DOCUMENT RESUME

ED 100 760 SO 008 045

AUTHOR Bonello, Frank J.; And Others
TITLE Alternative Approaches to Introductory Economics.
PUB DATE Dec 74
NOTE 40p.; Paper presented at the Joint Council for Economic Education, American Economic Association Meetings (December 1974)

EDRS PRICE MF-$0.75 HC-$1.85 PLUS POSTAGE
DESCRIPTORS Affective Objectives; Affective Tests; Cognitive Development; Cognitive Measurement; Course Content; *Course Evaluation; Course Objectives; Course Organization; *Economic Education; *Economics; Educational Alternatives; Educational Objectives; Educational Research; *Experimental Programs; Experimental Teaching; Higher Education
IDENTIFIERS Introductory Courses

ABSTRACT

This document examines the educational output of three alternative approaches to introductory macroeconomics at the University of Notre Dame. The framework for evaluation consists of the cognitive and affective tradeoffs entailed by using a new experimental course as opposed to two more traditional courses. The experimental course is a freshman level course which emphasizes laboratory and computer programing to teach hypothesis testing and economics methodology. The other two courses, lecture and discussion, are more traditional in scope, although one of the courses emphasizes current economic problems. Results of the scores on the Test of Understanding College Economics indicate that students in the experimental courses had higher scores than one of the traditional courses. In the skills of application of economic principles, the experimental course students score better than the two traditional courses, while scoring the same or less on recognition and understanding of economics principles. The affective goals are measured by course evaluations, change and final attitudes toward economics as a discipline, and changes in opinions toward economic issues. Results of the study indicate that the experimental course had the highest student rating in spite of the heaviest work load. Students found the experimental course more useful and relevant than the other two. The experimental course ranked second in its ability to change student attitudes about economic issues. (DE)

Alternative Approaches to Introductory Economics

by

Frank J. Bonello

William I. Davisson

Kenneth Jameson

Not to be quoted
without permission
of authors

Paper to be presented at the
Joint Council for Economic
Education session on economic
education at the American
Economic Association Meetings,
December, 1974

Introduction

Many economics departments offer several alternative introductory courses to the students of their colleges and universities. Presumably this is the result of the need to respond to a combination of institutional requirements as well as desired educational objectives. The rational department, of course, establishes a pattern of offerings which maximizes some objective function given institutional requirements and resource constraints.[1]

Such a maximization requires that the objectives of each alternative introductory course be carefully specified and measured. Besides providing information regarding the actual outcome of each course, this process allows a department to compare the alternative courses in terms of the tradeoffs between them and to determine whether these tradeoffs are acceptable. Unfortunately the methodology for such an evaluation is not well defined nor is the specification or measurement of objectives generally or easily available.

This paper addresses these problems by providing evidence on the educational output of three alternative approaches to introductory macroeconomics with a special focus on a new approach which has been developed at Notre Dame. It attempts to measure the broad educational effects, both cognitive and affective, of the courses. The framework for evaluation will be the tradeoffs entailed by using the new course as opposed to two courses of a more familiar nature. Hopefully our attempt in this difficult area will encourage further efforts

---

[1]The objective functions may differ in terms of what is maximized (or minimized). For example a department may choose from among the following alternatives: maximize the number of students who take introductory economics, maximize the economic knowledge of students who take the introductory courses, maximize interest in economics as a major, maximize the usefulness of introductory economics with respect to either the general educational background or vocational orientation of the student, and so on. Each one of these seems to imply a different pattern of offerings or at least a variation in the content of a given pattern.

leading to improvements in both the measurement of educational output and the methodology of evaluation.

The first section of the paper provides a brief discussion of the three courses, including a specification of their desired or expected outcomes as well as a simple profile of the students enrolled in each course. The next two sections present partial analysis of results and tradeoffs. The first of the two examines cognitive achievement of students as measured by the Test of Understanding College Economics (TUCE). The second section presents the available evidence concerning affective outputs, e.g. student attitudes toward the respective courses; towards economics as a course of study; and toward certain economic problems, institutions, and policies. The final section employs regression analysis to provide a "total analysis" of the output of the various courses.

Description of the Alternative Approaches

Each of the courses is an introductory macroeconomic course and thus they share certain common elements:

(i) Treatment of the nature of certain macro problems and institutions;

(ii) Treatment of certain tools of macro analysis; and

(iii) Treatment of the application of the tools to the problems.

