

Research on Teaching

ARTHUR M. SULLIVAN**

Many faculty members of Canadian Universities are, or are becoming, deeply concerned about teaching and involved in finding the most effective methods whereby teaching can be made as efficient as possible. To this end some faculty members carry out research designed to investigate effective methods of instruction and many more read with great interest reports and reviews of research on the process of instruction which appear in the journals.

By and large, it is a frustrating business. In carrying out research activities faculty members are frustrated by:

Unsympathetic administrators who refuse permission or funds for research, or both, or who interfere with the process of research by imposing annoying restrictions.

Prejudiced colleagues who argue from intuition or conviction rather than fact and who refuse to accept empirical evidence when it is provided.

Sceptical colleagues who react to the most important discoveries with the comment "I knew it all along."

Academic regulations which change drastically and dramatically before the investigation has been completed.

Numerous variables which cannot possibly be controlled in a practical learning situation.¹

Reviewing the research reported by others is also frustrating. Two trends may be noted:

*An earlier version of this paper was delivered as an invited address to Division 16, American Psychological Association, Montreal, August 26, 1973.

The research reported was carried out with the assistance of Grants S71-1869 and S69-1624 from Canada Council.

**Dean of Junior Studies, Memorial University of Newfoundland.

¹As an example of the difficulties which one might encounter I can offer my own experience during the 1972-73 academic year. After several years of research on teaching and learning at the first-year University level, I received a major research grant to continue and elaborate my research activities. Research associates and assistants were hired and the research was planned in great detail. Three different methods of instruction were introduced in three different courses and numerous student and instructor characteristics were measured precisely. For nine weeks everything proceeded smoothly and then approximately three weeks before the final examination we experienced the only student strike in the history of Memorial University. The strike continued for two weeks and disrupted the learning experiences and the typical patterns of achievement so badly that the results for the entire semester were obviously involved and the entire study had to be repeated during the 1973-74 academic year.

On the one hand, very elegant research studies are reported from artificial laboratory situations. All variables are carefully controlled and the results are significant and often replicated. Even reviews of research in major areas are consistent and encouraging. However, the conclusions reached, although undoubtedly of great theoretical importance, are not of great or immediate significance for learning in practical situations.

For example, Merrill and Boutwell² in a review of instructional development state:

The research cited has exciting implications for instructional design and suggests the following proposition for concept instruction. Using, during practice, either expository or inquisitory instance presentations, if examples are divergent from examples and matched to simultaneously or sequentially presented non-examples then correct classification of subsequent unencountered instances is more probable.

On the other hand, the research reported from applied or practical learning situations is very rarely elegant and results, although occasionally highly significant in one study, are rarely replicated. Summaries of research from practical learning situations are inevitably equivocal and pessimistic.³

Indeed, Trent and Cohen⁴ report that:

. . . relatively few of the educational innovations developed during the 1960's with great hope for their widespread usefulness are in operation today. . . .

There are, as has been noted above and elaborated elsewhere⁵, many difficulties in carrying out research in practical educational settings. Included prominently in these difficulties are the problems of insuring adequate controls — in both the practical and the experimental sense — and the problem of finding a precise yet common and readily understood measure of achievement and of other changes (for example, in attitudes) which have taken place.

The one additional important and characteristic difficulty in educational research which is of sufficient importance to be stressed is the tendency for researchers and reviewers to overgeneralize their results.

This tendency to overgeneralize is exemplified on the part of educational researchers

²Merrill, M. D. and Boutwell, R. C., "Instructional development: methodology and research," in F. N. Kerlinger, ed., *Review of Research in Education*. Itasca, Illinois, F. E. Peacock, 1973.

