletin or the final report

regarding the language laboratory's impact on speech production should be reviewed in the light of the following considerations:

(1) The total speech production test used in the Keating survey consists of an aspect in which no quantitative progress is possible beyond the first few weeks of the beginning course, while the building of other oral skills, such as active vocabulary, mastery of verb forms in meaningful context, accurate spoken use of a constantly growing number of grammatical structures—in short, the manipulative skills one must have to communicate effectively in another language—can proceed from level to level in arithmetical and, at times, geometric progression. These other vital skills, untested in the survey, really form the heart of the language laboratory's rationale.

(2) In a good program of foreign language instruction, only a limited qualitative gain can be made beyond the first level or two in the pronunciation of basic sounds. Keating's own results confirm this. The gain in the proportion of fourth level lab students evaluated as "good" over the corresponding second level group is only 24 per cent. For the non-laboratory group, this gain from the second to the fourth level was 37 per cent.

(3) The superiority of the advanced-level non-laboratory students in uttering the ten critical sounds should not be explained away entirely on the basis that suitable lab materials were unavailable in 1961–62, important as this consideration is.

Of perhaps even greater significance is the divergence of teaching philosophies. Those New York teachers who tended to cling to traditional goals would generally be the teachers of non-laboratory classes, while those who more or less accepted the total picture of language learning, with the heavy beginning audio-lingual emphasis, would normally have the laboratory students.

In the affluent schools Keating describes, where personnel relationships tend to be more democratic than elsewhere, the best equipment and materials to be found would normally be made available in support of the French department's particular philosophy. While some early laboratory installations could have been made for exploratory, teacher-training, or status purposes, without due regard for availability of learning materials, teacher inclination, and proper training already received, it is highly unlikely, on the other hand, that audio-lingually inclined foreign language departments in these particular schools would be deprived of equipment so manifestly wedded to their stated objectives.

If teachers using audio-lingual methods in the New York area are like those found elsewhere, they are concerned at all levels with a symphony of linguistic values, no one of which is allowed, through over-emphasis, or ill-timed emphasis, to inhibit development of the other vital skills.

Traditionally minded teachers, in general, are guided by vastly different principles. Where reading and translation are held in greatest esteem, pronunciation is about the only speech aspect that receives adequate attention. And even then its primary purpose is to facilitate oral reading, not to build speech habits.

It is entirely conceivable that the traditional teacher will continue for several more years to outshine his audio-lingual colleagues in training students to pronounce the ten critical sounds of French. He can afford to have his students spend many long hours at each level in such activities as detailed mastery of the International Phonetic Alphabet, memorization of the orthographic variations represented by the phonetic symbols, and lengthy phonetic transcriptions, since he does not have to bother with the exhausting, time-consuming job of building the rest of the speech habits.

The fact that no time limits were imposed in the speech production test would

seem to favor the non-laboratory student, particularly if the absence of lab training reflected the teacher's preference for the traditional mode. A student who has acquired a given language skill with the aid of extensive analytical processes should, if allowed unlimited time for his responses, show a reasonable degree of proficiency when tested.

A main objective of the language laboratory is to afford a type of learning experience that will produce unhesitating, automatic oral responses with little or no reliance upon analytical crutches. Fifty million Frenchmen just prefer not to wait indefinitely for the correct sound or phrase to be ground out through some mechanical process. Albert Valdman has very aptly described French conversation as a private form of war requiring spontaneous and rapid-fire reactions between interlocutors—in which a two-second silence can seem an eternity (*Audio-visual Instruction*, November, 1962).

One other consideration is worth noting in regard to Keating's evaluation of speech production. For this phase of the survey a 10 per cent sampling of the 5,000 students was made—a quite acceptable procedure, statistically speaking. From the pedagogical viewpoint, the study might have been a little more useful if the same sampling had been applied to all phases, so as to determine more accurately the inter-relationships among the skills tested; for example, it would be of interest to know whether high performance in one skill on the part of a given group of students was accompanied by low achievement in another.

This, however, is a minor point. What does cause some real statistical concern is that the researcher examined students from only 12 of the 21 school districts surveyed.

I feel sure Dr. Keating and most educational researchers would agree that among the characteristics of wealthy school districts are better organization and articulation of language programs through appointment of department heads or coordinators; reasonable unity in objectives and philosophy, whether this unity is arrived at through staff compromise and consensus or through acceptance of some respected leader's views; and, based upon these assumptions, a rather uniform imprint upon the students within a given district.