Beyond these shared elements there are significant differences in the courses. Of major interest is Economics 101-102, the new approach which teaches introductory economics as a "lab" course.[2]

---

[2]A more extensive description of this sequence as originally offered can be found in Economic Education Experiences of Enterprising Teachers, vol. 11, (New York: Joint Council on Economic Education, 1974), pp. 95-101.

The course is offered to groups of approximately thirty-five self-selected freshmen and it fulfills a university social science requirement. The students are introduced to economics as a science and learn both economic theory and the technical skills of testing hypothesis. There are three meetings each week, two lectures and a "lab"--a two-hour session which is conducted in a lab equipped with computer timesharing facilities. The first semester lectures cover two topics. The first is philosophy of science and here the students are introduced to the methodology of modern science with emphasis upon the process of analysis. The hypothesis testing methodology of economics is introduced within this broader scientific framework. The second topic is the content of macroeconomics. A typical model allowing for monetary and fiscal influences is constructed and then expanded to include aggregate supply. Policy implications are also highlighted. In this discussion the theoretical antecedants of the postulated relationships and the empirical evidence bearing on their validity are explored as well.

In the lab the students initially learn the basic tools of empirical research. They first learn to program in Fortran IV and attain the skills necessary to construct and manipulate large data sets. Next they learn the basic statistical concepts which are used in economic research. The final behavioral objective of this portion of the lab is the programming and use of a simple linear regression routine. The concluding lab sessions allow students to select certain macroeconomic hypothesis drawn from the lecture and to test and evaluate these hypotheses for the U.S. economy, e.g. the determinants of consumption and of investment demand. These final sessions integrate lecture and lab, theory and methodology, by having the students perform simple research tasks.[3] The materials used in the course include two textbooks,

---

[3]The second semester of the sequence deals with microeconomics. Here the entire laboratory time is devoted to testing micro relationships about education, migration, and regional inequality. Thus the integration of lecture and lab is much more complete during the second semester.

one of which is on philosophy of science, and a lab manual.

The other two courses are of a more familiar variety. Economics 223 is a sophomore level course required of all Business College students.[4] Two of the three meetings per week are lectures with approximately 300 students, the other is a small discussion class of approximately 20 students directed by a teaching assistant. The same basic macro model is specified with some micro, government activity within a market economy, and poverty treated as additional topics. There is one text, a standard introductory book.[5]

The third alternative, Economics 121, is another freshman course which fulfills the university social science requirement. Its 75 self-selected students meet three times each week, two sessions are lectures and the third may be either lecture, discussion, or problem session. While the basic macro model with its policy dimensions are specified, the major concentration of the course is on current policy issues in macro, e.g. one-third of the course was spent on wage-price controls during Fall, 1973. In keeping with this general emphasis, three paperbacks were used, one of which was a short macroeconomics text.

---

[4]At Notre Dame the Economics Department is in the College of Arts and Letters.

[5]Although this sequence may be considered conventional, it too has been subject to change. In particular, during the time covered by our investigation, Fall 1973, under a grant from the Alfred P. Sloan Foundation an experiment involving a specially designed system of computer assisted instruction was conducted. In the experiment certain students were allowed to use the computer materials in place of the regularly scheduled discussion classes.

The above discussion highlights some of the differences between the alternative approaches but does not explain why these differences exist. The structure for Economics 223 results from its being the first of a two course sequence required of all College of Business students. As such it places a heavy emphasis on conventional macroeconomic content. The 121 course exists because it allows students of all colleges to use economics to satisfy a social science requirement. It utilizes a current macro problems approach in order to stimulate student interest in economics. The 101-102 sequence was created because we felt it was a "better way" to teach economics and at the same time better served the educational objectives of the freshman year. With respect to the former the 101-102 students would become familiar with economic content as well as the type of activities performed by professional economists. In terms of the latter the sequence equips students with knowledge and skills which are applicable over a range of academic areas.

In attempting to evaluate and thus compare the courses, our concentration is on the educational side. In particular, we are asking "if the economics department were to teach all introductory courses in the form of 101, what cognitive and affective tradeoffs would be involved?"[6] Given our specification of the respective educational objectives our expectations are that there would be a loss in the pure content area as measured by TUCE when compared with the 223 approach. We would expect TUCE performance to be comparable to the national norm and to exceed that of 121. Turning to the affective domain, the impact of 101 should be less than that of the more topical course 121.

---

[6]Note that once the educational side has been determined questions regarding institutional requirements and resource constraints would then be addressed. For instance the educational tradeoffs may be acceptable to us as a department but not to the College of Business or the hardware facilities necessary to process 600 students through 101-102 may preclude its adoption as the norm for the introductory course.

