DOCUMENT RESUME

ED 273 169 HE 019 506

AUTHOR MacCuish, Donald A.
TITLE The Course Development Model in Higher Education: Improving Tests and Instruction.
PUB DATE 86
NOTE 20p.; For related document, see HE 019 505.
PUB TYPE Viewpoints (120)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Accountability; Classification; *College Faculty; Course Descriptions; *Course Objectives; *Educational Quality; Faculty Evaluation; Higher Education; Instructional Improvement; Student Evaluation; *Teacher Made Tests; Test Construction; *Test Reliability
IDENTIFIERS Course Development

ABSTRACT

A quality assurance model for assuring quality in the development of course objectives and classroom and exit examinations is presented. Called the Course Development Model, it was an outgrowth of a 1985 pilot study conducted at the University of Central Florida. The study found that more than 100 faculty who completed a questionnaire did not properly design nor adequately evaluate the tests they developed for their classes. Objectives of the model are to improve accountability, develop better testing instruments, and improve the instructional process. The model requires that course descriptions contained in college catalogs describe course content in concrete terms that would enable an instructor to derive sound course objectives. The course objectives can then be used to develop the tests and measuring devices used to evaluate student performance and learning. The model requires that a test construction taxonomy, such as Bloom's, be used as a guide to develop these tests. The final step is the evaluation of tests and measuring instruments for reliability and validity. The model promotes accountability without requiring the faculty member to give up independence. Use of the model will not only improve evaluation of student performance, but will help the professor improve instruction. The model will also be helpful in faculty evaluation. (SW)

The Course Development Model in Higher Education: Improving Tests and Instruction

Donald A. MacCuish
Manager
Southern Operations
ONLINE Computer Systems
20251 Century Blvd.
Germantown, MD 20874

If humans are born equal or can become equal with regard to learning, then the home and the school have responsibilities far greater than they have assumed in the past. (Benjamin Bloom,1976)

-Abstract-

Entrance and exit examinations play a very important role in higher education in the United States. For many the entrance examination has been used to exclude them from the world of higher education. As a result, we have prevented a portion of our population from traveling the road of "upward mobility" upon which many of our grandparents traveled. The exit examination is meant to evaluate those who have had the opportunity to travel that road, but in too many cases the examination and the procedures surrounding it are of questionable integrity.

This paper centers on the Quality Assurance Model, commonly referred to as the Course Development Model (the Model), and describes how it can be used to foster good test development and improve instruction in institutions of higher education. The Course Development Model as first discussed at the Testing and Quality Assurance in Higher Education Conference in Miami in February 1986 showed how the Model can develop faculty accountability as well as being used as one item in the faculty evaluation process.

The Model suggests that there is a gap between the stated curriculum and the method or methods of instruction used by the professor in the classroom. The Model bridges this gap in a logical fashion in that the professor provides an audit trail. In addition, the Model refines test development, fosters test evaluation, and encourages the improvement of instruction.

In the Spring of 1985, a pilot study was conducted at the University of Central Florida (UCF). The Model is an outgrowth of that study. Analysis of the data shows that 91% of the faculty at UCF develop 100% of tests used in the evaluation of student performance. One hundred percent of the faculty do do some test development as opposed to relying solely on department made, professional examinations, or standardized tests. What is interesting about this fact is that the faculty surveyed do not properly design nor adequately evaluate the tests they develop.

This paper analyzes the data in greater detail than was discussed at the Quality Assurance in Higher Education Conference. It is also used as a vehicle to discuss the relationship of test design and evaluation to the process of improving instruction in institutions of higher education.

It is quite possible that if this Model were to be followed by the teaching faculty at institutions of higher education they would become accountable to their superiors, control one of the elements in the faculty evaluation process, design high quality tests, and improve their instruction without any loss of independence. The Model provides them with a squential means of transition from curriculum to instruction. It is simple to follow and requires little additional effort. Most faculty, it

appears, at UCF, are already following portions of the Model. We only have to close the loop. If we do the result will be better examinations and improved instruction.

