IDEA TECHNICAL REPORT NO. 1

Development of the IDEA System

Donald P. Hoyt

William E. Cashin

Center for Faculty Evaluation and Development
Kansas State University
College Court Building
1615 Anderson Avenue
Manhattan, KS 66502-4073

INTRODUCTION

The primary purpose of the technical reports is to provide statistical and quantitative information about the IDEA system beyond that which is available in the Center's publications for the general user. These reports should be of particular interest to professionals in educational and psychological measurement. These technical reports presume that the reader is already familiar with the Center's materials for the general user; especially the Faculty Information Form, the Survey Form, the IDEA Report itself, and the Interpretive Guide.

Since it is our intention to continue to develop the IDEA system, we anticipate periodically publishing additional technical reports. This report, Technical Report No. 1, describes the development of the IDEA system up to the point when it became widely available, in the fall of 1975, to college instructors beyond Kansas State University.

Most of the research and development on the system up until that time was conducted by Donald P. Hoyt, director of the Office of Educational Research at Kansas State University. Dr. Richard E. Owens, director of the Office of Educational Improvement and Innovations, also made significant contributions to the development of the system, especially to the 1972 and 1973 revisions.

Because of his many commitments, it was not possible for Don Hoyt to write this Technical Report. Most of the Report, however, is based upon materials he had drafted or had written for other purposes. To the extent that the report's content is of value, the credit belongs to him. Shortcomings in the report, especially regarding its manner of presentation, are this writer's responsibility.

Technical Report No. 2 will describe in detail the computational procedures and comparative data base used in producing IDEA Reports during the 1976-77 academic year.

Because the data base will have doubled during 1976-77, we plan to reexamine the entire system during the summer of 1977. In the fall we plan to publish Technical Report No. 3 which will update the data presented in the second report. Additional reports will describe analyses of other aspects of the system.

March, 1977 William E. Cashin

TABLE OF CONTENTS

The IDEA system was originated at Kansas State University during the 1968-69 school year. Its historical, and political origins have been described in detail elsewhere (Hoyt, 1973a, 1973b, 1973c). While readers may discover a modest amount of humor and a remarkable experience of deja vu by reviewing the early trials and tribulations of the program, the most essential information about IDEA's invention concerns its purpose and rationale.

I. BACKGROUND

A. PURPOSE

During the fall of 1968, administrators, students, and faculty members at Kansas State University contributed to a lively debate about the purpose of a "student rating of instruction" program. At one point or another, the following observations were made:

1. Student feedback can provide information to the faculty member which will allow him/her to improve the quality of his/her instruction.

2. Student ratings can be published to give other students useful guidance in selecting instructors and courses which will best meet student needs.

3. Administrative judgments (merit salary recommendations; promotion; tenure) will be improved if student evaluations are used to inform them.

4. Faculty members who neglect their teaching or fail to keep up should be publicly exposed (embarrassed) even if the tenure policy protects them from losing their jobs.

All elements of the educational community agreed only on the first of these proposed purposes. Each of the other proposed purposes had both strong proponents and strong opponents! Since it seemed unlikely that the program could succeed without the support of all three groups (students, faculty, and administration), it was agreed that the sole purpose of the program would be to improve instruction.

B. RATIONALE

The initial problem was to define effective teaching. How can one recognize effective teaching? How does it differ from mediocre or poor teaching?

A false start was made by trying to compile a list of items which describe the attitudes and behaviors of the "ideal teacher." Thus, a list of 174 items was developed primarily from the work of Solomon (1966), Whitlock (1966), and Isaacson et al. (1964). But when local critics were asked to verify the value of each item as a descriptor of an effective teacher, some heavy resistance was met.

The most telling criticism of the approach was its intolerance of individual differences. Thus, our critics said, an item like "Induced students to participate" may be effective teaching for some faculty members in some courses; but it would be a mistake to suggest that no

teacher is effective unless he or she induces or seeks class participation. In other words, there are two serious flaws in the "ideal model" approach:

1. There may be many effective styles of communication; to suppose that there is one "best model" is to deny the importance of individuality among faculty members.

2. Instruction has many different purposes; methods which serve one purpose well may serve another one poorly. Therefore, a given teaching method may be "effective" in one situation but "ineffective" in another.

Because these criticisms appeared to be valid, a new approach was sought. There was general agreement that effective teaching could be recognized by its effect on students; if instruction was effective, students learned. But the type of change would differ depending on the subject matter, the level at which it was taught, and the intentions of the instructor.

Thus, one might expect lower division courses to concentrate on relatively basic cognitive development (e.g., gaining factual knowledge; learning principles, generalizations, theories; etc.). Upper division and graduate courses would be expected to stress higher level analytical skills (e.g., problem-solving or critical thinking) as well as professional applications. Similarly, some courses have "general or liberal education" purposes, serving primarily to increase students' understanding and appreciation of their cultural heritage; others may focus on the "personal development" of the student (e.g., self-understanding; career or life planning); still others may provide a foundation for advanced study; and some concentrate on practical skills (e.g., typing; blood analysis). Could a single student rating form be developed to provide useful means of judging "effectiveness" for courses with such diverse purposes?

Clearly, if such a device were developed, it would have to take into account these crucial individual differences. If effectiveness is to be assessed by student's learning, the types of learning must reflect the instructor's purposes.

The faculty members and students who served as consultants agreed that, conceptually, the effectiveness of teaching should be assessed by determining the amount of progress students made on objectives stressed by the instructor. This definition of teaching effectiveness is central to the IDEA system.

Two points require elaboration: (1) the focus on student progress and (2) the focus on instructor-chosen objectives.

With respect to student progress, we are concentrating on a concept not usually considered in higher education. In the typical class, we give letter grades (A-F) or, perhaps, some narrative rating or observation. In any event, these "traditional" reports tend not to be reports of student progress, but reports of end-of-course status. Thus they represent the admixture of many variables, the most important of which are probably the student's general academic ability or intelligence, previous background in the field, and academic motivations and habits. Such student characteristics will account for such a large part of the variance in final grades that one would be hard put to discover a unique effect due purely to the quality of instruction.

If we want to measure teaching effectiveness by examining the status of students on instructor-chosen objectives, we had best remove from that "status measure" as much "irrelevant" variance as possible ("irrelevant" in the sense that the variation was produced by factors independent of the instructor's influence). Hence "progress" (the difference between initial status and end-of-course status) appears to be a better way to infer teaching effectiveness than simple end-of-course measures (final examinations, GRE scores, CLEP scores, etc.).

The other matter needing further attention is the focus on instructor-chosen objectives. There is considerable debate about the matter of choosing educational objectives. There is no gainsaying the importance of student goals. One tends to get out of an activity what one puts into it; and what one puts in very often depends on his/her goals and purposes. But, in our view, the ultimate responsibility for selecting objectives must rest with the instructor. There are several reasons for this.