³See, for example, Duchastel, P. C. and Merrill, P. F., "The effects of behavioral objectives on learning: a review of empirical studies," *Review of Education Research*, 1973; Dessart, D. J. and Frandsen H., "Research on teaching secondary-school mathematics," in R. M. W. Travers ed., *Second handbook of research on teaching*. Chicago, Rand McNally, 1973; Dublin and Taveggia, T. C., *The teaching-learning paradox*. Eugene, Oregon, University of Oregon Press, 1968; Hartly, J., "Evaluating instructional methods," in I. K. Davies and J. Hartley, ed., *Contributions to an educational technology*. London, Butterworths, 1972; Mitchell, P. David, "Effectiveness of university teaching aids and a computer-based instructional planning laboratory," Paper presented at the Canadian Psychological Association annual meeting at Victoria, June 8-10, 1973.

⁴Trent, J. W. and Cohen, A. W., "Research on teaching in higher education," in R. M. W. Travers ed., *Second handbook of research on teaching*. Chicago, Rand McNally, 1973.

⁵Sullivan, A. M., "Psychology and teaching," *Canadian Journal of Behavioral Science*, Vol. 6 (1974), 1-29. (Hereafter referred to as Sullivan, "Psychology and Teaching.")

by the search for *the* method which will lead to improvement for *all* students in *all* courses. This general tendency has been referred to elsewhere⁶ as the Columbus complex, but in educational researchers this complex reaches its most extreme manifestation – the American obsession. Educational researchers will not be content with the discovery of a few offshore islands. They want to find, and immediately be credited with, the discovery of a whole readily identifiable continent. They do not appear to realize that the discovery of continents starts with the discovery of offshore islands and involves many years of patient exploration and charting of capes and bays.

The search for the one, all-sufficient method of instruction together with an almost invariably sparse and imprecise description of the characteristics of the subjects (learners) on whom the studies have been carried out has led researchers and reviewers alike to ignore or underestimate the effect of important variables and to miss crucial interactions entirely.

A hypothetical example of what may happen may be borrowed from agricultural research.

BRAND	REDDISH ROUNDISH FRUIT								
	A.			T.					
X	C	E	D	C	E	D			
	10	15	+5	10	5	-5			
Y	10	5	-5	10	15	+5	20	20	0

Figure 1. Results of three studies involving the effects which two types of fertilizer have on “reddish-roundish” fruit.

As Figure 1 shows Researcher A is carrying out a study which is intended to find out if Fertilizer X will produce an increase in the size of fruit. He selects a group of one type of fruit which he describes as “roundish-reddish,” divides the group into two and administers Fertilizer X to one group while the other is subjected to some appropriate control procedure. He finds a significant difference between the groups and concludes that Fertilizer X does indeed produce a significant increase in the size of reddish-roundish fruit.

Researcher B reads his report, tries the same procedure with what he describes as reddish and roundish fruit. He is dismayed to find that the treatment in fact has the opposite effect and the experimental fruit is definitely smaller than the non-treated controls. Researcher C tries out the procedure with another type of Fertilizer “Brand Y” on two types of loosely classified fruit, both of which may be described as reddish and roundish, and finds that the treatment has no significant overall effect.

⁶Sorokin, Pitrim, A., *Fads and foibles in modern sociology*. Chicago, Henry Regenery Company, 1956.

Researcher D reviewing all of the above studies concludes, quite reasonably I suppose, “there is no evidence that any type of fertilizer has been successful in increasing the size of fruit.”

A more precise description of the characteristics of the fruit would reveal, however, that two distinct types have been investigated, one roundish-reddish type of fruit – an apple (Column A) – while the other roundish-reddish type is a tomato (Column T) – a very different type of fruit indeed. Fertilizer X has consistently produced larger apples and smaller tomatoes while Fertilizer Y has consistently produced smaller apples and larger tomatoes. And yet, all researchers and the reviewer have missed this major difference and this extremely important interaction.

The situation is infinitely more complex in educational research where, because there may be a much greater variety of subject characteristics and an infinitely greater variety of fertilizers, the possibility of interaction becomes more pronounced and important. Recent reviews of trait-treatment interactions,⁷ suggest with cautious optimism that significant and important interactions are in fact not a rare occurrence in educational settings.