However we would expect that both 101 and 121 would be superior to 223 in this domain. The following sections will indicate that our expectations are generally correct and that we are able to specify the types of tradeoffs that would occur if 101 became the norm for introductory courses.

Before turning to the results of the analysis of the tradeoffs among the courses it will be useful to indicate the characteristics of the students in each of the courses since no attempt was made to assign students randomly.[7]

TABLE I

Selected Characteristics of Students

(standard deviations in parenthesis)

Characteristic	101	121	223
Verbal Scholastic Aptitude Test Score (VSAT)	539.44 (76.34)	526.35 (86.08)	567.51 (86.54)
Mathematical Scholastic Aptitude Test Score (MSAT)	602.78 (69.27)	590.00 (68.98)	620.09 (78.04)
Projected University Grade Point Average*	2.81 (0.668)	2.66 (0.668)	2.855 (0.567)
Percent Male	78%	85%	82%
Percent White	97%	96%	95%
Pre TUCE Score	12.39 (3.57)	12.37 (3.39)	13.04 (3.86)
Number of Students**	36	74	205

*The projected university grade point average is calculated using a formula which includes VSAT, MSAT, and rank in high school class. Weights are assigned to each variable with rank in high school class receiving the highest weight. The range here is from 0.00 to 4.00.

**Although there are more than 300 students in 223, we have selected only those students who have had no previous college economics. This was done so that the comparison is in effect between students who are experiencing their first college economics course.

[7]Account of the difference in student characteristics will be taken in the regression section by including these as independent variables. However, no such correction will be made in the following sections.

In examining the data presented in Table I it would appear that the students in three courses do differ in some major respects. The 223 students have higher MSAT and VSAT scores as well as a higher projected university GPA. The 101 students follow with the 121 students having the lowest MSAT, VSAT, and projected university GPA. In terms of the other two characteristics, percent male and percent white, the three classes seem quite similar. The students do not differ significantly (at the 5% level) on the TUCE pretest. Finally, as stated previously, all the 101 and 121 students are freshmen -- indeed first semester freshmen. The 223 students on the other hand are exclusively non-freshmen with about 80 percent of them being sophomores.

Cognitive Domain

As stated we would expect that a course such as 101 will yield less cognitive attainment as ordinarily measured by researchers on economic education than a course like 223. The basis for this expectation is straightforward; the time spent on economic content in lecture is less and the lab period develops skills which do not translate into content knowledge. Thus the question to be answered in this section is: what is the extent of the tradeoff?

Before attempting to answer this question, it is important to state explicitly why we are willing to accept less cognitive attainment in 101 (as measured by TUCE). There are two reasons. First, other knowledge and skills

are developed as a substitute--philosophy of science, computer programming, use of statistics and hypothesis testing--and we feel that these are valuable in the context of modern economics.[8] Second, we feel that the reduction in cognitive attainment will not occur evenly; that is, 101 emphasizes the "process of economic analysis" and thus these students will do relatively better on problems which require analysis and thought.

In measuring the cognitive dimension of the courses we followed standard procedure and utilized the TUCE Part I (macroeconomics), Form A, as both a pre and a posttest. Several different types of information are available from the TUCE and are presented below.

The first of these is student performance on the entire 33 question exam. (See Table II).

TABLE II

Test of Difference of Means (t-test) on TUCE Posttest*

Mean of Economics 101 20.89 (15.78)**

Comparison Course	Mean	t-value	Significant at .05
National Norm	19.22	1.70	yes
Economics 121	18.31 (14.35)**	2.27	yes
Economics 223	23.35 (17.91)**	-2.27	yes

*There was no significant difference in performance on the pre-test across the courses (see Table I).

**The values in parenthesis are the means on 25 questions which the instructors felt were covered in all three courses although the coverage was not the same in breadth or depth. The results of the comparison of means is the same for this common core as for the overall TUCE. Thus the discussion will concentrate on the whole TUCE.

---

[8]The success of 101 in these areas is measured by the students' ability to attain behavioral objectives set prior to the course. In general the students do succeed in attaining the goals.

These results correspond to our expectations. The 223 course is most effective in teaching content as measured by TUCE. The 101 course, although yielding significantly less cognitive attainment than 223, does produce significantly better results than either the national norm or 121. Given that maximization of TUCE type performance is not the sole goal of 101, we find that the tradeoff between broader types of knowledge attained in 101 and TUCE is not severe and the students actually perform quite well. In our view this tradeoff is more than acceptable.