First, the instructor should not be held accountable for failure to achieve objectives which he/she neither chose nor endorsed. Second, the instructor often assumes responsibilities to many others -- to society at large, to employers, to colleagues teaching advanced courses -- which assure them that "successful" students in the instructor's course have achieved certain understandings or competencies. The instructor cannot neglect such responsibilities on the grounds that the objectives they imply were of little interest to students. Third, there is no reason why, in selecting objectives, the instructor cannot seek and use student input. In many cases, it is entirely appropriate to let student needs or desires dictate how certain options will be selected. But in the final analysis, the instructor must be responsible for selecting the objectives to be pursued, because only the instructor has an understanding of the diverse expectations of all who are legitimately concerned.

II. DEVELOPMENT OF IDEA SURVEY FORM

A. SELECTING OBJECTIVES

If effectiveness is to be judged by the amount of student progress on instructor objectives, a method must be found for identifying the objectives of a given instructor. The most direct route to this goal involves asking the instructor. However, there are some practical drawbacks to this plan. There are the usual problems in semantics; one instructor's "problem-solving" may be the same as another's "application to practical situations." In addition, given the opportunity to make an open-ended identification of objectives, some instructors will provide a detailed list of behavioral objectives (e.g., "Given a map of the 48 contiguous states, the student will be able to name and geographically locate the four which produce the most wheat"); others will provide only very broad objectives (e.g., "Teach the students economic geography"). It would be desirable if all instructors used a common vocabulary and the same level of abstraction in describing their objectives. These desiderata are most likely to be attained if a common set of objectives is prepared from which instructors may select those descriptive of a given course.

As a first step, a review was made of two taxonomies of educational objectives (Bloom, 1956; Krathwohl, et al., 1964). The classifications in these taxonomies were too elaborate to be used directly; neither students nor professors could be expected to respond meaningfully to such a detailed listing.

The possibility of synthesizing the specific objectives into a smaller set of general objectives was suggested by the work of Deshpande and Webb (1968). These investigators showed by factor analysis that a large number of specific objectives endorsed by the faculty of Georgia Institute of Technology could be reduced to a much smaller set of general objectives. On inspection, the Deshpande-Webb "factors" bore a close resemblance to several of the major classifications given in the taxonomies.

This correspondence encouraged us to prepare a list of general objectives which might be used to describe any undergraduate course. A tentative set of eleven major objectives was developed. This list was submitted to a group of five professors who had previously won outstanding teaching awards at Kansas State University. On the basis of their critique the list was reduced to six objectives. A further critique was offered by members of student-faculty committees on effective instruction in the colleges of Agriculture, Engineering, and Home Economics at Kansas State University. As a result of suggestions from these groups, two additional objectives were selected.

The set of eight objectives which survived this process are listed below:

1. Gaining factual knowledge (terminology, classifications, methods, trends).

2. Learning fundamental principles, generalizations, or theories.

3. Learning to apply principles to solve practical problems.

4. Understanding myself--my interests, talents, values, etc.

5. Learning attitudes and behavior characteristic of professionals in the field most

closely related to this course.

6.Developing skill in effective communication.

7.Discovering the implications of the course material for my personal and professional conduct.

8.Gaining a broader understanding and appreciation of intellectual-cultural matters (music, science, literature, etc.).

The above eight objectives were those used in the first version of the system (see Appendix A, Form 1, items 59-66). Based upon instructors' experience with the first version, and upon suggestions for additional objectives which were gathered during some administrations of Form 1, the original eight objectives were modified and expanded to the ten objectives listed below. These were the ten objectives used in the 1972 revision (Appendix A, Form 2, items 50-59), and the 1972 short form (Form 3, items 23-32), and are the ten objectives which are now used (Form 4, items 21-30).

1.Gaining factual knowledge (terminology, classifications, methods, trends).

2.Learning fundamental principles, generalizations, or theories.

3.Learning to apply course material to improve rational thinking, problem-solving, and decision making. (Modification of original objective 3)

4.Developing specific skills, competencies and points of view needed by professionals in the field most closely related to this course. (Modification of 5)

5.Learning how professionals in this field go about the process of gaining new knowledge. (New objective)

6.Developing creative capacities. (New objective)

7.Developing a sense of personal responsibility (self-reliance, self-discipline). (New objective)

8.Gaining a broader understanding and appreciation of intellectual-cultural activity (music, science, literature, etc.). (Modification of 8)

9.Developing skill in expressing myself orally or in writing. (Modification of 6)

10.Discovering the implications of the course material for understanding myself (interests, talents, values, etc.). (Combination of 4 and 7)

B. SELECTING TEACHING METHODS

The investigation sought not only to develop a dependable way of evaluating instruction, but also to discover correlates of effectiveness which could provide insights into how improvements might be made. Do teachers whose students make considerable progress on a given objective behave differently in the classroom from those whose students make little progress?

Of course, the research hypothesis was that such differences did exist. Further, it was hypothesized that the specific behaviors correlated with effectiveness would vary depending on the teaching objective. To test these hypotheses and to fulfill the purpose of discovering clues as to how instruction could be improved, it was necessary to construct a device which would provide a suitable description of instructor behaviors.

A questionnaire approach was selected. An effort was made to write items which (a) represented a meaningful dimension along which instructors might vary, (b) gave evidence of being related to instructional effectiveness, and (c) described teacher behavior in sufficiently specific terms that, if it seemed desirable to alter the behavior, this could be communicated clearly.

Instruments used by other researchers were reviewed. Three appeared especially promising in terms of their comprehensiveness, specificity, and the thoroughness with which they have been examined statistically. These included the 72-item questionnaire developed by Solomon (1966), the 46-item questionnaire developed by Isaacson et al. (1964), and the 56-item device developed by Whitlock (1966). In each of these studies, factor analyses of the items have been performed. The factors identified, and the items loading heaviest on each, were compared for ostensible similarity and overlap. By this process, 87 of the 174 items were eliminated as "redundant."

The remaining 87 were submitted to a group of five faculty members who had won "outstanding teaching awards" at Kansas State University. They were asked to edit the items, identify remaining redundancies, suggest any types of potentially important behaviors which had been overlooked, and to delete items which they believed were irrelevant to teaching effectiveness. This process resulted in further editing of 32 items, elimination of 21, and suggestions for additional items having to do with characteristics of the course (assignments, examinations, etc.) rather than those of the instructor.

The remaining items were reduced to 36 on the basis of their relationship to a measure of "overall teaching effectiveness." Items from the Isaacson et al. (1964) inventory were retained if they consistently loaded on their "skill" factor. Solomon items were retained if they loaded on either of the two factors which differentiated his "effective" from "less effective" instructors. And Whitlock items were retained if they (a) differentiated faculty members nominated for teaching awards from those not nominated and (b) differentiated among the top and bottom 27 percent of the nominated instructors.

