

EXPERIMENTATION IN THE LANGUAGE CLASSROOM:  
GUIDELINES AND SUGGESTED PROCEDURES  
FOR THE CLASSROOM TEACHER

Harlan Lane  
University of Michigan

*Learning: Sunburn or Behavioral Change.* Scientists are beginning to develop an image of the optimal learning situation. The teacher may not be surprised to discover that current and traditional pedagogical techniques are greatly at variance with this image. Many teachers continue to be burdened with the "sunburn" model of learning. The teacher, prime source of knowledge, light, (and, occasionally, heat) "exposes" students to his ideas; they "soak it up" and, in turn, become "enlightened." Students who fail to learn are simply not "sensitive" or "receptive," they do not "see the light." A newer, more workable model is emerging from current behavioral research, a model that defines learning in terms of a *change of behavior*.

Consider the student who is about to learn French. He does not distinguish properly among French sounds; he does not respond appropriately when addressed in French; he does not produce most French sounds correctly; he cannot read French from a text; and so on. The teacher's task is to *modify the student's* behavior so that he *will* hear, understand, speak, and read French. To change the student's behavior from what it is now to what it should be: (1) the student's current behavior must be carefully assessed; (2) the desired terminal behavior must be carefully analyzed; and (3) a program must be set down that will lead in small steps from initial to terminal behavior.

Exactly what are the desired terminal behaviors in language learning? Descriptive linguistics can provide an account of the terminal behavior required for foreign language fluency. How can this terminal behavior best be developed from the initial behavior repertory of the student? Psychology is building the bridge between initial and terminal behavior by specifying programming techniques that will facilitate learning. What role

---

<sup>1</sup>This article was prepared in conjunction with the staff of project ALLP, University of Michigan, at the request of the U. S. Office of Education.

can the language teacher play? The language teacher can conduct important research within the context of the traditional classroom. Lest we too quickly abandon tried (if not true) methods and succumb to nothing more than a fad, we must use the classroom as a proving ground for new techniques. Furthermore, small-scale but rigorous research in the classroom can generate a wealth of provocative ideas and experimental findings.

A most valuable resource in improving modern language pedagogy is, therefore, you—the language teacher. This article has been written with the hope of stimulating your interest in the techniques and findings of behavioral science and in the pursuit of research in the language classroom.

*Some Questions for Classroom Research.* Each class hour can be part of a learning experiment. You introduce a controlled change in technique or content and observe a related change in technique or content and observe a related change in the performance of your students. No matter what the outcome of this experiment, if you know what you did and what your students did, you can make some positive statement. In this sense a properly performed experiment always "works."

There are no absolute rules for generating good experiments, but a recurrent feature is that the experimenter is interested in the experiment; he is curious about a question that the experiment will answer. Perhaps some of the following questions will seem interesting to you, worthwhile asking and answering, and will prove suggestive of other experimental questions.

1. What would happen if . . . your Russian students learned Cyrillic script from a specially prepared program? When you are ready to teach orthography in your course, you section the class at random into three homework groups. Group A, the control, is assigned the task of copying the dialogue appearing in Cyrillic in the textbook; they are to do "the best they can" and to hand in their work the next day. (This may be the technique you are using now.) Group B learns Cyrillic script from a "program" that you specially prepare.

Here's how you might do it: Bear in mind the writing skills that the student now possesses and those that you wish to develop (the "terminal behavior"). Based on your experience as a teacher write out a sequence of symbols in increasing order of difficulty. The first symbols may not be Cyrillic "letters" at all, but parts-of-letters that are not difficult to draw. Do not be afraid of too slowly increasing the difficulty of the symbols you choose. (Almost every programmer begins by increasing the difficulty of his teaching program too rapidly.) After

this sequence of parts-of-letters and letters is completed, join the letters into groups of two's and three's, then into words and, finally, sentences. This is your "program" for teaching Cyrillic script.