Another type of insight into cognitive attainment is available from a disaggregation of the TUCE posttest scores into "objective categories." These include recognition and understanding, simple application, and complex application. This categorization was suggested by the authors of TUCE and results in three 11-question divisions.[9] As noted above we would expect the 101 students to do relatively better on questions which require more complex

---

[9]It is difficult to say whether the names given the categories represent the types of knowledge tested. The TUCE manual defines the categories as follows: "Identification or 'recognition and understanding' refers to questions which could, in principle, be answered by reference to a text book."
"A simple application of a principle or concept may be defined as one in which the student demonstrates that he can utilize that principle or concept when its use is specified or clearly implied."
"A complex application requires that the student demonstrate ability to select and utilize a concept or principle when its use is not specified." See Manual Text of Understanding College Economics (New York: The Psychological Corporation, 1968). It has been argued however that the categories represent different degrees of hardness as evidenced by student scores. See Darrell Lewis and Tor Dahl "The Text of Understanding College Economics and Its Construct Validity," Journal of Economic Education, Spring 1971, 2, pp. 77-85.

thought and reasoning patterns, i.e. on complex application as opposed to recognition questions. Table III presents the results of this disaggregation of TUCE performance.

TABLE III

Test of Difference of Means on Objective Categories

TUCE Posttest

	Mean	t-value	Significant at .05
Recognition and Understanding			
Econ 101 = 7.39			
Comparison Course			
National Norm**	7.41	-	
121	7.64	-0.34	No
223	8.16	-1.12	No
Simple Application			
Econ 101 = 6.97			
Comparison Course			
National Norm**	6.17	+	
121	5.49	+2.06	Yes
223	8.21	-1.89	Yes
Complex Application			
Econ 101 = 6.53			
Comparison Course			
National Norm**	5.03	+	
121	5.17	+2.11	Yes
223	6.98	-0.62	No

*There was no significant difference in performance on the pre-test across the course in any of the objective categories. The means on the objective categories were as follows: recognition and understanding 5.00 (101), 5.11 (121), 5.14 (223); simple application 4.16 (101), 3.99 (121), 4.27 (223), and complex application 3.22 (101), 3.27 (121), 3.63 (223).

**Note: No standard deviations are available for the national norming data and thus tests cannot be made.

In examining these results, we find some confirmation of our hypothesis that the loss in TUCE performance for the 101 students would decline as we moved to more complex skills. Although the 101 students attained a significantly lower total TUCE score than the students in 223 there was no significant difference on the complex application questions--a result in accord with our expectations. However, the major difference between the 101 and 223 students is on the simple application questions which is not the expected result. The comparison of 101 and 121 shows exactly the pattern expected, 101 returns a lower (insignificant) score on recognition and understanding questions and in the other two categories yields higher (significant) results. The same pattern seems to be indicated when 101 is compared to the national norm. There 101 is very slightly below the national norm on recognition and understanding and then rises to 12% and then to 30% superiority as we move to simple and complex application questions. These results when added to the overall results in effect reduce the negative impact of the TUCE tradeoff, the loss in TUCE performance occurs in "skill areas" that are not stressed within the 101 course.[10]

The final type of insight into cognitive attainment as measured by the TUCE involves a disaggregation by "content categories." The TUCE developers have established seven content categories. An analysis of the courses on this basis is important for the courses differ in topical coverage and the emphasis given to covered topics. It is also important because the objective categories of the TUCE also vary with content.

Given the large number of categories and the difference in the number of items per category, it will be easier to deal with these results in terms of the

---

[10]One might add that the complex application area is the most important not only because it requires the more advanced skills, but also because the retention of this type of knowledge appears to be greatest. These results were reported by Philip Sanders in a paper given at the 1974 Midwest Economics Association Meetings.

percent of questions in each content category answered correctly on the post-test. Our expectations were that the 223 students would do better in most or all categories while the 101, while not performing as well as the 223 students should perform better than the 121 students in most or all categories. The one possible exception was the content category labeled determinants of economic growth. Here 121 might be superior because this is the only course which spends any time on growth per se. Simply performance across content categories should be responsive to the breadth and depth of coverage given these areas in any course.

These results, (Table IV) in general also confirm our expectations.

TABLE IV

Content Category Results of the Tuce Posttest

Percent Correct*

Category	101	121	223
A. Scarcity, functioning of economic systems, base elements of supply and demand	70.8%	65.7%	76.2%
B. Macroeconomic Accounting	86.0	69.8	87.0
C. Determination of GNP (income-expenditure theory)	46.5	49.3	59.5
D. Money, banking, and monetary policy	64.8	50.3	70.9
E. Government fiscal policies	60.0	47.0	68.1
F. Determinants of economic growth	45.8	70.9	63.4
G. Policies for stabilization and growth	56.4	52.2	61.6

*On the pre test there were only two differences which were significant: on Category D 223 was significantly above 101 ($t = 2.77$) and Category E 101 was significantly above 121 ($t = 1.91$). In terms of the above posttest, 223 is significantly better than 121 on categories A through E and significantly better than 101 on topics C and F. In turn 121 is significantly better than 101 on category F with no other significant differences appearing.