These 36 items, and the 30 items not selected, were presented to student-faculty committees on the improvement of instruction in three colleges at Kansas State University. These committees were invited to criticize these items and suggest potentially useful

additions. On the basis of their recommendations, six other items were selected. Four of these were related to "overall teaching effectiveness," though not to the degree required of the original 36 items.

On the recommendations of consulting students and faculty, a set of 16 additional items were constructed to describe the course. Four items were devoted to each of four aspects of the course--examinations, (Form I, items 43, 47, 51, and 55), out-of-class assignments, (items 44, 48, 52, and 56), textbooks (items 45, 49, 53, and 57), and the course content (items 46, 50, 54, and 58). These 16 items were all original and therefore could not meet the same selection standards employed with the other 42 items.

As can be seen by examining the various forms, these teaching methods items have undergone considerable revision over time. The 1972 version (Form 2) reduced the number of items to 36; the 1973 "Short Form" (Form 3) contained 18 items; and the present version (Form 4) has 20 teaching methods items. The primary rationale used for including items in these various revisions was the relationship between each teaching method and students' reports of progress on the objectives. This will be discussed in more detail in Part III, Section B., Determining Relationships Between Methods and Objectives.

C. SELECTING OTHER VARIABLES

The major assumption which guided the development of the instrument was that specific teacher behaviors did influence certain types of student progress under certain circumstances. Therefore, to provide assistance in improving instruction, it would be necessary to establish the relevance of a given instructional method or approach for a particular goal. At the same time it was recognized that progress might also be contingent upon course management decisions. This was the reasoning behind including the set of 16 items on examinations, out-of-class assignments, textbooks, and course content (Form 1, items 43-58) described above. But in addition, "success" (student progress) might also be contingent on certain characteristics of the students enrolled in the course. Therefore, questions were included about their previous grades, interest in the course, and curriculum requirements (Form 1, items 67-72).

Probably no set of items studied during the development of the IDEA system has undergone more change and experimentation than these items. On Form 2 these items were grouped under two headings: The Course, consisting of 13 items (38-49); and Self-Rating, eight items (60-67). On Form 3, the Short Form, using the same two headings, the number was reduced to four items (19-22) and four items (33-36), respectively. On the IDEA Survey Form presently used (Form 4), still using the same two headings, the Course section contains four items (31-34) and Self-Rating contains five items (35-39).

Throughout all of the experiments with the various items, the guiding principle has been how much useful information did the item yield, especially with respect to being able to sort out effects due to the instructor versus those resulting from course or student characteristics. More will be said about this in the next part.

III. ASSESSING INSTRUCTIONAL EFFECTIVENESS

With the exception of developing the items for the instructor's course objectives, the process of selecting items for the IDEA Survey Form was fairly similar to that used by many student evaluation of teaching approaches. To develop the more unique aspects of the IDEA system, three problems had to be addressed: 1) measuring the extent to which the students made progress on the objectives the instructor had selected for the course; 2) determining what relationships, if any, existed between the teaching methods and progress on the instructor's objectives, and 3) controlling for effects other than the instructor's teaching which might influence the students' ratings (e.g., the students' level of motivation and the size of the class). Each of these three problems will be discussed in turn. Statistical data bearing upon each problem will be presented in Part IV.

A. MEASURING STUDENT PROGRESS

A key problem was estimating student progress on each objective. Solutions involving the direct measurement of student progress were ruled out as unrealistic. Such an approach would require the development of as many as ten psychometrically reliable and valid measures of student progress for each, individual course. Likewise, pre- and post-testing was considered to be an unsatisfactory solution on purely pragmatic grounds because of the amount of class time required.

The most straightforward solution would be to ask the student to judge his or her own progress. While such a solution has some clear advantages (simple, quick, provides for differences among courses, focuses on intra- rather than inter-individual comparisons), it is not without its dangers.

There are a number of reasons to suspect that not all students are capable or willing to make an accurate self-appraisal (e.g., Combs et al., 1963). It would be difficult to justify using self-ratings to assign grades to individuals. But it was proposed that class averages be used to evaluate instructional effectiveness. Given two classes which stress the same objective, all that is required is that there be more true progress in the one with the higher average rating. Most likely, some students would be too optimistic, some would be too pessimistic, some would be careless, and some would be mistaken. But if these errors were not systematic (i.e., if they occur in all classes in about the same proportion), and if there was at least a substantial minority who give responses which were reasonably accurate, then the difference in average ratings would still be meaningful.

There is considerable evidence bearing on the question. For example, a number of studies have shown that student estimates of their probable grade point average is about as predictive of first year performance as are college aptitude tests or high school rank (e.g., Keefer, 1965). Other studies have shown that a self-rating of vocational interests is more predictive of future occupational choice than are interest test scores (Holland & Lutz, 1968). Still other studies show that the amount of distortion which occurs in making self-reports is minimal even when there is considerable motivation to distort (Walsh, 1967); for example, both scholarship applicants and non-applicants report quite accurately the special honors and recognitions they have received.

Encouraging as these findings are, they are not directly related to the present question. A report by Solomon, Rosenberg, and Bezdek (1964) is of more direct relevance. A carefully constructed achievement examination was used to measure mastery of relevant factual information in 24 college classes in American Government. Pre- and post-tests were administered, so that "gain" could be studied, thus controlling for aptitude and experience differences. At the end of the courses, students provided a number of self-ratings, including one on how much factual information they had learned. This rating correlated .52 with gain scores, indicating a substantial degree of overlap between the two. When considered with the other evidence on self-ratings and the minimal assumptions which must be met before these could be used in a program of instructional evaluation, there was reason to believe that this simple approach might have much to offer, and so it was decided to adopt this approach.

B. DETERMINING RELATIONSHIPS BETWEEN METHODS AND OBJECTIVES

It was originally hypothesized that progress on different objectives would entail different teaching methods. The logic of this seemed compelling. Lecture-type activities like clearly stating objectives and summarizing in ways which aided retention seem better suited to gaining factual knowledge or learning principles than to improving critical thinking or problem-solving skills. Activities like helping students to answer their own questions or encouraging students to express themselves seem more appropriate for the achievement of the last two.

There were no direct studies linking specific teaching methods with specific course objectives in the research literature, but the question was empirically testable. Given a set of teaching methods items and a set of course objectives items, one need simply determine if significant correlations existed between the students' ratings of the frequency with which the instructor used various teaching methods and their ratings of progress on the instructor's objectives for the course. Although the discovery of such correlations would not "prove" a causal relationship between employing a given teaching method and students' making progress on a given objective, such evidence would provide strong support for such a relationship. At the very least, an instructor might reasonably be advised to consider the use of such relevant methods when it was found that his or her students were not making satisfactory progress on an objective.