INTRODUCTION

In his opening remarks to the Testing and Quality Assurance in Higher Education Conference in Miami, Florida on 12 February 1986, Dr. Robert McCabe, the President of Miami-Dade Community College, told the assemblage that assessment was a growing issue with both state legislatures and the public. There is, he said, a mistrust of what educators are doing. We as educators must ask ourselves the question: "Do we really know what we are doing?" Our goal should be the assurance that more people are achieving higher competencies. Assessment then, according to Dr. McCabe, is of enormous importance. Assessment can and must have a positive impact on programs, curricula, and instruction.

Kenneth B. Clarke, the keynote speaker, followed Dr. McCabe at that same conference. He said that part of the problem was in the concept of diagnostic versus rejective use of tests and measurements. We have not explored this relationship, nor have we expounded on it, nor have we integrated the differences between them in our educational system. In order to increase the opportunities of higher education for all our citizens, improve the instructional process, and develop diagnostic measures to evaluate learning we must review and analyze the relationship between curriculum and instruction and find methods for strengthening this association.

COURSE DEVELOPMENT MODEL

There are many reasons for tightening this relationship between curriculum and instruction. The first is to improve accountability which is one of the

growing issues in higher education. A second is to develop better testing instruments. The third reason is to improve the instructional process itself.

Glenny and Schmidtlein (1983) state that state legislatures and departments of higher education will scrutinize the budgets for state supported colleges and universities more closely as the student population continues to decline and as the availability of funds becomes more scarce. They believe that one method of control will be intrusion into the educational process itself. The most likely areas of this invasion will be to develop criteria for the evaluation of new and existing academic programs and faculty effectiveness. A closely related area will be the oversight of exit evaluations of student preparedness to demonstrate competence in their fields of specialization.

How can educators contribute to this inevitable process to preserve academic freedom, protect the interests of students, and at the same time provide legislatures and the citizenry as a whole the assurance they want. The Course Development Model meets all of these objectives and provides a high degree of accountability discussed in this author's presentation to the Quality Assurance in Higher Education Conference in February 1986. The Model, depicted in figure 1, requires that course descriptions contained in college and university catalogues sufficiently describe the course content in concrete terms that would enable an instructor to derive sound course objectives from them. The course objectives can then be used to develop the tests and measuring devices used to evaluate student performance and learning. It is important to note, however, that the model requires that a test construction taxonomy, such as Bloom's, be used as a guide to develop these tests. The final step is the evaluation of tests and measuring instruments for reliability and validity. The

cycle is completed only when the evaluation of the tests leads to improvement of the instructional process.