The 223 students perform best in six of the seven content categories and are second best in the remaining category. The 101 students are second in five of the seven categories while 121 students return the highest percentage in the growth category, second in another, and on the other five are in third place. Again 101 performs in an acceptable manner yielding its poorest performance in the growth area--an area not specifically covered in the course. It was also lowest in the determination of GNP, a result which was not expected but whose negative impact is reduced because of the absence of complex application questions.[11]

There then are the cognitive results. As we have indicated 101 does not maximize TUCE performance. But the loss in performance whether measured in terms of total loss, loss by objective category, or by content category is more than compensated for by gains in areas which are not captured by the TUCE exam. It should also be noted that these losses are only relative to a course (223) which has as its objective the maximization of TUCE like performance.[12] Perhaps even more important is the fact that performance of students can be influenced by the structure and coverage of materials.

Affective Domain

There is clearly another output of the educational process which must rate on a par with the purely cognitive effects. This is the impact of our educational endeavors on the attitudes and beliefs of the students. The major problem in dealing with this important area is the absence of standarized

---

[11]Further analysis by content category is presented in Appendix I.

[12]It should be pointed out that micro semesters of 102 and 224 return TUCE results which roughly correspond to the pattern noted in the macro courses. In addition, we found that the gap between the courses, the tradeoff, is less (102 Posttest of 23.19 compared with a 224 score of 23.58 with t = .8). This result was to be expected since there is much more direct concentration on economic content in both the lecture and the lab of 102, for philosophy of science and computer programming are not taught directly.

instruments or techniques for assessing these impacts of teaching. Nonetheless the attempt must be made and this section of the paper examines three possible measures for use in the affective area: University-wide course evaluations, change and final attitudes toward economics as a discipline, and changes in opinions toward economic issues.[13]

Our expectations are that there will be a tradeoff in this dimension among the three courses. In particular we would expect the 121 course to have more of an impact in the affective area than either 101 or 223. This is because of the nature of the course which takes a much more topical approach to macroeconomics and relates the material directly to current economic questions. The 223 course should yield the worst results because it is a large lecture class and is taken as a required course by a majority of students.

One commonly used measure in the affective domain is the course evaluation form. On the particular form used at Notre Dame students are asked to rate the instructor and the course on a variety of items using a scale from 1 to 4 with 4 being the highest possible rating. Within the Economics Department the

---

[13]There is some question whether all three of these elements should be placed in the affective domain. This is especially true of the changes in opinions toward economic issues which some might label a cognitive element rather than an affective element. See: Benjamin S. Bloom, ed., Taxonomy of Educational Objective Handbook I (Ann Arbor: McKay Company, 1956) pp. 185-189. Opinion hence decision changes on the part of the student regarding specific economic issues reflects the attainment of an objective in the cognitive domain in a different sense than is measured on the TUCE examination. For example, a student may be able to replicate the economists' analysis of comparative advantage (e.g., TUCE, complex application or some other application question) as that relates to foreign trade. However, if at the end of the course the student still feels that free trade is an economic evil have we achieved our objective in the cognitive domain? Finally, the affective area might be expanded to includes measures such as the number of subsequent economics courses taken by the students and the number of students who become majors. In the current case such measures are not possible or useful: at Notre Dame students do not formally declare majors until their junior year and thus for the base year we would formally have no majors generated by 101 and 121 and Business students who make up the majority of the 223 course are much more restricted in course choice than Arts and Letters students.

scores on the items seldom fall below 2.5 with a typical range between 3.0 and 3.5. The evaluation is taken towards the end of the semester, normally within the last two weeks.

The course evaluations for the three alternative approaches are reproduced in Table V. In summarizing these results we can begin by making three observations: (i) in general, 101 receives the highest rating even though it engendered the heaviest work load; (ii) 121 was rated close to 101 with major differences appearing on three items--teacher's ability to encourage independent thinking, quality of text, and work load and (iii) 223 receives the lowest overall rating, and, even though this course yields superior TUCE results, it was rated lowest in terms of work load.