C. CONTROLLING OTHER EFFECTS: CLASS SIZE AND STUDENT MOTIVATION

Any attempt to measure teaching effectiveness succeeds to the extent that it can isolate outcomes which are effects of the instructor's teaching from outcomes which result from characteristics of the student (e.g., previous knowledge; lack of effort), and of the course (e.g., class size; difficulty of subject matter). Asking the student to compare the progress he or she made in this course with progress made in other courses was one attempt to control such influences. The inclusion of several items asking about student effort, course difficulty, etc., was the other approach used.

Class size was one of the factors considered. It seemed logical that some teaching methods might be affected by the size of the class; for example, discussion is more feasible in a class of 10 students than in one containing 100 students. To answer this question one would need to examine both the frequency with which a teaching method was used in different size classes and its relationship to the students' ratings of progress. Both were done.

Most student rating forms contain an item asking if the course was "required" on the grounds that students will probably give higher ratings to "elective" courses. Such an item was included on the first version (Form 1, item 67). However, experience showed that some "required" courses were very popular with students (especially required courses in the major), and some "elective" courses were regarded negatively (especially science or mathematics electives taken to satisfy distribution requirements). Measures of the students' level of interest (item 68) or motivation to take the course proved more useful, resulting in the eventual selection of a single item "I had a strong desire to take this course" (Form 4, item 36).

Responses to this item, which provide a measure of the students' motivation at the beginning of the course, have proved to have a statistically potent influence on students' progress ratings, as well as on responses to other items. This seems reasonable; highly motivated students are more likely to put more effort into a course. It is not surprising that they then report getting a great deal out of it.

IV. STATISTICAL DATA

The initial form was administered in 708 classes at Kansas State University in the spring of 1969. Of these, 65 were excluded because they included fewer than 10 students. The remaining 643 were divided into "Studio" courses (N = 37) and "Non-Studio" courses (N = 606). Extensive analyses were done to determine if the two types of courses differed in any important ways. Very sizable differences were found, particularly with regard to the adequacy of the list of objectives for describing "Studio" courses. It was concluded that either a different form should be constructed for studio courses or increased flexibility would have to be introduced in the form to accommodate such courses. Because of their apparent uniqueness, studio courses were excluded from further analyses.

Between the spring of 1969 and the spring of 1975 various versions of the form were used in thousands of classes at Kansas State University, as well as in trial uses at a few other institutions. The data pool used for the fall 1975 IDEA system was based upon 3,663 classes at Kansas State University.

A. RELIABILITY

To estimate reliabilities, data were examined for 184 medium-size classes, classes enrolling 30-49 students. Students in each class were numbered consecutively. Then, for each measure, two average scores were obtained for each class, one for the odd-numbered students and one for the even-numbered students. These two "scores" were correlated for the 184 classes. The results were taken as an estimate of the reliability of the various measures when the number of observers was half the average number in these medium-sized classes.

Reliabilities were computed for six a priori scales composed of subsets of the "methods" items, for the four a priori course description scales, and for the eight progress ratings. Results are shown in Table 1.

Table 1

Estimated Reliabilities of A Priori Scales and
Progress Ratings for Various Numbers of Raters

	Number of Raters			
A Priori Scales	10	25	50	100
Prep. & Org.	.84	.93	.97	.98
St. Involv.	.84	.93	.97	.98
Clar. Comm.	.78	.91	.95	.97
Stimulation	.84	.93	.97	.98
Spk. Style	.81	.92	.96	.98
Personalism	.75	.89	.94	.97
Total Methods	.85	.94	.97	.98
Examinations	.73	.89	.93	.97
Assignments	.37	.60	.75	.86
Textbook	.61	.80	.89	.94
Content	.65	.83	.91	.95
Progress Rating				
Fact. Knowledge	.78	.90	.95	.97
Princ., Theories	.68	.85	.92	.96
Applications	.67	.84	.91	.95
Self-Understanding	.73	.88	.93	.97
Prof. Att., Beh.	.68	.84	.92	.96
Communication	.74	.89	.94	.97
Implic. for Conduct	.69	.85	.92	.96
Gen-Lib. Educ.	.73	.88	.94	.97
Progress, Rel. Objectives	.84	.93	.97	.98

The reliability of individual methods item was estimated by following the same procedure for each of these 42 items. For classes in which 25 students provided ratings, these estimated reliabilities ranged from .81 to .94, and averaged .866. Standard errors of measurement were calculated on a number of sample items and averaged approximately 0.3.

It should be pointed out that reviewers of the literature on student ratings of teaching have consistently concluded that such student ratings are reliable (Costin et al., 1971; Menges, 1973; Kulik & McKeachie, 1975). Thus it is not surprising to discover that those conclusions also apply to the IDEA items.

B. VALIDITY

Insofar as the IDEA Survey Form asks students to give the same kind of ratings as other student rating forms, the conclusions about their validity also apply to the IDEA system (Costin et al., 1971; Menges, 1973; Kulik & McKeachie, 1975). The unique aspects of the IDEA system require further discussion. The validity of student ratings of progress, the relationships between teaching methods and objectives, and the influences of class size and student motivation will each be discussed.

1. Students' Reports of Progress. In the absence of objective measures of the amount of factual knowledge gained; skills and competencies developed; interest, talents, and values developed, etc., in each course, it was impossible to test directly the validity of the students' ratings of their progress. Those studies in the research literature which bear on the issue have already been discussed in Part III., A. The only direct test reported in the literature prior to 1968 was the study by Solomon et al. (1964). Therefore, an indirect test was applied.

The indirect test which was conducted involved correlating the students' average progress ratings on each objective with the instructors' ratings of the importance of the objective.

A positive correlation should be obtained if the following assumptions were valid:

a. Teaching was effective at Kansas State University;

b. Faculty members gave careful attention to the identification of objectives for each class; and

c. Student ratings of progress were valid.

If any of these assumptions was completely erroneous, there should be no correlation between student ratings of progress and instructor ratings of importance. To the extent that any of these assumptions was only partially true, the correlation between importance and progress should be lowered. Of course, this correlation would also be attenuated by the limited (3-point) range of the importance ratings.

Results are shown in Table 2.

Table 2

Intercorrelations Between Average Progress Ratings and
Instructors' Ratings of Importance of Objectives

		Progress Ratings							
		1	2	3	4	5	6	7	8
	1	20	01	02	-25	-09	-23	-08	-25
	2	22	18	12	-16	-08	-21	-07	-11
Instructor's	3	07	09	32	-01	-01	-03	08	-24
Ratings of	4	-14	01	09	40	29	35	25	19
Importance	5	-03	01	15	17	32	23	26	-07
	6	-06	00	05	20	18	50	24	17
	7	-02	-03	15	16	27	17	26	-02
	8	-02	-03	-21	-01	-10	-03	-16	40

Note: Decimals have been omitted. Critical correlations are underlined.
N = 606 classes for which 10 or more student ratings were available.
Correlations of .11 are significant at the .01 level of confidence.