There are, in addition to the obvious variable “methods of instruction,” *three* extremely important variables which must be taken into account in planning research on methods of instruction, and in interpreting the results of such research and in searching for possible interactions.

These are:

- 1 The characteristics of the learner.
- 2 The characteristics of the subject matter.
- 3 Time.

We shall examine each one briefly.

1. Characteristics of the Learner: The most important individual difference variables – or characteristics of the learner – are:

- (a) Age.
- (b) Sex.
- (c) Intelligence.
- (d) Level of anxiety.
- (e) Degree of extroversion or introversion.
- (f) Cognitive – or learning – style.

All of these factors have obvious, well documented, and important effects on learning. All, as well, have effects which are not so obvious or well documented but which have nonetheless been noted and investigated by researchers.⁸ No attempt will be made in this

⁷See, for example, Cronbach, L. J. and Snow, R. E., *Individual differences in learning ability as a function of instructional variables*. Final report to LISOE, March 1969, Stanford University, School of Education, Contract No. OEC 4-6-061269-1217; Berliner, D. C. and Cahen, L. S., “Trait-treatment interaction and learning,” in F. N. Kerlinger ed., *Review of Research in Education*. Itasca, Illinois, F. E. Peacock, 1973.

⁸See, for example, Gagné, R. M., ed., *Learning and individual differences*. A symposium of the Learning Research and Development Center, University of Pittsburgh. Columbus, C. E. Merrill, 1967.

presentation to deal with even the most obvious direct effects of these variables because what is more important for our purposes at this time is the finding that each of these variables may interact with methods of instruction to produce very interesting and important consequences.

For example, Domino⁹ selected two types of students: one as being high in a personality trait identified as independence and the other high in a trait identified as conforming. He subjected one group of each type of student to either a high structure (teacher-dominated lecture) or a low structure (student-led discussion) learning experience. One of his dependent variables was factual knowledge in introductory psychology. As Figure 2 will show, the high conforming students demonstrated a higher level of achievement in the high-structured situation and a lower level of achievement in the low-structured situation when compared with the high independent students.

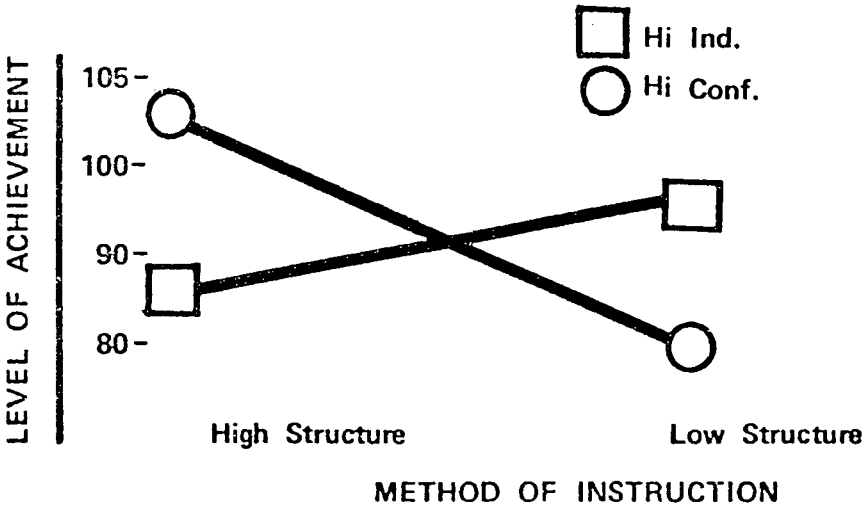


Figure 2. Effect of two types of instruction on the achievement of two different types of students (after Domino 1971).

Studies carried out at the Institute for Research in Human Abilities at Memorial University of Newfoundland also reported results which are relevant to this area of interest.¹⁰

The dependent variable in these studies was scores on a difficult Letter Score Test. The research design included four practice conditions.

⁹Domino, G., "Interactive effects of achievement orientation and teaching style on academic achievement," *Journal of Educational Psychology*, Vol. 62 (1971), 427-431.