Examination of the patterns on the specific items or of the overall average indicates partial confirmation of the expectations of the tradeoff in the affective domain. As expected both 101 and 121 yield ratings which are superior to the 223 course. Contrary to expectation the 101 course is more successful than 121; on most of the items and, thus, on the overall average 101 receives a higher rating.

A second measure within the affective domain is student attitudes regarding economics as a discipline. Specifically students were given five descriptors of economics and were asked to rank their attitudes on a scale from 1 to 5. For example, one descriptor was "hard" with a 1 implying that economics was not very hard while a 5 means that it was very hard. The data were collected on a pre and post basis. Table VI presents these results for the three courses. Here the 101 and 223 results are quite similar, for all five descriptors the post survey yielded higher values than the pre survey. Thus these students found economics harder, more abstract, more useful, more theoretical, and more relevant than they had originally thought. The 121 students moved up

TABLE V

University Course Evaluation

(4.0 Point Scale)

Category	101	121	223
1. Knowledge of subject matter.	3.93	3.91	3.94
2. Enthusiasm for subject.	3.96	3.94	3.86
3. Effectiveness of class teaching or direction.	3.64	3.56	3.44
4. Ability to stimulate student interest in the subject.	3.58	3.43	3.14
5. Interest in students--e.g., availability, helpfulness, etc.	3.61	3.41	2.89
6. Fairness in dealing with students.	3.67	3.58	3.31
7. Respect for student's viewpoint.	3.70	3.55	3.23
8. Organization and management of course.	3.58	3.65	3.59
9. What overall rating would you give this teacher as compared with other teachers you have at Notre Dame?	3.83	3.78	3.57
10. Teacher's ability to lecture in a clear, interesting manner.	3.80	3.70	3.71
11. Teacher's ability to encourage independent thinking.	3.60	3.25	3.05
12. Teacher's carefulness of preparation for class.	3.75	3.67	3.74
13. Teacher's promptness in returning student's work.	3.73	3.70	3.88
14. Intellectual level of the lecture.	3.67	3.63	3.43
15. Quality of text.	3.48	3.18	3.22
16. Personal value of the course to you.	3.60	3.52	3.29
17. Standards for students performance required for this course.	3.46	3.31	3.27
18. Compared with other courses you have taken or are taking at Notre Dame, the work load for this course was:	3.45	3.13	2.79
Average	3.67	3.55	3.41

the scale on these descriptors but they found economics less abstract and less theoretical than they had anticipated.

TABLE VI

Attitudes Towards Economics as a Discipline

(Scale of 1 to 5)

Descriptor	101		121		223	
	Pre	Post	Pre	Post	Pre	Post
Hard	2.99	3.61	3.16	3.86	3.19	3.58
Abstract	2.69	2.80	2.71	2.63	2.81	3.00
Useful	4.38	4.54	4.48	4.55	4.15	4.34
Theoretical	3.72	4.06	3.64	3.55	3.57	3.62
Relevant	4.36	4.70	4.45	4.55	4.20	4.42

In interpretating these results we felt that economics courses, particularly introductory courses, should be perceived by students as useful and relevant. Indeed the 121 course is designed to stimulate interest in economics and thus presumably should yield the highest post values on the useful and relevant descriptors.[14] Relative to the 223 course this result was obtained but contrary to our expectations the 101 course matches 121 on the useful descriptor (a difference of only .01) and surpasses 121 on the relevant descriptor (a difference of .15). Simply the tradeoff expected between 101 and 121 in this area of the affective domain is not realized; indeed, this measure as well as the previous measure imply that 101 is the most successful of the three courses.

---

[14]It is interesting to note that a course can be made more useful and more relevant without at the same time making it less abstract and less theoretical at least as evidenced by 101 and 223. This is a tradeoff which many people seem to feel exists in any given course. Our feeling is that usefulness and relevancy can be maintained or increased even with heavy theoretical emphasis if examples of a realistic nature are continuously used in the process of theoretical development or if an involved theoretical development is subsequently applied to to a problem.

The third and final measure of the affective domain is student opinion regarding certain economic institutions, problems, and policies. Again surveys were administered on a pre and post basis. On these surveys an item--say inflation--is presented and under it are listed a series of semantic differentials such as "bad-good," "inevitable-controllable," "demand pull-cost push," etc. Each semantic differential represents then a continuum ranging from 1 to 5 with a 3 representing a middle of the road position or no opinion. Eight macro items were given with a total of forty-one semantic differentials.

The interpretation of student attitudes from such instruments is far from clear.[15] It can be agreed however that change in student opinion is a measure of educational impact; the educational process operates on a student opinion which is either poorly formulated or which is not based on knowledge and evidence. Consequently the educational process brings forth a change in that opinion. This implies that a possible measure of the success of a course is the size of the opinion change.