Key to objectives:
1 = Factual knowledge
2 = Principles, theories, generalizations
3 = Applications
4 = Self-understanding
5 = Professional attitudes, behaviors
6 = Effective communication
7 = Implications for personal-professional conduct
8 = General-liberal education

The eight critical correlations are underlined. All were significantly greater than zero (P<.01), ranging from .18 to .50; the average was .32. The 56 "irrelevant" correlations (progress on one objective versus importance on another) averaged +.02. These finds offer strong support for the contention that self-ratings of progress were made with acceptable validity.

2. Relationships Between Methods and Objectives. A crucial validity question for the IDEA system concerns the magnitude of the relationship between teaching methods and progress ratings. This was approached in two ways. First, multiple regression equations were developed for predicting mean student progress ratings from scores on the a priori "methods" scales; and second, item analyses were performed to determine the relationship between a given progress rating and a given teaching behavior.

a. Multiple regression study. Table 3 gives the intercorrelations among the six scales and the zero-order correlation of each scale with each of the eight progress ratings. All correlations were positive, but obviously the scales were not measuring six independent dimensions of instruction. There was a very substantial overlap among Clarity of Communication, Speaking Style, and Personalism. In general, if an instructor demonstrated positive characteristics on one dimension, he or she demonstrated positive characteristics on all other dimensions.

Table 3

Intercorrelations of Teaching Methods Scores and
Their Correlations with Average Progress Ratings

	(1)	(2)	(3)	(4)	(5)	(6)
(1) Prep. & Org.	-					
(2) Student Involv.	12	-				
(3) Clar. Commun.	79	42	-			
(4) Stimulation	56	48	71	-		
(5) Speaking Style	59	49	80	76	-	
(6) Personalism	57	59	82	69	74	-
Factual knowledge	58	18	63	55	48	47
Principles, theories	51	33	64	59	53	56
Applications	40	36	53	49	43	47
Self understanding	31	54	50	62	54	56
Prof. attitudes, behaviors	38	48	53	65	52	57
Communication	29	61	49	55	51	52
Implications for conduct	40	50	57	64	54	58
General-liberal education	26	32	38	49	39	41

Note: Decimals have been omitted.
N = 606 classes for which 10 or more student ratings were available.
Correlations of .11 are significant at the .01 level of confidence.

In the lower part of Table 3, correlations of the a priori scales with mean progress ratings are shown. These correlations make the methods scales appear somewhat more distinctive than was apparent from their intercorrelations. For example, Student Involvement was correlated substantially with progress on Communication Skills, but bore only a slight relationship to progress on Factual Knowledge. The reverse was true for Preparation and Organization. Clarity of Communication correlated higher than the other scales with the first three (cognitive) objectives, while Stimulation appeared to be the most effective scale for predicting gains on the affective objectives.

Major results from the multiple regression analyses are shown in Table 4. Only beta weights which were significantly greater than zero (P less than .05) were retained.

Each of the six scales made an independent contribution to the prediction of at least one progress rating. The Stimulation scale was a useful predictor for all eight progress ratings, and was the most important predictor for four of them: the three affective criteria and the one criterion which had both affective and cognitive elements. Clarity of Communication contributed the most to the prediction of three cognitive criteria, while Student Involvement was the chief predictor of the fourth (Communication Skills).

Table 4

Significant Beta Weights and Multiple Correlations for Each of Six Teaching Methods Scores Used to Predict Each of Eight Class Progress Ratings

	Teaching Methods Scores						
Progress Rating	Prep-Organ.	Student Involv.	Clar. Commun.	Stimu-lation	Spkg Style	Person-alism	R
Factual knowl.	146	-096	480	312	-186	-	671
Princ. theories	-	-	440	275	-	-	665
Applications	-	144	431	233	-165	-	566
Self-und.	-	270	-	399	-	126	680
Prof. att., beh.	-	161	-	471	-	151	686
Eff. commun.	-	429	127	253	-	-	679
Impl. conduct	-	229	192	396	-	-	679
Gen-lib educ.	-	-	-	390	-	138	495

Note: Decimals have been omitted.
N = 606 classes for which 10 or more student ratings were available.

The high intercorrelation among the scales was responsible for three significant negative weights: one for Student Involvement and two for Speaking Style. In these instances, the scales with the negative beta weights acted as suppressor variables. Their overlap made it possible to "suppress" some of the error variance of these variables, thus increasing their predictive potency.

b. Item analyses. A relatively elaborate scheme was developed for examining the relationship of individual teaching methods to progress ratings.

1) All classes were sorted into one of three sizes: less than 30 students (small), 30-49 students (medium), and 50 or more students (large). As was already mentioned, this was done because logic suggested that size of class may condition which methods will be effective.

2) The instructor's ratings of the importance of the first objective (Gaining factual knowledge) were then considered. All small classes for which this objective was rated "Essential" were grouped together. Collectively, they were called Group A. A similar grouping was made for large classes (Group B). Medium sized classes were ignored as were small and large classes where this objective was not rated "Essential."

3) The classes in Group A were numbered consecutively. Odd-numbered classes were placed in the test-development group, Group A-1. Even-numbered classes were placed in the cross-validation group, Group A-2. Groups B-1 and B-2 were formed in the same way from Group B, and so on through H for all eight objectives.

4) Classes in Group A-1 were then sorted into one of six categories, depending on the average progress rating on "Gaining factual knowledge". Category 1 classes reported relatively large amounts of progress, while Categories 2 through 6 reported progressively less progress on this objective. An identical process was followed for Group B-1, etc.

5) The number and percent of students in each category who said "True" or "False" to each of the first 58 items on the instrument was then determined. Chi squares and corrected contingency coefficients were computed. An item was retained: (1) if a linear trend was apparent, such that the percentage saying "True" regularly increased or decreased from Category 1 through Category 6, and (b) if the chi square value were significant beyond the .001 level, and (c) if the corrected contingency coefficient was at least .25.

6) This process was repeated for objectives 2 through 8, so that a total of 16 item analyses were performed.

These item analyses resulted in the construction of 16 empirical scales--a separate scale for large and small classes for each of the eight objectives. The cross-validation classes (those in Groups A-2, B-2, etc.) were then scored on these special scales. Finally, these scores were correlated with average progress ratings.

Of the 58 individual items, 47 were linearly related to effectiveness to a degree which was both statistically and practically significant. For the most part, whether or not an item was selected depended upon the objective being studied and the size of class. Of course, in every instance, students who described instructors of Category I classes (highly successful) gave a higher percentage of responses in the keyed direction than did their Category 6 counterparts; the median difference in percentages between these two groups was 32.4.