In 1961–62, foreign language learning was in a state of such rapid transition that the 21 school districts doubtless represented 21 different shadings of philosophy, objectives, materials, and methods. Under these circumstances it would seem almost imperative to have had the 10 per cent sampling include all 21 districts to insure the representative quality of the sampling; that is, to preclude the possibility, in such a limited number of cases, of significantly affecting the outcome by the mere replacement of one district's learning situation for another.

Professor Keating's choice of the French language for his survey merits no special censure in the general sense, for French and the other common Western languages have served as vehicles in the large majority of research projects undertaken to date. In examining the potential of the language laboratory with specific regard to the reading skill, however, a question of considerable importance arises.

There is no need to recall that the Anglo-French cultural exchange program goes back at least to Alcuin and the Carolingian renaissance of the 8th century, even a few decades before the official birthdate of the French tongue, and extends through the Norman period, with all of its linguistic impact upon English, through long centuries of both rivalry and alliance.

And it is almost insulting for me to point out that the English language, as an instrument of intellectual endeavour— as a tree of knowledge, if you will—while it has Germanic roots, trunk and branches, bears foliage largely of Latin, French, and Greek nourishment.

Nor would it surprise anyone that my eleven-year old son, without any training in French

reading as such, could decipher with near-perfect understanding sentences like the following:

> **Marie prépare le dîner.**
> **Paul Martin arriva le 5 novembre 1962 pour participer au match de football.**

What I am saying to Keating and other researchers who may follow him is that, for purposes of evaluating the true potential of the language laboratory with specific regard to the development of reading skill, French, because of its unique relationship to English, is the worst language he could have chosen. Russian would have been better, since it does have a somewhat different alphabet and, as a Slavic tongue, is more remote from English; but the linguistics purist would tell us that any real or abstract appraisal of the language laboratory's full potential as effecting the reading skill would have to involve a language outside the Indo-European family, such as Arabic or Chinese. In a very true sense the American ninth grader who elects a French course has been reading that language from the time he received his first set of alphabet blocks from Santa Claus at the age of two.

This is not to say that the laboratory can contribute little to the skill of reading French and the other Western European languages. On the contrary, there is a growing body of theory and practice that leads us to believe the cumulative mastery of grammatical structure, as reinforced by adequate materials now here or in process of completion, may go a long way toward reducing the reading problem on advanced levels to one of lexical novelty. The laboratory even has implications for the teaching of English reading, particularly with students who have had a defective background of audio-lingual skills in earlier life.

The high correlation between educational expenditure and teacher competence, together with instructional quality in general, lends powerful support to the argument Keating and Vincent advance to show why the Metropolitan School Study Council schools should be able to make the laboratory work if any schools can. But how do we answer the haunting question of why the great majority of these admittedly capable teachers were employing a frequency pattern so at variance with what has become almost professional dogma?

The 1960 New York State guide mentioned above (which does not appear in Keating's fairly extensive bibliography) recommends a minimum of 60 minutes a week, preferably in three 20-minute sessions. A Northeast Conference committee, about a third of whose membership were from Columbia University, simply advised teachers to use the laboratory as often as possible. Frederick Eddy and Alfred Hayes, considered among the nation's leading authorities on both the technical and pedagogical aspects of the language laboratory, have in separate publications advocated daily use for the secondary school.

While both Vincent and Keating express disappointment over the prevailing once-a-week practice, they do not attempt to explain why highly competent teachers would have such a situation exist. The reader is left with the impression that this kind of use results from some local professional consensus which, in the light of other reports, runs counter to national practice.

In a recent communication from the U. S. Office of Education, Hutchinson and Gaarder report a 1961–62 study of 133 New Jersey high schools showing that something over 50 per cent of schools there had lab sessions only once a week, not quite as bad as in the schools studied by Keating. Data from 9 states, representing all sections of the country show a much closer relationship to orthodox practice, with 3–5 times per week in 37 per cent of schools; two times a week in 36 per cent, and once a week in only 27 per cent. South Carolina's pattern is reasonably close to this.

All of this seems to indicate that the teacher's use of the language laboratory more closely approaches acceptable national practice as one moves away from the metropolitan areas of New York and New Jersey. This backwardness is all the more paradoxical when one considers that teachers from this very area have played such an influential role in formulating national doctrine on effective use of the laboratory.