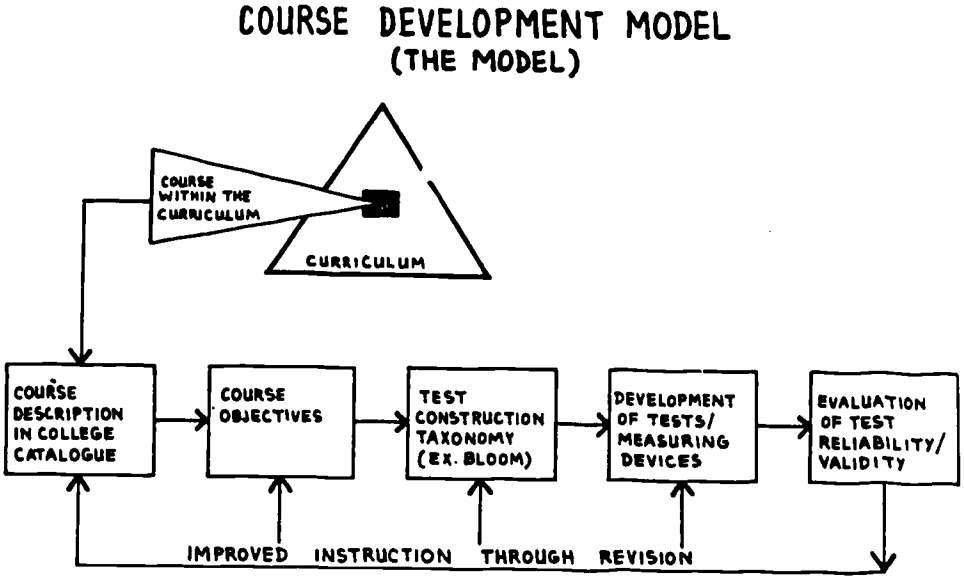

Figure 1

The obvious question at this juncture is why do we need such a model? The intrusion of state legislatures, boards of trustees of private colleges and universities, and the public at large are not sufficient justification for this additional work which will be required as the benefits will be minimal. Quite the contrary!

When we review the various theories for development of curriculum we notice that the theorists have not provided us with the means to get to the instructional process. When we review the theories of instruction we are not provided with the how to base the instruction on the information provided by the curriculum. There is a gap which clearly emerges when we view these two side-

by-side. The Course Development Model provides the administrator, faculty member, and evaluator with a clear, easy to follow procedure and audit trail leading from the curriculum to the instruction. This model does not impact on the way the curriculum is developed, nor does it infringe on the faculty member and his or her method of instruction. It does, however, insure that the two are well connected and interrelated.

PURPOSE AND BACKGROUND

Dr. Clifford Aldeman of the National Institute of Education says that college faculty are not trained in the design, development, and evaluation of tests. If they were trained in these areas the instructional process would be improved. His statement suggests that there is a need for some kind of faculty development program.

Elbe (1972) suggests that faculty development should be an integral part of all institutions of higher education. The Course Development Model provides a tool for helping an institution identify areas for faculty development. Based on Gaff's (1975) findings the input factor of the model is important.

> "The kinds of change that emanate from instructional improvement programs are more in the areas of process, teaching methodology and techniques, learning materials, and interpersonal relationships. And the changes that do occur will probably have more impact directly on individuals and small groups than indirectly through changes in organizations as a whole" (p. 164).

Of the areas Gaff identified as needing attention, the Model will successfully impact on design of new courses, redesign of existing courses, and improvement of instructional materials.

Carlberg (1981) states that about 15% of the instructional time available to teachers is devoted to testing activities. Therefore, it is essential, that

that faculty know how to develop good measuring devices. If we assume that Dr. Adelman is correct then one major area for faculty development is in the area of test development. Carlberg also indicates that careful consideration of testing is important because of the impact testing has on the planning of instructional strategies.

Hopkins and Stanley (1981) state that research has shown that tests can improve student learning. As a result, measurement specialists have identified a large number of strategies for employing tests in the classroom. Goslin (1967) supports the idea that faculty have instruction on how to develop high quality instruments. The Bloom Taxonomy provides the faculty member with a proven methodology and process for test design, and development.

Knapper (1979) states that effective instruction is closely linked to effective evaluation. The problem, according to Knapper, is few instructors evaluate the way instruction is presented. He also notes that teaching practices and learning effectiveness can be substantially changed if the evaluation process is conducted properly. Instructors must be willing, something which all too many are unwilling to do, to evaluate their current teaching practices. The Course Development Model forces faculty members to evaluate the way in which they teach their students.

Bloom (1976) says that most students can learn what most schools have to teach if the subject matter is approached in a systematic and sensitive fashion. Variations in levels of learning by students is based not only on the student's learning history, but also by the quality of instruction they receive. "In general, we can estimate that quality of instruction can account for at least one fourth (r=+.50) of the variance on relevant cognitive achievement measures" (p. 135).