To arrange that the students' behavior have some consequence at each step you might try this: Write all the symbols in order on index cards (and number them). Leave every other card blank. The student is instructed to examine the stimulus card, turn it over, write his response on the next (empty) card, and then compare the two. Then, he is to go on to the next stimulus card and proceed in this manner through the pack. On the following day, the student turns in his work so that it may be graded. Your third experimental group (C) can do both: work through the script and copy the text.

Your "independent variable" was the script program. What shall be the dependent variable--what change in behavior should you measure? Perhaps someone, unacquainted with the experiment, will grade the work of the three groups for you and you will compare their average grades. You may also use other measures of learning. For example: By administering a writing test at a later date, you can determine how the three groups compare in their ability to *retain* the writing skills they have mastered.

The time required to do the homework should be roughly equal for the three groups. If you control this variable it will not confound your results. This is an example of exercising experimental control. It is reasonable to assume that time spent in learning script, by whatever method, affects performance on a writing test. Let us say that your three groups learned script by the different methods and also spent different amounts of time in learning. Suppose, that the group scores on the writing test were found to be different. Are these differences in score due to different learning methods or to the different amounts of time spent in learning? As you can see, the effects of these two variables—method and time—would be confounded in your results.

What would happen, after all, if your Russian students learned Cyrillic script from a specially prepared program?

2. What would happen if . . . one of your Spanish classes learned the first three or four beginning dialogues from a text that had numbers in place of vowels? Since English and Spanish use similar written symbols, you may have observed students who use English sounds in response to the letters in their Spanish textbook. One way of preventing this transfer of English speech habits in the reading of Spanish text is to remove the

stimuli that elicit the English responses--namely, the letters common to both languages. You might want to try a completely new, arbitrary symbol system. Short of this, the present experiment proposes that you try removing the most common symbols and source of trouble, the vowels. Copy the first few dialogues in the text onto a mimeograph stencil, substituting "1" in each place that "e" occurs, "2" for "u", and "3" for "a", and so on. As you have done perhaps in prior courses, read the Spanish materials aloud (you *may* need the original text for this) and drill your students in pronunciation. If you have a second class using the unaltered textbook, these students may serve as a control group. The details of the experimental design and the choice of a dependent variable are left to you.

3. What would happen if. . .you taught French vocabulary with pictures? One group of subjects learns French words in response to pictures only; a second group learns French words in response to their English "equivalents;" a third group is presented with both the pictures and the English words when learning French vocabulary. How would these groups compare on a subsequent vocabulary test? How would they compare on a re-test several weeks later? (Or, better, how would they compare if the first test were postponed a few weeks?) And, incidentally, how would the experimental and control groups compare on a test of pronunciation for these words?

4. What would happen if. . .you used the SRS models in preparing your language laboratory tapes? In line with our earlier distinction between two conceptions of learning: sunburn vs. behavioral change, you may now be merely "exposing" your students to a second language in the language laboratory. What would happen if your tape recordings were prepared in this manner: first, the acoustic stimulus (S) (an isolated sound, a word, phrase, or sentence), then a pause during which the student gives an imitative response (r), then a repetition of the stimulus (S), yielding "confirmation." Again, the details of design and choice of a dependent variable are left to your ingenuity.

5. What would happen if. . .

(Left blank to be filled in by the reader)

*On the Significance of Results.* Since you are actively engaged in language teaching, you probably have an image of the ideal language-learning situation and you may be convinced that it exists rarely, if at all, in our classrooms. (As indicated earlier, the psychologist will readily agree.) If you are willing to allow that there is great room for improvement in language teaching, you will probably agree that the only important changes in technique are those that show dramatic effect. At this point in our knowledge, changes in the learning situation that produce marginal changes in behavior are not significant—in the sense that they are not very interesting. These "small effects" may, however, encourage you to further research along the same lines. Small effects often grow to become large ones when the experimenter "refines" his technique and extends his control to more of the learning situation.