Since Vincent and Keating do not explain this for us, we must again consult statistics.

The September, 1962, issue of *PMLA* reported that New York State, with 37.9 per cent of its high school population taking modern foreign language courses, ranked second only to tiny Rhode Island in 1960. When we consider the characteristics of the schools and communities surveyed by Keating—high percentage of college graduates among the parenthood, heavy suburban concentration of business and professional patrons, extremely high proportion of high school students headed for college with practically no financial obstacles in the way—we could expect possibly 85 per cent of the students to be enrolled in foreign language courses. There are only one or two school-community situations like this in my own state; but the one South Carolina school most closely resembling the New York suburban, high-rent group recently reported a 92 per cent foreign language enrollment. The national modern foreign language percentage in that same year was 21.7. But the picture is still not complete. NEA figures (Research Report 1961–R22), together with the U. S. Office of Education's 1963 *Digest of Educational Statistics*, reveal that the average total enrollment of New York State's secondary schools in 1961–62 was around 815, while the national mean was approximately 483. Furthermore, since Keating was dealing, not with upstate rural communities, but with metropolitan area schools that sometimes run up to two or three thousand students, we could conservatively expect his schools to number around 1,500 as an average.

This leads us to a rather startling conclusion: Keating's schools, most of which possessed only one equipped room, may have had total foreign language enrollments of around *twelve* times the national average!

In the light of this situation, both the abundance of capable, well-paid teachers and the lack of adequate learning materials for the upper three French levels take on secondary importance with relation to any effective use of the language laboratory. With 1,000 to 1,200 highly motivated foreign language students competing for use of the one room, it becomes mathematically impossible for the laboratory experience to approach optimum frequency and contact hours.

Hutchinson and Gaarder, in the communication cited above, report that results of a recently completed research study on the relative effectiveness of various types of language laboratory experiences for second-year French in ten New York City high schools indicate the inadequency of the one laboratory practice period in achieving gains, whereas dramatic gains ensued when daily practice was afforded. A previous study conducted in the same schools revealed that positive results could be shown with two laboratory periods per week.

Accessibility of equipment and materials is probably the most vulnerable area of instructional design to the break-down of cost-quality correlations. If the Keating survey proves anything, it is that the New York area schools, as well as others similarly situated need very badly to equip more rooms with the language laboratory facility in order to catch up with the rest of the nation.

Any interpretation teachers or administrators may have read into the reports of the study to the effect that the Metropolitan School Study Council schools in 1961–1962 could either demonstrate or predict any reasonable use of the language laboratory as a model for the great majority of high schools over the country is severely compromised by the above considerations.

It is at this point, too, that the Keating study loses any real validity, nationally speaking, in the "general" versus "selective" federal aid question which inspired the survey in the first place.

If further evidence is needed to document the special nature of the New York situation, the research report itself is the best source.

Keating's comparison of "high" and "low" expenditure schools is both interesting and useful. By the designation, "high", he means that group of schools surveyed whose expenditure per pupil in 1961–62 exceeded a median of $600, while the "low" expenditure schools fell below that sum. The national average for the same year was approximately $414, which would seem to place the "high"-"low" comparison somewhat in the category of ivory towerism.

But the statistics dealing with this phase of the survey are very significant. The mean reading comprehension score for the "low" school

lab students at Level I, 13.34, actually exceeds, in a very slight degree, the 13.06 score of the "high" school lab group, while achievement of non-lab students in the "low" schools, 15.00, falls well below the 20.56 score recorded for the comparable "high" expenditure group. The drop in reading achievement from the "high" to "low" schools at Level II is about the same for lab and non-lab groups; mean scores for lab students, 27.72 to 19.50; for non-lab students, 31.84 to 23.80.

In the remaining figures that deal with this aspect of Keating's study, the trend strongly reasserts itself. Level II lab students in "low" expenditure districts scored 19.77 in listening comprehension, only 5.86 points below their "high" district counterparts, while the non-lab, "low" district students fell 8.88 points below the non-lab, "high" district French students to approximate parity (20.65) with the "low" school lab students.

Thus, if we accept Keating's findings as valid, setting aside the vast limitations we have already discussed, and if we should also agree with his assumption that the achievement differential between "high" and "low" expenditure schools results from the wealthier schools' ability to hire superior teachers whose students presumably profit less from the laboratory, what does the survey tell us?