Thus, it is incumbant upon educators to design, develop, model, and use a systematic sensible approach to linking curriculum with instruction in order to transit the gap between the two. Only by this linkage can we accurately and fairly evaluate students, and improve our instruction. The Course Development Model provides the tools necessary to accomplish these tasks. Do faculty use this or a similar model now, or is such a suggestion simply a theoretical fantasy? To answer this question a pilot study was conducted at the University of Central Florida in the Spring of 1985. The results of that study suggest that this model is realistic, feasible, and could be readily implemented with minimal behavior changes required on the part of the teaching faculty.

ANALYSIS OF THE DATA

This discussion of the research is the second of a two part analysis of the data obtained by this pilot project. The first part of the analysis is contained in a previously published paper, "A Quality Assurance Model for Higher Education", and is available through the Educational Resources Information Center (ERIC). The second part analyzes the data as it pertains to testing and improvement of instruction.

All four hundred five members of the teaching faculty at the University of Central Florida were requested to particpate in this study by completing a survey on the use of statistics by college faculty. One hundred and thirty-one completed surveys were returned. This represents a thirty-one percent return rate. Although this sample is biased, it does represent one third of the population, thus it is assumed to accurately represent the attitudes and behaviors of the UCF faculty. The breakdown of returned surveys by colleges within the

university is as follows: Arts and Sciences - 44%, Business Administration - 18%, College of Education - 21%, College of Engineering - 12%, and College of Health Sciences - 5%.

The first statistical procedure to be performed was a frquency distribution which provided the moments of the distribution and other statistical information. Table One contains this data for the review of the reader. For those cases when the between group significance was significant at the 0.5 level it is so indicated. I would like to call to the reader's attention that there is a significance between groups in the areas of "using course objectives to develop tests" and "using a taxonomy for test development". These are brought to your attention because the Course Development Model stresses their importance.

An important element of the model is the generation and use of teacher made tests by faculty in the evaluation of student knowledge gained in the classroom. Over ninety-one per cent of the faculty design and develop all of the tests and measuring devices they use to evaluate student performance. The mean was 97.672. What makes this of great concern to us as educators is the statement by Dr. Aldeman concerning the need for faculty training or test procedures. The question in the survey which asks about the use of a taxonomy had a mean score of 3.33. This substantiates Dr. Aldeman's position. Further research will help determine whether or not this is unique to UCF or typical of higher education. I believe the latter to be correct.

The number of tests developed for each course was found to be 4.12. The average number of tests actually given in each course averaged 4.07. The difference in these two is assumed to be a result of rounding. The median and mode