¹⁰Sullivan, A. M. and Skanes, G. R., "Differential transfer of training in bright and dull subjects of the same mental age," *British Journal of Educational Psychology*, Vol. 41, No. 3 (1971), 287-293; Skanes, G. R., Sullivan, A. M., et. al., "Intelligence and transfer," Institute for Research in Human Abilities, Memorial University of Newfoundland, 1973.

Condition 1. Practice on Letter Scores Items.

Condition 2. A pre-test followed by practice on LST.

Condition 3. Practice on Number Scores Items.

Condition 4. A pre-test followed by practice on LST.

The results (see Figure 3) demonstrate that for dull subjects the condition which produces the highest level of performance (that is, highest scores on the Letter Scores Test), is Condition 1, that is the condition which involves no pre-test but only practice on material which is highly similar to that found on the test. This same condition produces the lowest level of performance for the bright students. The condition which produces the highest level of performance for bright students is Condition 4 which includes a pre-test followed by practice on material which is related, but not highly similar, to the material found on the test. This condition produces the lowest level of performance for dull students. (Interestingly, this finding holds even when mental age is held constant – that is when bright and dull subjects are of the same mental age – in which case an intersecting interaction similar to that in the study reported above has been found.)

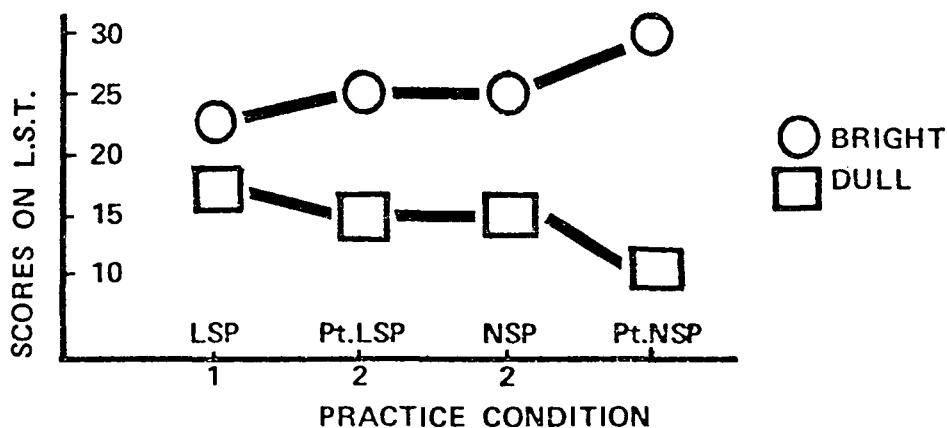


Figure 3. Effects of four different practice conditions on the Letter Score Test, scores of bright and dull students.

No attempt will be made in this paper to present a comprehensive list of such interactions.¹¹ The important point which emerges and which must be stressed is this:

Any given method of instruction may produce an improvement in the performance of one group of learners, but that same method may not necessarily facilitate the performance of other groups of students who do not have the same characteristics – and may in fact actually produce a decrement in the performance of students whose characteristics are markedly different.

It is obvious that such interactions are of considerable importance in the investigation

¹¹See, for example, Witkin, H. A., *The Role of Cognitive Style in Academic Performance and in Teacher Student Relations*. R. B. 73-11, Educational Testing Service, Princeton, New Jersey.

of effective instruction in practical learning situations. It is equally obvious that such interactions will be missed unless the major characteristics of the learners are documented and reported in all research studies.

2. Characteristics of the Subject Matter: This factor is very rarely acknowledged by researchers or reviewers and yet it is undoubtedly of considerable importance.

For example, five years ago I introduced a method of instruction for teaching introductory psychology.¹² The method of instruction (see Figure 4) involved a very careful specification of objectives followed by an opportunity for individual learning followed by a test to determine whether the objectives had been attained. This initial test was followed by group instruction and eventually, for those who had not yet attained criterion level of performance, an individual tutorial. It has worked extremely well in psychology and we are satisfied with the results after five years have elapsed.