In addition to the "size" of an effect, there are other criteria you may take into account in estimating the importance of your findings. "Reasonableness" is one. Do the results of your experiment "make sense"? Do they agree with other experimental findings? If they do not, you may be on the verge of a new discovery and will want to check up on it with further research. More likely, however, you have made an old discovery—some uncontrolled variable is wreaking havoc. As an example, consider the experiment on programmed learning of Cyrillic script. You will remember that Group A copied the text, Group B received the script program, and Group C did both. Suppose that, in the writing tests, Group B did the best, Group A second-best, and Group C poorest. These results don't quite "make sense;" you may wonder how to account for them. If programmed learning (Group B) is better than copying (Group A), why should both combined (Group C) give poorest performance? One possibility is that Groups A, B, and C were not truly comparable before the beginning of the experiment and their penmanship grades reflect two confounded variables: learning method and prior skill.

You will observe that the criteria for importance of results, how dramatic are they and how reasonable, draw heavily on your experience as a language teacher and on your knowledge of psychology and linguistics. There is no other course; it takes experience and knowledge—that is, sophistication in your field—to assess properly the importance of your findings.

The size of an effect and its reasonableness tell you something about its *reliability* too. A reasonable but small effect will probably turn up again in the same or similar experiments. A reasonable and large effect is even more likely to recur.

If another person who does not share your private sophistication wishes to assess for himself the reliability of your findings, he has two courses open to him. First, he may replicate your experiment and see if he gets the same results. Alternatively, he may use a public criterion of reliability, employing statistics. Many psychologists publish statistical tests of their findings along with their report of research with this reason in mind—to aid the unformed reader in arriving at an opinion about the reliability of their findings. Essentially, the statistical tests (unfortunately called significance tests) tell you what the odds are that the difference between your experimental and control groups is just a chance happening.

There are many pitfalls in applying statistics in assessing the significance of data. Perhaps, the most dangerous is that your devotion to statistics, may deflect interest from the practical and theoretical importance of your findings, which are quite another matter. Statistical significance does not guarantee either practical or theoretical importance. You incur a second danger in selecting a statistical test to be applied; often, statistics are applied to data for which they are not suited. Undoubtedly, the best course to follow, where possible, is to increasingly refine your technique and experimental control until your results are unequivocal.

*Telling the World.* There are many advantages in informing others of your experimental findings. In addition to receiving prestige as a researcher, you may receive helpful criticisms of your experimental design, references to related studies by other experimenters, indications of the range of applicability of your findings, and suggestions for follow-up research. In preparing your findings for publication, you may want to include the following steps: (1) tell others about your work; (2) then, write it up informally and distribute a dittoed copy to your fellow teachers and to someone who is trained in research methods, such as a psychologist or a linguist; (3) look over the journals in your field, and consider which one contains articles like your own; which one is read by the audience you wish to attract. Note the format in which the articles are presented and bear this in mind in your "write-up". Most journals have a manual of style to which you can refer. (4) Submit your article to the journal. Remember that the editors can also aid you in preparing the final manuscript by criticizing both form and content. Since it is true that neither piety nor wit will serve to retract an article once published, we strive for perfection before publishing. Nevertheless, suggestive findings from small-scale experiments deserve communication as well as the more definitive findings from large-scale research.

## SOME REFERENCES:

- Carroll, J. B. "Research on teaching foreign languages," *Handbook of Research on Teaching*. New York: Rand McNally and Company, 1962.
- Lumsdaine, A. A. and Glaser, R. *Teaching Machines and Programmed Learning: a Sourcebook*. Washington, Department of Audiovisual Instruction, National Education Association, 1960. (See, especially pp. 94-229.)
- Morton, F. R. *The Language Laboratory as a Teaching Machine*. Ann Arbor, The University of Michigan Language Laboratory, 1960.