The set of items related to effectiveness was different for each objective and each class size. In essence the results describe 16 partially overlapping, yet distinct, "models" of the successful teacher.

The technique for selecting items was designed to insure that only highly relevant behaviors would be identified. There remained the problem of estimating the validity of the entire set (scale) of items for predicting progress on a given objective.

Classes in Group A-2 (cross-validation classes for which the instructor designated a given objective as "Essential") were "scored" on the corresponding empirical scale developed from Group A-1. The score for a given class was:

$$100 \times \frac{\text{(number of responses in the keyed direction).}}{\text{total number of responses}}$$

Scores on the empirical scale were correlated with average progress ratings.

This process was followed for each of the 16 analyses. The results are shown in Table 5. The level of predictive validity varied from .50 to .83. Only on the general-liberal education criterion, "Gaining a broad understanding and appreciation of intellectual-cultural matters (music, science, literature, etc.)," were the correlations for both large and small classes below .60. For small classes, the empirical scales for "Factual knowledge" and "Principles and theories" correlated in the low .50's with progress ratings on these objectives.

Table 5

Correlations of Empirical Scales With Progress Ratings in Cross-Validation Samples

Progress Rating	Large Classes r	N	Small Classes r	N
Factual Knowledge	.729	32	.500	100
Prin., theories	.703	46	.543	112
Applications	.644	31	.673	83
Self-understanding	.623	34	.691	90
Prof. att., behavior	.708	44	.612	104
Eff. communication	.829	26	.669	125
Implic. for conduct	.620	22	.773	60
Gen-lib. education	.549	38	.521	72
Average	.676		.623	

Note: All correlations are significantly greater than zero (P less than .01).

The degree to which a class makes progress on a given objective is undoubtedly a function of many variables, including instructor work load, the adequacy of the teaching facilities and equipment, the degree of congruence between student expectation and instructor objectives, instructor "personality" (warmth, rigidity, authenticity, etc.), and the presence or absence of disruptive or hostile students. The data of Table 5 establish that the instructor behaviors surveyed in this investigation represent one important variable related to progress. The higher the correlation the more influential this particular variable can be inferred to be.

Based upon these data, it was decided to incorporate the "methods by objectives" relationships as a permanent feature of the IDEA system. With a change from a "True-False" to a five-response format, it became possible to compute Pearson product-moment correlations instead of contingency coefficients. Empirical data and administrative experience dictated the decision to reduce the number of teaching methods to twenty and to increase the number of objectives to ten. Each of the teaching methods items had correlations of at least .30 with one or more of the ten objectives. There continued to be some variations in the correlations for different class sizes, probably because size was directly related to item variance. Tables 6, 7, 8, and 9 present these correlations for Small (1-14 students), Medium (15-34), Large (35-99), and Very Large (over 100) classes.

Table 6

Correlations Between Teaching Methods and Progress Ratings, Small Classes (1-14)

1975-76 Data Pool
(N = 659[a] classes)

Progress Ratings	Teaching Methods																			
	1	2	3	4	5	(6)[b]	7	8	(9)[b]	10	11	(12)[b]	13	14	15	16	17	18	(19)[b]	20
21	33	50	38	41	51	-20	36	57	-48	58	33	-33	36	52	58	54	53	40	-24	57
22	38	51	38	41	49	-22	37	49	-45	52	33	-26	35	49	65	48	47	31	-27	52
23	45	56	41	37	52	-30	34	50	-41	46	32	-24	37	45	66	43	39	42	-30	54
24	43	56	44	40	55	-31	40	57	-48	56	43	-27	38	49	50	56	47	49	-34	54
25	39	45	36	31	42	-17	25	44	-32	39	29	-20	35	37	52	41	32	40	-19	54
26	45	57	45	45	52	-33	39	49	-51	44	44	-29	37	41	46	45	42	36	-40	56
27	48	57	46	38	48	-31	35	46	-43	44	45	-22	40	39	54	43	38	38	-34	51
28	30	33	34	37	28	-18	32	30	-39	30	31	-29	31	30	45	26	34	07	-24	49
29	47	44	44	28	32	-18	28	30	-29	24	36	-23	41	28	52	22	22	18	-23	48
30	48	54	46	40	46	-30	36	46	-44	42	37	-30	42	42	57	38	38	38	-36	66

Note: Decimals have been omitted.

[a]For items 12 and 20 the N = 331.

[b]( ) refers to items where low scores are desirable.

Table 7

Correlations Between Teaching Methods and Progress Ratings, Medium Classes (15-34)

1975-76 Data Pool
(N = 1,285[a] classes)

Progress Ratings	Teaching Methods																			
	1	2	3	4	5	(6)[b]	7	8	(9)[b]	10	11	(12)[b]	13	14	15	16	17	18	(19)[b]	20
21	29	50	25	46	47	-20	35	61	-47	61	35	-43	30	48	62	54	47	49	-22	52
22	30	51	28	47	48	-27	37	59	-48	58	38	-45	32	50	65	53	47	43	-28	50
23	41	56	37	43	49	-36	36	55	-42	49	39	-40	37	41	65	48	36	51	-36	49
24	39	54	35	44	51	-32	39	61	-46	60	41	-44	35	42	57	55	42	56	-35	53
25	34	47	30	41	41	-19	30	55	-34	50	36	-27	30	33	56	44	30	57	-20	53
26	48	55	48	41	53	-37	42	49	-46	45	58	-44	40	36	53	44	32	42	-43	57
27	48	58	47	46	54	-34	45	57	-52	52	54	-45	42	45	62	53	-40	45	-42	57
28	40	43	45	44	42	-25	41	38	-42	38	44	-40	38	36	46	31	33	22	-35	54
29	47	39	47	28	35	-17	36	32	-28	29	47	-29	38	23	42	31	21	33	-28	44
30	52	58	53	48	55	-32	46	57	-49	53	54	-44	46	42	58	50	39	53	-41	63

Note: Decimals have been omitted.

[a]For items 12 and 20 the N = 568.

[b]( ) refers to items where low scores are desirable.