Two years ago, however, we decided to use essentially the same method, with minor modifications, for teaching in mathematics. Surprisingly, we found no significant overall difference between the performance of experimental and control subjects. When, however, we compared the performance of subjects according to their previous level of achievement in mathematics we found that the approach was of considerable benefit for those whose previous level of achievement was high but not of benefit for those whose previous level was low. When we examined the previous academic performance of students who had been taught by the same approach in psychology we found that exactly the reverse was true. That is, the approach was of more benefit for those students whose level of prior academic achievement was low than it was for those students whose prior level of achievement was high. We found other important differences between psychology and mathematics, for example some of the characteristics of successful instructors in mathematics and psychology are quite different.¹³

Successful instructors in psychology are characterised by an ability to get the student to work very hard. But the successful instructor in mathematics is characterised by an ability to prevent students from becoming discouraged. This is especially true for a conventional non-structured approach in which the objectives are not precisely defined and evaluations are relatively infrequent. Other differences have been noted. For example, a student will avoid tutorials in mathematics, but go willingly to tutorials in psychology; and a low level of achievement in mathematics is associated with a lower than predicted level of achievement in other subjects. This is not true of psychology.

There are doubtless many factors which might account for this difference, but the most important concerns the prerequisite skills needed for both subjects. For first-year psychology, the student must know how to read, write, and compute. Psychology texts are plainly written and most students can study the subject on their own. Since students have not had previously discouraging experiences with psychology, knowledge of results does not distress them, but rather serves as a useful stimulus to further work. In mathematics, however, the case is different. At the university level a great number of complex

¹²Sullivan, A. M., "A structured individualised approach to the teaching of introductory psychology," *Programmed Learning and Educational Technology* 6 (1969), 231-242.

¹³Sullivan, "Psychology and Teaching," p. 21.