Table VII presents these results. To summarize these data we can establish ranges of opinion change: (i) no effect--average opinion change of less than .100, (ii) intermediate effect--average opinion change of between .100 and .500 and (iii) significant effect--average opinion change greater than .500. Using these ranges 223 generates the largest number of significant changes--eight--followed by 101 and 121 each with five significant changes.

---

[15] As stated previously, one can argue that the type of effect we are attempting to capture is more appropriately cognitive than affective. We have some sympathy for this argument but include our discussion within the affective section because the affective area itself, as is this particular dimension, fairly undefined.

TABLE VII

Opinions of Economics Institutions*
Problems and Policies
(Scale 1 to 5)

Issue	101 Ending Position **	101 Change ***	121 Ending Position	121 Change	223 Ending Position	223 Chang
Decision-Making for National Economics Policy						
Simple-Complex	4.581	+.275	4.452	+.074	4.585	+.119
Random-Rational	4.032	-.032	3.932	+.040	4.126	+.329
Political-Apolitical	2.400	-.145	1.986	-.109	2.189	+.134
Important to me-Irrelevant to me	2.097	-.264	.905	+.243	1.881	-.041
Important to society-Irrelevant to society	1.355	-.589	1.432	+.094	1.333	-.217
Government Spending Deficit						
Undesirable-Desirable	3.097	+.375	2.527	-.068	3.433	+.802
Democratic-Republican	3.065	+.176	2.986	+.027	2.936	-.034
Promotes Prosperity-Promotes Depression	2.414	-.142	2.342	-.239	2.234	-.48
Increases Aggregate Demand-Decreases Aggregate Demand	2.290	-.571	2.055	-.769	2.241	-.50
Government Controls such as Wage and Price Freezes						
Undesirable-Desirable	2.871	-.407	3.027	-.041	2.975	-.08
Democratic-Republican	3.433	+.044	3.351	-.128	3.509	+.15
Free Enterprise-Socialism	3.677	+.306	3.730	+.525	3.509	+.09
Effective-Ineffective	3.914	+1.081	3.027	-.189	3.224	-.02
Facilitate Market Operations-Distort Market Operations	3.677	+.177	3.658	+.266	3.696	+.22
Poverty						
Laziness-Lack of Opportunity	3.355	+.469	3.644	+.320	3.873	+.323
Serious-Unimportant	1.163	-.054	1.419	-.054	1.503	+.003
Market Determined-Discrimination	2.677	-.209	2.890	+.196	2.615	-.103
No Jobs-No Skills	3.194	-.149	3.423	+.342	2.949	-.239
Inevitable-Can be Eliminated	3.000	+.056	2.905	-.234	3.283	+.159
Individual Responsibility-Social Reponsibility	3.571	+.257	3.676	+.224	3.870	+.357
Inflation						
Bad-Good	2.633	+.462	1.932	-.257	2.704	+.508
Recession-Prosperity	3.500	+.250	2.973	-.203	3.660	+.635
Inevitable-Controllable	3.452	-.005	3.365	+.014	3.403	-.176
Democratic-Republican	2.900	-.267	3.203	+.054	3.019	-.031
Demand Pull-Cost Push	3.267	-.019	3.563	+.320	3.342	+.179

TABLE VII
(Continued)

Issue	101 Ending Position *	101 Change **	121 Ending Position	121 Change	223 Ending Position	223 Chang
High Unemployment						
Democratic-Republican	3.233	+.073	3.315	+.099	3.208	-.018
Recession-Prosperity	2.194	+.033	2.324	+.256	2.120	-.042
Inflation-Deflation	3.167	+.024	3.189	+.459	3.465	+.449
Controllable-Fact of Life	2.467	+.238	2.541	+.365	2.132	-.098
Bad-Good	1.935	+.324	1.589	+.008	1.767	+.159
Federal Income Tax						
Progressive-Regressive	2.097	-.292	2.486	+.222	1.742	-.833
Good-Bad	2.097	-.292	2.351	+.036	2.006	-.556
Democratic-Republican	2.839	-.078	3.027	+.137	2.881	-.045
Free Enterprise-Socialism	2.710	+.127	2.662	+.216	2.648	-.017
Too high-Too low	2.806	+.112	2.541	+.089	2.722	+.356
Fair-Unfair	2.710	-.290	3.014	-.081	2.673	-.644
Increasing the Money Supply						
Helps People-Hurts People	2.484	-.766	2.669	-.861	2.296	-.838
Inflation-Recession	2.033	-.189	2.028	+.731	2.184	+.106
More Gold-More Paper	3.677	+.594	3.919	+.824	3.829	+.464
Democratic-Republican	3.033	+.144	2.959	-.069	3.006	+.031
Free Enterprise-Spcialism	2.710	-.033	2.608	-.172	2.491	-.297