	MEAN: ARTS & SCIENCES	MEAN: BUSINESS ADMINISTRATION	MEAN: EDUCATION	MEAN: ENGINEERING	MEAN: HEALTH	GROUP MEAN	STANDARD ERROR/ DEVIATION	SKEWNESS	KURTOSIS	BETWEEN GROUP SIGNIFICANCE
TEACHER MADE TESTS	98.42	95.00	98.14	100.0	93.57	97.67	0.828 9.473	-5.034	27.58	NO
NUMBER OF TESTS CREATED/COURSE	4.228	4.250	3.407	3.875	6.142	4.122	0.187 2.141	1.901	6.090	NO
NUMBER OF TESTS GIVEN/COURSE	4.105	4.041	3.666	3.750	6.142	4.068	0.172 1.974	1.976	7.420	NO
HOW FREQUENTLY DO YOU CURVE	2.631	2.750	3.030	2.250	2.714	2.694	0.093 1.066	-0.328	-1.113	NO
COURSE DESCRIPTION TO COURSE OBJ	2.122	1.833	1.481	2.000	1.857	1.908	0.099 1.133	0.827	-0.842	NO
USE COURSE OBJ TO DEVELOP TESTS	1.754	2.041	1.370	2.062	1.285	1.740	0.084 0.957	1.131	0.219	YES
USE A TAXONOMY FOR TEST DEVELOP	3.789	3.625	2.074	3.750	2.428	3.328	0.089 1.019	-1.186	-0.055	YES
OBTAIN RELIABILITY FREQUENTLY	3.122	3.291	2.851	3.250	2.142	3.061	0.093 1.065	-0.782	-0.704	NO
USE RELIABILITY DATA	3.140	3.250	2.814	3.375	2.428	3.084	0.091 1.045	-0.744	-0.767	NO
AVERAGE RELIABILITY	3.140	3.541	2.814	3.687	2.714	3.190	0.107 1.229	-0.928	-0.867	NO
VALIDITY ONE	3.526	3.875	3.222	3.937	2.571	3.526	0.108 1.236	-0.745	-0.388	YES
VALIDITY TWO	0.754	0.416	1.074	0.500	0.857	0.732	0.124 1.419	1.617	1.006	NO
VALIDITY THREE	0.263	0.375	0.407	0.000	0.428	0.290	0.083 0.973	3.157	8.377	NO
CALCULATE STAT VALIDITY	3.543	3.458	2.888	3.437	2.285	3.313	0.079 0.904	-0.981	-0.276	YES
AVERAGE VALIDITY	4.596	4.625	3.481	4.375	3.285	4.274	0.121 1.387	-1.472	0.357	YES
DO ITEM ANALYSIS	3.403	3.041	2.333	3.312	2.285	3.045	0.095 1.087	-0.639	-1.047	YES
REVISE TEST BASED ON ITEM DISCRIM PWR	3.386	3.250	2.592	3.625	2.571	3.183	0.081 1.036	-0.755	-0.957	YES
OBTAIN STANDARD ERROR	3.596	3.583	3.222	3.312	2.285	3.412	0.081 0.927	-1.503	1.173	YES
OBTAIN STANDARD DEVIATION	3.280	3.166	2.925	3.000	2.142	3.091	0.101 1.160	-0.872	-0.814	NO

TABLE 1

for each is 4.00. The Pearson Correlation between the two scores was .72 and significant to the .05 level.

In order to look at the relationship of the variables comprising the Model, correlations between each were calculated. Table 2 shows the results of this investigation. The individual relationships are all significant except the correlation between using objectives to develop tests and obtaining reliability tests and measuring devices by faculty. Thus we can conclude that faculty indicate a desire to bring curriculum and instruction closer together but lack the Model to do so.

MODEL: CORRELATION MATRIX*

	1	2	3	4
1. DESCR				
2. OBJ	43+			
3. BLOOM	26+	23+		
4. ORELI	19+	12	28+	
5. URELI	16+	18+	28+	89+

* Decimals omitted
\+ Significant at .05

TABLE 2

Three separate three variable log linear analyses were conducted to test the strength of the relationships between selected variables. The first concerned, (1) using course descriptions to develop course objectives, (2) using the course objectives to develop tests, and (3) using a taxonomy, such as Bloom's, to design the testing instrument. The second was (1) using course

descriptions to develop course objectives, (2) using course objectives to develop tests, and (3) obtaining the reliability of the tests used in the classroom. The third was similar to the second with the substitution of using reliability to improve tests in place of obtaining the reliability of the tests. The interpretation of the log linear analysis depicted in Table 3 shows that Model 4 (variables 1 and 2 impact on 3) satisfies the goodness of fit test in all three cases. Thus, faculty at the University of Central Florida do use Course Descriptions to Develop their Course Objectives. They use their Course Objectives to develop their tests. These two variables impact on the use of a taxonomy to actually develop the tests, and on both obtaining the reliability of the tests, and on using the reliability of the tests to improve instruction. The problem, however, is that the impact is in a negative direction. If the faculty member follows the first two steps in the process he or she is less likely to follow through and use a taxonomy, obtain the reliability, or use it once obtained. In order to improve both the tests used in higher education and the instruction, faculty members must be trained in the use of a taxonomy, how to obtain the reliability of their tests, and what to do with this information once they have it.