In essence it proclaims that the comparative gain in instructional value of the language laboratory, when one moves down from the "Rolls-Royce" schools of the nation to the "Cadillac" variety, is such that administrators struggling for greater quality of learning on "Ford" and "Volkswagen" budgets should be falling all over themselves to acquire this wonderful facility.

The language teaching profession, however, will not accept this over-simplified interpretation. While it is only reasonable to assume that the laboratory can and will go a long way toward closing the qualitative gap between the great majority of teachers who have limited audio-lingual proficiency and those relatively few who passess superior command of a foreign language, it is also predictable that the over-all orchestration of linguistic skills as carried out by a superior teacher with the aid of a laboratory and adequate materials will be more effective than that of an equally superior teacher with the same materials, but deprived of this electronic aid.

Both Vincent and Keating seem unaware of the fact that the major contribution of the language laboratory transcends teacher qualifications in the conventional sense.

Without a language laboratory, what teacher is physically qualified to evaluate one student's oral responses and correct him individually, if necessary, while the other 29 students proceed simultaneously and uninterruptedly in a practice session, hearing only correct speech models and unexposed to the errors of their classmates? What teacher can, with his own voice, phonograph, or tape recorder, distribute sound evenly over a classroom so as to give all students equal opportunity in the vital aural discrimination phase of language learning?

In short, what foreign language teacher, without the aid of a laboratory, can physically provide a learning situation in which the supervised, individual experiences necessary to build language habits are multiplied many times; where each student has equal access to perfect, untiring speech models and can with high fidelity compare his voice with those of various native speakers; where each student is given the opportunity to iron out his deficiencies in unembarrassed isolation, gaining all the while both self-confidence and a proficiency little contaminated by the errors of his peers? If critics of the language laboratory can find answers to these questions, we language teachers shall be the first to recommend some other use for our instructional funds.

In fairness to the researcher, it must be stated that he clearly describes his study as not intended to show what the language laboratory can do in "some ideal and highly creative program." On the other hand, Keating fails to perceive the polarization that emerges between the schools he identifies and the ideal situation.

Yet it is in the great middle ground that NDEA supporters and the language teaching profession have staked out their claim. Dependable laboratory equipment is available, and has been for several years; highly creative learning materials are now on the market for a steadily increasing number of levels; from 25 to

30 per cent of our teachers have attended summer institutes, and many others have received some training in special workshops during the past two or three years.

It was unfortunate that the Keating study was undertaken at the time and in the manner and place in which it was, with apparently so little awareness of or attention to the professional recommendations and warnings regarding the years it would take before suitable recorded programs for several language levels arrived on the scene.

This criticism is not intended to condemn the "candid camera" approach to educational investigation, since this is essentially the way by which ongoing school programs are evaluated. All that we language teachers ask is that the investigator know fully the limitations of his camera, that his instrument be in focus—and with particular reference to curricular areas in a state of rapid transition—that he employ a wide-angle lens.

Perhaps even more unfortunate was Vincent's pre-publication research bulletin in which Keating's findings were elevated to the "general" versus "selective" federal aid debate two or three months before linguists could begin to scrutinize the full report as it came off the press.

To those who know what it is to develop creative materials for the newer total approach to foreign language learning for the high school, the study, together with the premature publicity surrounding it, calls up visions of some Shakespearean fantasy in which the new-born Macduff, "from his mother's womb untimely ripped"—with little attention given to the maternal well-being, and before any clinical judgment could confirm the infant's breathing—had already been dispatched to slay a Scottish despot.

As for the careless and/or uninformed interpretation of the Vincent and Keating reports given here and there in the mass media, this so-called public enlightenment can best be described, not as educational illumination, but, in Miltonian terms: "not light, but rather darkness visible."

In conclusion, what can we say of the Keating study? Certainly, the limitations we have considered should make clear to any thinking person that the language laboratory was not really on trial in the situation described in the survey, nor, indeed, were the audio-lingual learning goals this instrument is primarily designed to achieve.

The basic philosophy underlying the total approach to language learning is actually strengthened, almost incidentally, by our study of the report.

More than ever we high school and college teachers of the common Western languages should stop taking exaggerated pride in those of our students who perform well in the reading-grammar-translation skills only, when we know perfectly well we must share the laurels these particular skills bring us with eight or more language arts teachers and the English instructor across the hall, not to speak of the bright student's natural ability to perceive cognatal relationships.