Table 8

Correlations Between Teaching Methods and Progress Ratings, Large Classes (35-99)

1975-76 Data Pool
(N = 624[a] classes)

Progress Ratings	Teaching Methods																			
	1	2	3	4	5	(6)[b]	7	8	(9)[b]	10	11	(12)[b]	13	14	15	16	17	18	(19)[b]	20
21	39	60	38	58	61	-24	55	72	-67	74	53	-38	44	69	78	68	69	51	-28	59
22	44	64	43	57	67	-36	56	71	-69	74	59	-44	48	71	75	68	68	48	-39	60
23	61	73	56	51	68	-49	53	69	-66	67	60	-48	57	64	75	66	61	61	-47	61
24	51	69	48	55	68	-38	56	71	-66	70	64	-42	48	63	75	69	62	53	-41	61
25	44	61	39	44	62	-28	53	71	-59	68	55	-33	44	56	68	66	57	61	-32	67
26	57	70	57	51	69	-49	54	62	-62	61	64	-48	53	54	68	61	53	51	-49	67
27	59	74	58	55	70	-46	59	67	-68	65	68	-46	56	62	79	67	59	49	-47	65
28	32	46	36	51	52	-28	50	47	-51	48	50	-37	36	43	56	48	45	32	-31	67
29	58	57	54	36	50	-40	40	45	43	44	54	-38	48	38	54	47	36	46	-42	54
30	64	72	64	57	68	-46	58	66	-66	65	64	-48	62	59	68	65	57	63	-50	74

Note: Decimals have been omitted.

[a]For items 12 and 20 the N = 325.

[b]( ) refers to items where low scores are desirable.

Table 9

Correlations Between Teaching Methods and
Progress Ratings, Very Large Classes (100 or more)

1975-76 Data Pool
(N = 203[a] classes)

Progress Ratings	Teaching Methods																			
	1	2	3	4	5	(6)[b]	7	8	(9)[b]	10	11	(12)[b]	13	14	15	16	17	18	(19)[b]	20
21	49	66	50	58	67	-32	56	71	-71	71	64	-36	52	69	77	68	69	52	-38	62
22	52	66	52	54	71	-41	52	67	-68	66	60	-30	57	71	75	65	68	52	-40	57
23	59	73	57	56	73	-57	54	73	-69	72	64	-52	59	73	67	71	68	71	-52	72
24	49	64	50	52	66	-40	51	67	-64	69	64	-39	50	65	69	68	65	57	-42	54
25	44	55	47	50	60	-30	50	59	-58	58	57	-24	47	52	59	59	54	58	-30	62
26	50	64	53	57	68	-58	58	69	-69	68	63	-56	54	67	65	70	65	63	-61	74
27	53	72	54	59	63	-41	55	69	-67	67	66	-48	53	64	79	71	63	50	-47	60
28	27	48	35	52	50	-41	48	52	-57	53	47	-47	38	51	55	52	51	32	-51	64
29	41	53	46	37	46	54	38	46	-44	47	55	-60	40	44	45	48	41	41	58	62
30	58	68	60	56	68	-49	57	69	-69	73	65	-56	57	68	60	72	68	72	-56	78

Note: Decimals have been omitted.

[a]For items 12 and 20 the N = 104.

[b]( ) refers to items where low scores are desirable.

Table 10

Correlations Between Progress Ratings
and Selected Variables by Class Size

	Class Size																			
	Small (N = 327)					Medium (N = 715)					Large (N = 298)					Very Large (N = 99)				
Progress Ratings	31	32	33	35	36	31	32	33	35	36	31	32	33	35	36	31	32	33	35	36
21	05	09	17	31	43	-02	17	23	42	56	07	18	36	58	62	02	33	30	62	66
22	19	19	40	47	50	02	23	32	51	53	05	20	40	60	59	05	36	39	63	59
23	24	31	37	48	50	04	26	18	45	49	03	23	22	46	54	07	36	16	41	50
24	-18	11	00	20	41	-24	21	-02	35	61	-16	27	23	52	63	-03	44	22	54	71
25	19	22	34	42	44	-11	13	09	35	60	-06	06	15	36	57	01	18	14	35	65
26	-16	07	-04	21	39	-26	15	-14	29	49	-26	15	-05	32	68	05	24	-00	34	61
27	01	20	16	35	46	-12	26	00	44	54	-14	34	16	51	63	10	50	27	62	54
28	13	-00	18	26	35	-03	-11	-04	18	38	-20	-00	-01	20	52	06	-02	05	25	37
29	36	31	32	44	38	12	12	-10	19	25	07	18	-09	23	42	06	13	-12	13	39
30	15	17	25	41	57	-07	04	-09	28	58	-11	06	-09	24	69	09	06	-12	23	71

Note: Decimals have been omitted. Item numbers refer to items on IDEA Survey Form (Form 4)

3. Other Effects: Class Size and Student Motivation. A major objection to the approach typically used by student rating forms was their implication that a single, monolithic model of "good" college teaching existed. When student ratings of progress were used as the criteria for effective teaching, the IDEA system replaced the single model with a multitude of models (the number of possible combinations that could be made from the ten objectives). This constituted a significant step toward providing a standardized approach which was sensitive to the variations among courses.

As the data were gathered, they were examined for other variables which might significantly influence students' ratings, especially of their progress on the instructor's objectives. The section on Item analysis has already described what was done with respect to class size after it was determined that this factor did effect student ratings. While a number of variables which might affect progress ratings (besides the 20 teaching methods items) were investigated, only eight remain on the present IDEA Survey Form (Form 4): four are related to The Course (items 31-34), and four describe the student (items 35-38); item 39 was not used until the fall of 1975.

Of these eight variables, two are simply measures of affective outcomes which are assumed to be desirable for all instructors, regardless of other objectives:

37. I would like to take another course from this instructor.

38. As a result of taking this course, I have more positive feelings toward this field of study.

It can be argued that item 34:

34. Degree to which the course hung together (various topics and class activities were related to each other).

describes a teaching approach or method, since it provides a measure of the degree to which the instructor was able to integrate the course.

There remain then five items of variables which require examination: amount of reading (item 31), amount of non-reading assignments (32), difficulty of the subject matter (33), student effort (35), and student motivation to take the course. (See Form 4 for the actual wording of these items).

Table 10 presents the correlations of each of these five items with the ten progress ratings for each of the four class sizes. An examination of Table 10 will reveal that student motivation (36) has the highest correlations with the progress ratings. (There are four exceptions, out of the 200 correlations presented, where student effort (35) has higher correlations.) The decision to control for student motivation in the IDEA system was based upon this kind of data, which clearly demonstrated the strong influence of student motivation on progress ratings. Additional adjustments for responses to others will be considered in future revisions.

When Form 2 and Form 3 were introduced in 1972 and 1973 respectively, the comparison groups simply used two levels of motivation (above and below the mean). When the present version of the IDEA system was developed in the fall of 1975, the decision was made to use five levels of motivation. This decision was based on statistical evidence that a more precise consideration of the motivation measure was desirable. The score ranges for each motivation level were based on the senior author's clinical judgment. Approximately 22% of the classes fell into each of the first four levels: below 3.0, 3.0-3.4, 3.5-3.9, and 4.0-4.4. Only 12% fell in the highest level 4.5-5.0.