Next a four variable log linear analysis was performed. The four variables used were: (1) using course descriptions to develop objectives, (2) using the course objectives to develop the tests used in the course, (3) using a taxonomy, such as Bloom's, to design the tests, and (4) obtaining reliability information on the tests. Heeding the advice of Dr. John Kennedy, the log linear analysis was performed on the last eight of the sixteen models (see Table 4). Analysis of the results show that Model 0, the null hypothesis model, satisfies the goodness of fit test. "If this model fits observed data reasonably well, we must conclude that there are no effects of sufficient magnitude

LOG LINEAR ANALYSIS

MODEL	2	df	VARIABLES	COMPONENT	df
0	94.90447	66			
1	71.97225	60	DESCR	22.93219	6
2	53.82140	57	OBJ	18.15088	3
3	36.08543	54	BLOOM	17.73597	3
4	8.08543	45	DESCR-OBJ	28.00401	9
5	5.37327	36	DESCR-BLOOM	2.70815	
6	2.24921	27	OBJ-BLOOM	3.12406	
7	0	36	DESCR-OBJ-BLOOM	2.24921	

3.a.

0	134.57344	62			
1	103.95084	60	DESCR	30.62260	2
2	65.45541	52	OBJ	38.49543	3
3	49.37898	54	ORELI	16.07643	3
4	15.86690	45	DESCR-OBJ	33.51208	9
5	8.16124	36	DESCR-ORELI	7.70566	9
6	7.50713	27	OBJ-ORELI	.065401	9
7	0	22	DESCR-OBJ-ORELI	7.50723	5

3.b.

0	126.38297	62			
1	92.30995	60	DESCR	34.07320	2
2	61.30515	57	OBJ	31.0048	3
3	45.04537	54	URELI	16.25978	3
4	14.44798	45	DESCR-OBJ	30.59739	9
5	8.30819	36	DESCR-URELI	6.13979	9
6	6.93657	27	OBJ-URELI	1.37162	9
7	0	23	DESCR-OBJ-URELI	6.93657	4

3.c.

TABLE 3

to be entertained. Adoption of the null model is analogus to performing a three-way ANOVA and discovering that, of the seven omnibus F tests, not one has achieved statisitcal significance" (Kennedy, p.166).

Hierarchical Logit Models for Four-Way Tables Where D Is the Logit Variable

Logit Model No.	Log-Linear Logit Models	Marginals Fitted
(0)*	λ +a+b+c+d+ab+ac+bc+abc	ABC,D
(1)	λ +a+b+c+d+ab+ac+bc+abc+ad	ABC,AD
(2)	λ +a+b+c+d+ab+ac+bc+abc+ad+bd	ABC,AD,BD
(3)	λ +a+b+c+d+ab+ac+bc+abc+ad+bd+cd	ABC,AD,BD,CD
(4)	λ +a+b+c+d+ab+ac+bc+abc+ad+bd+cd+abd	ABC,ABD,CD
(5)	λ +a+b+c+d+ab+ac+bc+abc+ad+bd+cd+abd+acd	ABC,ABD,ACD
(6)	λ +a+b+c+d+ab+ac+bc+abc+ad+bd+cd+abd+acd+bcd	ABC,ABD,ACD,BCD
(7)	λ +a+b+c+d+ab+ac+bc+abc+ad+bd+cd+abd+acd+bcd+abcd	ABCD

Note. Constituent terms (parameter estimates) in models are denoted by their respective superscripts with the exception of the first term, .

*Numbers assigned to logit models in this table no longer reflect the number of constituent terms. They do, however, reflect the number of interaction terms that involve the logit variable. (Kennedy, p. 164)

a = DESCR b = OBJ c = BLOOM d = ORELI

TABLE 4

At this point something must be said about the log linear analysis. A four variable analysis requires a great deal of computer memory. The running of this particular study literally shut down the computer center for over a half hour. No other programs were run during the time that the computer made the calculations for this analysis. Later a stepwise regression analysis was run which provided more tangible interpretations of the data. The other factor worth noting is that with a sample size of only one hundred thirty-one (131) meant that a large number of cells were empty.