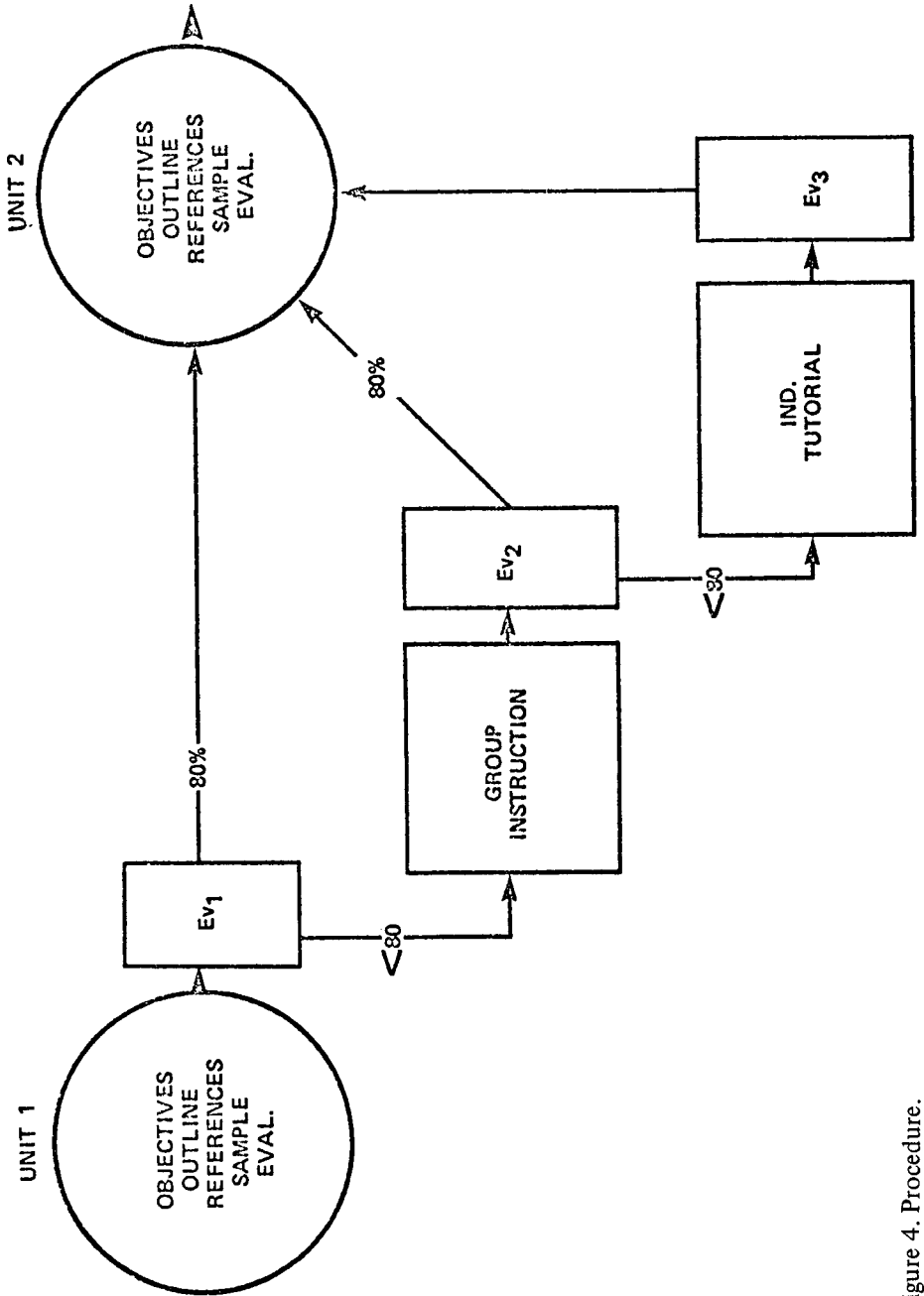


Figure 4. Procedure.

prerequisite skills is needed. Without them the student cannot learn even from well-written self-instructional texts. In addition, knowledge of results, if the student is not doing well, reminds him of previous frustrations in mathematics and is likely to diminish his already weak motivation. Therefore, it is extremely naive to expect the same system of instruction to be equally successful in psychology and in mathematics.

Cognitive (for example, number and type of prerequisite skills) and motivational factors (for example, nature and amount of previous discouragement with the subject) may enable us to place science-oriented subjects on a continuum from psychology on the one hand through biology, chemistry, physics to mathematics on the other. One would, therefore, expect that a highly structured method of instruction which emphasized specific objectives, a pre-test and self-learning, such as the PSI method, would be most successful in psychology, somewhat less successful in biology, less again in chemistry and least successful in mathematics. Although no comprehensive data have been collected concerning this hypothesis, the preliminary review of the literature which I have been able to carry out suggests that this is in fact the case.

Again, however, the most important point is that there are major and important differences in the degree of success which a particular method of instruction will have with various subjects. There may in fact be interactions in that *one method of instruction may produce an increased level of performance in one subject but a decreased level in another*. Such variations and interactions can only be found if the course on which the study is carried out is reported and if the characteristics of the course are noted and attended to by reviewers.

3. Time: It is likely that the passage of time will reveal important changes and differences, but it is very difficult to obtain relevant information since most studies last for only one semester and are not repeated — or if repeated the results are not reported.

Our own experiences in psychology at Memorial University suggest that important differences do in fact occur with time. Not only with regard to variations in overall achievement but also in less obvious ways.

The results of our original experimental and control group were quite dramatic and highly significant.