More than ever, college foreign language departments should evaluate the genuine effectiveness of individual high school programs on the basis of a total range of language skills and cultural knowledge, and reconsider the Procrustean discipline their antiquated placement and credit policies tend to impose upon conscientious high school instructors.

More than ever—in view of the rumored delay in implementing the Presidential directive to require new Foreign Service officers to demonstrate broad linguistic proficiency—we teachers should in all humility and with all dedication appreciate the vastness of our task.

In the tricky, uncertain decades ahead, where a probable nuclear stalemate will more than ever transform the Cold War into a communications race—with all of the economic and social pressures, propaganda and subversion we must face around the world—the nation can scarcely afford any longer the luxury of half-learned foreign languages or the international condemnation of our 200 million voices of silence.

Appendix

Backwards Via 'Aid'

SURPRISING as it may be to some people, tax dollars distributed by Washington to educational institutions do not always result in better education.

There has been considerable publicity over the past few years about the benefits and improvements which accrued when the National Defense Education Act was passed and money made available to schools.

Now, years later, someone conceived the idea of checking up on some of the results. The survey was conducted by William S. Vincent, and published in the Institute of Administrative Research Bulletin put out by Teachers College, Columbia University.

Using matching funds available under the defense education act, many districts purchased equipment and established laboratories for the teaching of foreign languages. It was simply assumed that the laboratory equipment was an effective aid to teaching language.

Some districts purchased heavily of this equipment, and some poor districts diverted money from other meager programs to take advantage of the "federal dollars." In other words, some programs suffered so that language instruction could, it was assumed, be improved.

Researcher Vincent made a startling discovery. Using students studying French, he found that the achievements of the "non-laboratory" group were significantly greater at the first through the fourth year level in every instance. His experiment and testing were carefully designed and carefully controlled.

Now, tragically, the nation is finding out that the hard-earned tax dollars were worse than wasted. The federal dollars encouraged local districts to adopt a program which resulted in a lowering of the quality of education being offered students. The results of this lower quality education will cost the nation millions more in the long run than was wasted orginally.

This obviously was a case of leaping before looking, and now the whole nation has a bigger mountain to scale than before.

The State, Columbia, South Carolina, June 7, 1963, p. 10-A, reprinted with permission of the editor, H. F. Cauthen.

The Shortcomings of Language Laboratory Findings in the IAR-Research Bulletin

FRANK GRITTNER, *Wisconsin Department of Public Instruction*

THE Institute of Administrative Research (hereafter "IAR"), Teachers College, Columbia University has released a bulletin which purports to prove that students perform better without language laboratory instruction than those who are taught with it. The printed leaflet, *IAR-Research Bulletin*, Vol. 3, No. 3, was mailed out to various parts of the country before the complete report was made available. Half of the five pages are devoted to abstracting from, generalizing upon, and stating implications of the study conducted by Raymond Keating.

The practice of bulk mailing preliminary findings seems to be a modern research innovation. There is something to be said for the older approach under which such data were examined in toto by knowledgeable members of appropriate professional groups. In this case language professors, teachers, administrative heads, and others involved in the installation and utilization of laboratories might have been contacted. For some reason, the IAR group did not adhere to this policy. The study was first mentioned in Wisconsin at a meeting of the Associated Public School Systems in October of 1962. Repeated requests for data upon which the report was based produced only promises

that it would be sent. Finally, nearly a year later, our office received a copy of the bulletin described above. The following comments relate to that publication.