Changing the dependent variable to "using course descriptions to develop course objectives", a stepwise-multiple regression reveals that both the "use of course objectives to develop tests" and "use of a taxonomy, Bloom for instance", does correlate well with the course use of course descriptions variable. The correlation is statistically significant (r=.43). The two indicies of reliability were unimportant. As a result, we can say that faculty members who use course objectives to develop tests and use a taxonomy for test construction are in all likelihood going to use the course descriptions to develop their course objectives. The problem is we can only account for 19% of the variance which is not encouraging.

The last two statistical procedures used in this study are a multiple regression and a stepwise multiple regression. Analysis of the data shows that there is a correlation among the four variables being considered. The stepwise multiple regression shows exactly what this correlation looks like. The R squared value is .1. The stepwise multiple regression reveals that the relationship of using a taxonomy, like Bloom's, on reliability is only .08 (R squared value). Thus, the other two independent variables have no impact at all and the impact of Bloom is minor although it is statistically significant.

FACULTY DEVELOPMENT

Analysis of the data leads to some very interesting interprteations which can have a strong impact on testing and improvement of instruction in higher education. At the present time faculty members do not have a theoretical model which enables them to move from curriculum to instruction in a logical, sequential, and auditable fashion (figure 2). Yet, behaviorally, they exhibit some of

the behaviors called for in the Course Development Model. What is needed is instruction on how to bridge the gap between the curriculum and the professor's preferred method(s) of instruction in the classroom.