Because the IDEA system was based on the strategy which required comparisons with similar classes, the size of the data pool severely limited the number of variables which could be taken into consideration. The original data pool was large enough to accommodate two variables. For logical and statistical reasons class size and student motivation were chosen as the two variables to use. As the data pool is increased, it will be possible to consider other variables, e.g., student effort, amount of reading. Whether or not these variables will make an independent contribution to the "explanation" of the variance in student ratings remains to be seen; the amount of variance accounted for by class size and student motivation is already quite sizable.

REFERENCES

Bloom, B. S. (Ed.). Taxonomy of educational objectives: cognitive domain. New York: David McKay, 1956.

Combs, A. W., Soper, D. W., & Courson, C. C. The measurement of self-concept and self-report. Educational and Psychological Measurement, 1963, 23, 493-500.

Costin, F., Greenough, W. T., & Menges, R. J. Student ratings of college teaching: reliability, validity, and usefulness. Review of Educational Research, 1971, 41, 511-535.

Deshpande, A. S., & Webb, S. C. Student perceptions of instructor teaching goals. III. Internal structure of ratings (Research Memorandum 68-5). Atlanta: Office of Evaluation Studies, Georgia Institute of Technology, 1968.

Holland, J. L., & Lutz, S. W. The predictive value of a student's choice of vocation. Personnel and Guidance Journal, 1968, 46, 428-434.

Hoyt, D. P. Identifying effective instructional procedures. Improving College and University Teaching, 1973, 21, 73-76. (a)

Hoyt, D. P. Measurement of instructional effectiveness. Research in Higher Education, 1973, 1, 367-378. (b)

Hoyt, D. P. The Kansas State University program for assessing and improving instructional effectiveness. In A. L. Sockloff (Ed.) Proceedings: faculty effectiveness as evaluated by students. Philadelphia: Measurement and Research Center, Temple University, 1973. (c)

Isaacson, R. L., McKeachie, W. J., Milholland, J. E., Ling, Y. G., Hofeller, M., Bairwaldt, J. W., & Zinn, K. L. Dimensions of student evaluations of teaching. Journal of Educational Psychology, 1964, 55, 344-351.

Keefer, K. E. Self-prediction of academic achievement by college students. Unpublished doctoral dissertation, University of Tennessee, 1965.

Krathwohl, D. R., Bloom, B. S., & Masia, B. B. Taxonomy of educational objectives: affective domain. New York: David McKay, 1964.

Kulik, J. A., & McKeachie, W. J. The evaluation of teachers in higher education. In F. N. Kerlinger (Ed.), Review of research in education (Vol. 3). Itasca, IL: Peacock, 1975.

Menges, R. J. The new reporters: students rate instruction. In C. R. Pace (Ed.), Evaluating learning and teaching. San Francisco: Jossey-Bass, 1973.

Solomon, D. Teacher behavior dimensions, course characteristics, and student evaluations of teachers. American Educational Research Journal, 1966, 3, 35-47.

Solomon, D., Rosenberg, L., & Bezdek, W. E. Teacher behavior and student learning. Journal of Educational Psychology, 1964, 55, 23-30.

REFERENCES (Continued)

Walsh, W. B. Validity of self-report. Journal of Counseling Psychology, 1967, 14, 18-23.

Whitlock, G. H. The evaluation of teaching effectiveness. Proceedings, Sixth Annual Forum of the Association for Institutional Research. May, 1966.

APPENDIX A

THE FOUR VERSIONS OF THE IDEA SURVEY FORM

Form 1: Student Reactions to Instructors and Courses, 1969.

Form 2: Student Reactions to Instruction and Courses (2nd ed.), 1972.

Form 3: Student Reactions to Instruction and Courses--Short Form, 1973.

Form 4: IDEA Survey Form--Student Reactions to Instruction and Courses, 1975.

APPENDIX B

Items Comprising A Priori Scales

TEACHING METHODS

Preparation and Organization

1. The instructor seemed to have a well developed plan for each class session.
13. On several occasions, he seemed unprepared for class.
19. He made it clear how each topic fit into the course.
25. He failed to state clearly the course requirements and deadlines.
31. He stated clearly the objectives of the course.

Student Involvement

2. There were discussions between teachers and students (as opposed to mere responses to questions).
8. The instructor seemed to lack energy.
14. Students made comments to the instructor without being asked.
18. He requested and obtained student's questions and reactions.
20. He encouraged student comments even when they turned out to be incorrect or irrelevant.
26. He attempted to induce silent students to participate.
32. He became angry or sarcastic when corrected or challenged by a student.

Clarity and Communication

3. He explained course material clearly, and explanations were to the point.
5. The instructor answered student questions as completely as reasonable.
9. He was often incoherent and/or vague in what he was saying.
21. He presented examples of what he wanted by way of homework, papers, etc.
27. He summarized material in a manner which aided retention.
33. He introduced stimulating ideas about the subject.

Stimulation

4. The instructor seemed to lack energy.
10. The instructor seemed enthusiastic about the subject matter.
16. He demonstrated the importance and significance of his subject matter.
22. He sometimes presented material in a humorous way.
28. He stimulated students to intellectual effort beyond that required by most courses.
34. He introduced stimulating ideas about the subject.
37. He related course material to real life situations.

Speaking Style

11. He generally spoke too rapidly.
15. He spoke with expressiveness and variety in tone of voice.
17. His presentations were dry and dull.
23. He explained the reasons for his criticisms of students' academic performance.
29. He lectured in a rambling fashion.
35. He repeated material to the point of monotony.
39. He was available for individual help.

TEACHING METHODS (Continued)

Personalism

6. He adjusted his pace to the needs of the class.
12. He changed his approach to meet new situations.
24. He explained the reasons for his criticisms of students' academic performance.
30. He understood student comments and questions even when these were not clearly expressed.
36. He displayed favoritism.
38. He was available for individual help.
40. He often dismissed class late.

COURSE REACTION

Exams

43. The examinations gave a balanced coverage to major topics.
47. Examination questions were often unclear.
51. Examinations stressed memorization of information for which later recall seems unreasonable.
55. Examination questions were frequently too detailed or picky.

Assignments

44. The instructor gave ample notice for lengthy assignments.
48. Out-of-class assignments were reasonable in length.
52. Assigned readings were pertinent to the topics presented in class.
56. I usually had no difficulty in obtaining outside reading materials.

Textbook

45. The textbook (or substitute reading materials) seemed out-of-date to me.
49. The textbook (or substitute reading materials) contained too little illustrative material.
53. Assigned readings (including text) were reasonably clear and understandable.
57. Reading materials (including text) were organized in a logical, orderly fashion.

Content

46. Too much of the course material repeated content covered by courses I had taken previously.
50. Too much time was spent on too few topics -- the course needs more breadth.
54. The instructor failed to make clear the relationship between this course and other courses.
58. There were too many topics to understand any of them well.