Perhaps the most glaring shortcoming of the IAR study is its failure to account for certain obvious variables. The language laboratory, after all, has been in widespread use for only five years. That a five-year-old device is not a mature, perfectly functioning educational instrument should surprise no one. Yet Keating makes no mention of those problems which are considered primary by virtually everyone who is involved in the administration of laboratories, namely: (1) Is the equipment properly designed and in a good state of repair? (2) Are teachers well-versed in its use? (3) Are taped materials correlated with the text and are they suitable for lab use? (4) Is there an adequate supply of such materials? (5) Are students evaluated regularly on their progress in the audio-lingual phases of language study? A study which fails to examine such matters carefully is apparently questionable. What must be compared in such a study is a given quality of non-laboratory teaching with a comparable quality of teaching involving appropriate use of the laboratory. That Keating missed this point completely is evidenced by the following statement: "Good *language teachers* using whatever devices they care to employ are superior to *language laboratories* in teaching reading and listening comprehension. . . ." Not even the most avid proponent of laboratories would juxtapose the device with the teacher. Instead laboratory advocates have maintained that a good teacher who believes in the development of speaking and listening skills can do a far better job with a properly-equipped, properly-used language laboratory. Where studies have been conducted on the basis that it is a supplementary device for developing conversational skill, the results have strongly favored the laboratory. In Brooklyn, for example, where more than 10,000 students are being taught with the *aid* of the language laboratory, a three-year study reveals that (1) ". . . improvement in speech or in listening comprehension was made without detriment to traditional skills," and that (2) the laboratory group has "superior achievement" in language fluency and intonation.[1] It is difficult to understand why the IAR report made no effort to explain or even acknowledge this and other reports *favorable* to language laboratories. Findings which present only one type of evidence when contrary evidence is available may justly be suspected of being propaganda rather than research. To avoid this appearance of propagandizing many of us in the language field have not publicized the favorable findings. The laboratory movement seems too embryonic to make final judgments one way or the other.

Herein lies a second major limitation of the IAR study; it is probably too early to expect the laboratory to have reached its full potential. Note, for example, the following copyright dates on materials which were specifically designed for laboratory use: French, 1961–1962; German, 1962; Russian, 1962; Spanish, 1963. Practically all of the materials commonly used in high school laboratories were produced since 1960.[2] Consider also that many of the early laboratories were poorly designed and difficult to service when measured against current standards. Add to this the problem of re-training tens of thousands of teachers most of whom were educated in the era of the grammar-reading-translation methodology. Consider in addition that most teachers did not even have a tape recorder five years ago.[3] And finally, take into account that the lab's potential for individualizing instruction has hardly been tapped. All of these factors seem to indicate the wisdom of giving the equipment more time to prove (or disprove) itself.

In the one area where the laboratory has had sufficient time to prove itself—the education of adults—its great value is unchallenged. In the December 1962 issue of *Hispania* (pp. 829–836), R. C. Reindorp describes the evolution of foreign language teaching in the Institue of Latin American Studies. The most relevant point from this study is that even

[1] From a study conducted by Mrs. Sarah W. Lorge, Language Consultant, Bureau of Audio-Visual Instruction, Board of Education, Brooklyn 1, New York.

[2] We refer here to audio-lingual materials by Harcourt-Brace, Holt-Rinehart-Winston, Chilton, D. C. Heath, Encyclopaedia Britannica, Electronic Teaching Laboratories, Regents Publishing Co., and McGraw-Hill, to name a few.

[3] See Finn, Perrin, & Campion, *Growth of Technology*, p. 48, Occasional Paper No. 6, 924 West 37th St., Los Angeles 7, California.

with well-motivated adults (people who had immediate need of the foreign language for business or educational purposes) optimum use of the new methods and equipment was realized only toward the end of the ten-year study. A decade of experience and analysis of results lead Reindorp to the following conclusions regarding modern methods and the laboratory: "The aim is to provide sufficient drill, practice, and repetition so that the student may develop and fix the necessary automatic speech habits . . . Like walking, 'knowing' the multiplication tables, or swimming, talking is a complex set of habits and it is a well established fact that the way to develop habits or skills is through long hours of practice. For this purpose the language laboratory becomes the peerless drill master. This is the basic purpose of the electronic equipment which is made available to the student in the room called the language laboratory."

In addition to its questionable timing and failure to control variables, the IAR study appears to contain a number of internal flaws. The first of these would be enough to discredit the entire report since, by its own admission, it measures *failure to use* the laboratory rather than *use* of the laboratory. The bulletin states that "the degree and frequency of employment of the language laboratory in the schools in which this investigation was made were found to be disappointing. The modal practice was one period per week." Incredibly this shortcoming is passed off with the remark that "it is hard to see how more frequent periods in the laboratory would change the outcome." This is in direct opposition to studies on the learning of skills which have been conducted over the past half century. Invariably distributing practice into a number of short periods has proved superior to massing it into one long one.[4] An example which would seem to have application here is the teaching of Morse Code in the military during World War II in which the seven-hour daily practice sessions were changed to four hours and were distributed over a longer period of time. The four-hour group "ended up markedly superior to the massed group."[5] Applying the IAR remark to other areas where skill development is involved one would have to say, for example, "if a football team achieves mediocre results by practicing once a week, it is hard to see how more frequent periods of practice would change the outcome." Logic, common sense, and bona fide research clearly contradict such an assertion.