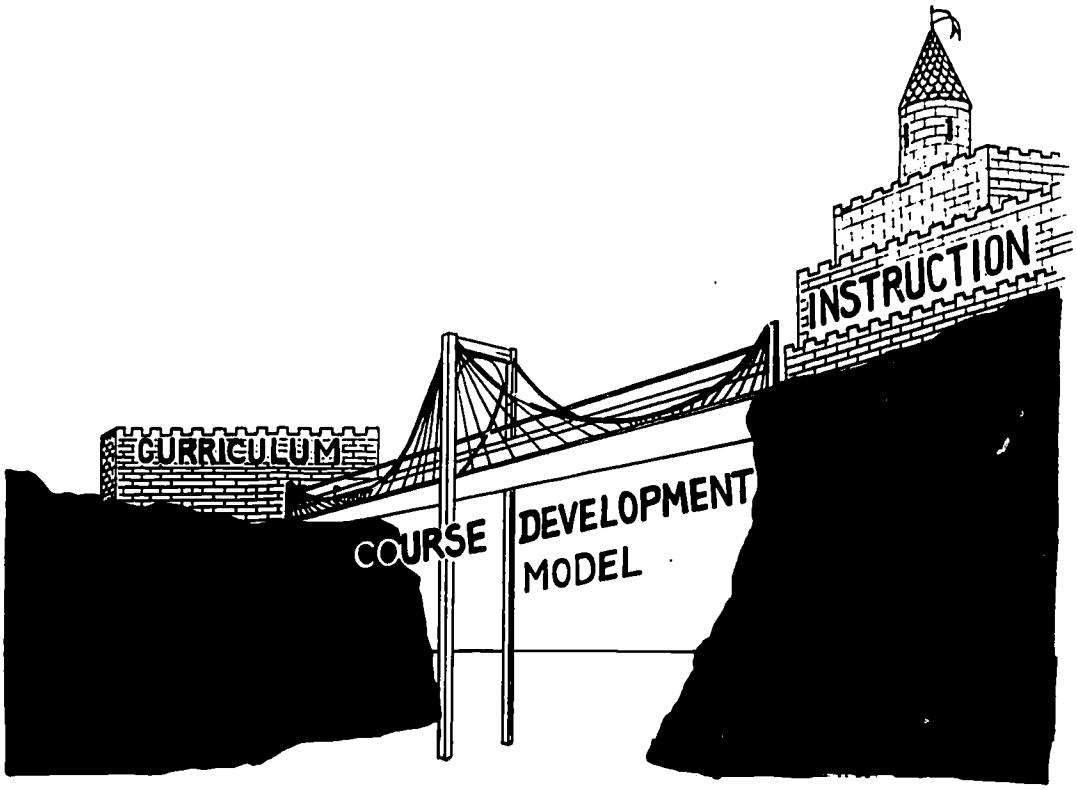

Figure 2

For the most part, the behavioral changes are minor. Many faculty members already use the course descriptions contained in university and college catalogues to develop their course objectives. Both test development and instruction can be improved if college faculty will consistantly perform this task. For the most part it is a simple documentation procedure. Next, faculty members must consistantly use their course objectives to develop the tests and other measuring devices used in their classes. By using course objectives the tests should have content validity because of the process by which they are developed.

Faculty members must also be informed about taxonomies of educational objectives such as that developed by Benjamin Bloom and his associates. By using a taxonomy the tests which are used will be properly targeted to a domain of learning. In addition to the logical sequence of steps used and the audit trail they provide, the tests will have construct validity, since the tests will have been developed to test the construct. This also will significantly improve the the tests used in the classroom.

Finally, by teaching faculty how to obtain reliability information on their tests and use this information the entire course development process will be improved. The reliability information can be used to not only improve the tests themselves, but to improve instruction by having the professor examine whether or not the instruction meets the requirements of the learner. If properly schooled, the instructor can use the reliability information to improve the course descriptions and or objectives themselves.

CONCLUSION

But, why should college faculty adopt such a model? There are several reasons. First, the Model makes good pedagogical sense. Second, the public is demanding accountability. This Model provides for that accountability without requiring the faculty member to give up the independence now enjoyed. In addition, the Model does not require the writing of a new curriculum, nor does it mandate a change in the instructional practices now employed. Third, use of the Model will not only improve tests and other forms of evaluating student performance, but it will also help the professor improve instruction. Lastly, this model satisfies the state legislatures' desire for accountability and evaluation of faculty, provides a methodology for improving tests and instruction, yet preserves faculty independence.

References

Bloom, Benjamin S., et al. A Taxonomy of Educational Objectives (New York: Longman Publishers, 1956).

Bloom, Benjamin S. Human Characteristics and School Learning. (New York: McGraw Hill, 1976).

Bruner, Jerome. The Process of Education. (New York: Vintage Books, 1963).

Bruner, Jerome. "Models of the Learner", Educational Researcher. June/July 1985, pp. 5-8.

Carlberg, C. South Dakota Study Report. (Denver, Colorado: Mid Continent Regional Laboratory, 1981).

Eble, Kenneth E. Professors as Teachers. (San Francisco: Jossey-Bass Publishers, 1972).

Gaff, J.D. Toward Faculty Renewal. (San Francisco: Jossey-Bass Publishers, 1972).

Glenny, Lyman A. and Frank A. Schmidtlein. "The Role of the State in the Governance of Higher Education," Educational Evaluation and Policy Analysis. 5, No. 2, 1983, pp. 113-153.

Goslin, D.A. Teachers and Testing. (New York: Russel Sage Foundation, 1967).

Hopkins, K.D. and J.C. Stanley. Educational and Psychological Measurement. 6th Edition, (Englewood Cliffs, NJ: Prentice Hall, 1981).

Kennedy, John J. Analyzing Qualitative Data. (New York: Praeger Publishers, 1983).

Knapper. C.K. "Evaluating Instructional Development Programs," Paper presented at the International Conference on Improving University Teaching, London, England, July 1979, ED 177 968.

MacCuish, Donald A. "A Quality Assurance Model for Higher Education: A Pilot Study," Paper presented at the Testing and Quality Assurance in Higher Education Conference, Miami, Florida, February 1986.