In subsequent years, however, the results although encouragingly consistent and significantly better than those which we obtained in years when the lecture approach had been used, nonetheless were not nearly as impressive as those obtained during that first year. Our results then suggest that as far as overall results are concerned there is an initial and highly dramatic effect in the first year which is reduced considerably in the second year. However, there is little overall variation after the second year and achievement in the second and subsequent years remains at a relatively high level in comparison with a conventional lecture approach.¹⁴

In addition to overall achievement effects, there are other effects which can only be noted with the passing of time. During the first years in which the structured approach was used in psychology we found very little variability associated with instructor, however as time went on, the variability became much greater and was found to be associated with

¹⁴ Student satisfaction has also been consistently high and student achievement in subsequent psychology courses has been both high and highly correlated with achievement in the introductory course.

experience of the instructor, and appears to be related primarily to the nature of the instructor's involvement in the programme over a period of time. It appears to me that our present inexperienced instructors, not having been part of the development of the system, now feel imposed on by the system and their consequent dissatisfaction with the system itself and with the materials which have been prepared is communicated to the students and is partially responsible for their lower level of achievement.

Thus, a method of instruction, although successful in one setting and for a limited period of time, may prove very disappointing when an attempt is made to transplant it to another setting. Therefore, such methods of instruction must be evaluated and modified continually. A detailed account of the research referred to above, together with statistics and tentative conclusions has been published elsewhere.¹⁵

In summary then, it appears to me that it is only in the context of these three variables – that is, characteristics of the learner, characteristics of the subject matter, and time – that methods of instruction can be investigated most fruitfully and valid results obtained.

Some implications are obvious. For example, a highly structured method which has produced significant improvement in the performance of introverted high-conforming high-anxiety students in biology is not likely to produce a similar degree of improvement if used with high-independent low-anxiety extroverted students in philosophy. For these students and that subject, a student-centered discussion approach might be expected to produce the highest level of achievement and satisfaction.

Other implications are not so obvious. Even within a highly structured system many variations are possible; for example, objectives can be more or less specific, a pre-test can be given or omitted and a criterion level of performance may or may not be required. A system which included specific objectives, a challenging pre-test, and a criterion level selected by the student might be expected to produce a high level of achievement for bright students in psychology, but might very well *reduce* the level of achievement in mathematics especially for those students whose prior level of achievement was low. For these students and that subject the system which would be expected to produce the highest level of achievement would include general objectives, no pre-test, and a moderately high criterion level selected by the instructor.

More, and more specific, predictions must await further research. From this point of view I would suggest that in the general area of research on methods of instruction there is too much of the wrong kind of research being done and not enough of the right kind.

By the wrong kind of research I mean studies which involve only one course, or one (usually atypical) group of instructors, which give no information concerning the characteristics of the students, which include only vague achievement data, which continue for only one semester and are never replicated. And yet such studies are not only common but also treated as if the findings could be generalized to other students and other courses and by reviewers as if they could form the basis for the formulation of important general principles concerning the psychology of instruction.

By right kind of research I mean research of two kinds.

1. Major research efforts which are directed towards the discovery of general principles and their areas of application. These studies must involve more than one course, more

¹⁵Sullivan, "Psychology and Teaching."

than one type of instructor, must report full, comprehensive data concerning achievement of students and other changes, (for example, in attitudes) and must be carried out over a period of years. Such major research efforts will require careful planning and considerable research support but, in my opinion, are absolutely necessary if general principles are to be found.

2. Purely practical research in which in one subject and one semester an individual instructor tries systematically to improve his teaching and to evaluate the results of his efforts. Such applied research would be intended to solve one practical problem and would not be intended to have theoretical significance. However, such studies could – if reported appropriately – be fitted into a general framework. By this I mean reported so that all details, for example, the type of subject matter, the characteristics of students, the achievement measures used, et cetera, were presented in detail or readily available.

Such purely applied research projects could then be collected by reviewers and fitted into their appropriate place in the general theoretical framework and eventually a relatively distinct and complete picture of the variables and the interactions which are of importance in the psychology of instruction will emerge.

Those who work in educational research must realize that each cannot, to use the previous analogy, discover and chart a complete continent all by himself. Educational researchers must realize that there is satisfaction and glory enough in discovering small islands and in charting small bays. It is only in this way that a complete and valid map of the entire continent, which represents teaching and learning in the university, will ever be produced.