A second basic flow involves the "speech production test" which appears to miss the point of what the lab was designed to do. It is described as "a test devised by Keating and James H. Williston of the Teachers College faculty to assay the quality of the student's pronunciation of 20 critical sounds of the French language." This might be an excellent test for a class in French phonetics. However, a parroting of isolated, unrelated sounds is *not* a measure of the student's ability to communicate in the foreign tongue. On the contary, the linguistic approach insists upon complete utterances. "A person has 'learned' a foreign language when he has, first, within a limited vocabulary, mastered the sound system (that is, when he can understand the *stream of speech* and achieve an understandable production of it) and had, second, made the structural devices (that is, the basic *arrangements of* utterances) matters of automatic habits."[6]

One of the main criticisms of the traditional approach relates precisely to this point; that is, while American language students in the past could produce a wide variety of correctly pronounced vocabulary items, they were helpless when confronted with a native. The student was simply unable to put the words together in such a way as to meet the demands of normal conversation.

A third basic flaw might conceivably arise from one of the virtues of laboratory instruction. The increased emphasis upon conversational skill (and who would use the laboratory for any other purpose?) tends to reduce attrition. In the Brooklyn study, for instance, it was noted that although three years of language study was sufficient to meet college entrance requirements, 71% of the students in the experimental (laboratory) group elected to continue French for a fourth year. This compares

[4] Herbert J. Klausmeier, *Learning and Human Abilities*, Harper & Brothers, New York, (1961).

[5] Klausmeier, *ibid*, p. 246.

[6] Robert Lado, "Linguistics and Foreign Language Teaching" *Language Learning, A Journal of Applied Linguistics*, Special Issue No. 2, March, 1961.

with 35% of the non-laboratory students.[7] If this were true in the IAR study, it would partly explain why the superior performance of the laboratory group did not extend beyond the first year. While the non-laboratory group was losing most of the low achievers in the second year, the laboratory group would tend to maintain a more representative cross-section of abilities, thus depressing the total test score.

A fourth basic flaw is the tendency in this study to base conclusions upon pure speculation. For instance, the assumption that "the financially well-favored" school will necessarily proceed more wisely than other less affluent systems might best be described as "naive." If modern research technique permits one to equate wealth with wisdom, could it not also permit a few suppositions in the other direction? In fact the present writer could cite evidence of "less well-favored" schools which—precisely *because* of their limited funds—tend to proceed cautiously, making certain that they have the staff, the materials, and the knowledge to carry on the drastic new program which the purchase of a language laboratory implies. A further specualtion might lead to the conclusion that wealthier districts are more inclined to purchase modern apparatus as a matter of status. Moreover, large urban schools with their larger staffs may well have more problems to solve regarding agreement on the use of modern methods, some teachers favoring traditional objectives, others advocating the newer approach. In the IAR study this might explain the meager utilization of the laboratory. It is possible that the teachers were not very enthusiastic about the equipment in the first place. The point of this is that speculation can lead in any direction. Actually, no statistical evidence has been given to support the supposition that the relative wealth of the community has anything to do with the problem. The use of such problematical statements in something called a "Research Bulletin" might well be questioned.

In summary, the study discussed above is open to serious question on the following points: (1) The low frequency of laboratory use tends to invalidate the study. (2) The manner of evaluation appears not to be consistant with the usual objectives of laboratory teaching. (3) Important variables are not mentioned, are uncontrolled, or are lightly set aside by means of questionable assumptions. (4) Other studies which contradict this one were not taken into account. (5) Since there are, as yet, no widely accepted standardized tests,[8] since teacher re-education lags behind technology, and since the profession itself has not settled upon what constitutes optimum equipment, it would seem that the far reaching conclusions drawn in this study are premature. It may be that no one has proved conclusively that laboratories are effective at the high school level. However, it would also seem, that, in the light of its many shortcomings, the findings of the IAR study have contributed little to an understanding of the effectiveness of the language laboratory.

[7] Lorge, *op. cit.*

[8] The MLA Cooperative FL Tests were not available at this writing.

Reprinted from
THE MODERN LANGUAGE JOURNAL
Vol. XLVIII, No. 4, April, 1964