* One often finds that the term "opinion" is used to refer to a type of "guess" that a person makes when asked a question, where that guess is based on little or no information. After being subjected to new information consistently presented (e.g., the process of education) the person may be said to have changed his "opinion" or to have made a decision or to "have taken a specific position" or to have "defined his own position" on that particular question or issue. In a sense a student might have "no information" on a particular issue at the beginning of a course and any response he would give in an opinion survey would be a guess or an "opinion". By the end of the semester the student may well have gained enough information (consistently presented) to have made a decision on the specific issue or question and he has therefore taken a position on that issue even if it is a "middle of the road" position. This is the basic problem of interpretation of opinion surveys that we have noted. By and large we should probably use the word "opinion" for the survey results at the beginning of the semester and "evaluation or decision" for the results at the end of the semester.

** Ending position is defined as $\frac{\sum_{i=1}^{N} Oi}{N}$ where Oi represents the opinions of the ith student and N the number of students. Change is defined as the difference between ending and beginning positions.

*** A plus sign indicates that ending position represented a higher score than initial position.

In only two instances did the significant changes occur for the same semantic differential in all three courses. For the intermediate range there were sixteen such changes for 101, and fourteen for both 121 and 223. Eight of these occurred simultaneously across all three courses and four are concentrated within the issue of poverty.

On this basis then it appears that 101 and 223 students have a slightly higher propensity to change opinions than 121 students and in the case of 223 the changes appear to be somewhat stronger.[16] With the exception of the poverty issue these changes do not occur consistently, each course seems to have its own impact. These results are contrary to our expectations and indicate that if there is a tradeoff it is not between 101 and 121 (101 performs better) but between 101 and 223.

Another possible measure is to examine the "no opinion" students, i.e., those who were noncommittal or 3 initially and remained as such on the post survey. This case is viewed as the extreme of "no impact"; a course did not provide any information--as perceived by the student--upon which to form an opinion.[17] To summarize the data we can compare the number of semantic differentials (out of a total of 41) for each course for which less than 50% of the initial 3's remained 3's. We find that this is the case for 20 semantic differentials in 101, 23 in 121, and 31 in 223. Thus 223 is most successful. This analysis is for the most part inconsistent with our expectations: inconsistent in that 223 is more successful than 101 and 121, but consistent in that 121 is more successful than 101.

---

[16]Note that the ordering of effects here is similar to the ordering on the cognitive results. This suggests that indeed this measure is part and parcel of cognitive achievement.

[17]It is necessary to admit that no impact for a continuing "3" position may be a misnomer. What may have in fact occurred is that the student established a solid foundation for a middle of the road position, changing in effect from no opinion to a middle of the road position.

While there is no agreed upon method of assessing the impact of a course in the affective domain and little hope of isolating course specific effects, the combined evidence indicates that the tradeoff expected between 101 and 121 did not occur. To use the positive approach 101 performed much better than expected relative to 121.

Regression Analysis of Cognitive Achievement

Previously it was argued that the expected tradeoff of less cognitive achievement as measured by the total TUCE for 101 relative to 223 was realized. That conclusion, however, is suspect for as Table I indicates the 101 students had lower MSAT and VSAT scores than their 223 counterparts. The counter argument then is that the tradeoff occurred not because of a difference between courses but because of differences between students in the courses. In order to control for this possibility and other differences we can employ multiple regression analysis of TUCE performance.

Before reporting the results of the regression analysis, it is useful to articulate our view or theory of the educational process as it pertains to cognitive achievement. It appears, at least to us, that three types of factors are operative in producing educational output although all center on or are processed through the student. The first may be grouped under the heading of human capital and measure the ability of the student. As a student enters an introductory course he has certain abilities. These abilities at this point in time are a product of his prior experience in the educational process as well as such things as home environment, peer group influences, and innate ability. In measuring the stock of human capital we are not however concerned

about sources, we merely seek an appropriate quantification of human capital as it bears on the students capacity to achieve cognitively in economics.

But just because a student brings a certain amount of human capital to an introductory course does not mean that he or she will use it. This obvious statement raises the need for our second category of factors which we will label utilization rates. Specifically we would expect that a student with a given amount of human capital would learn more the more frequently he attends class, the longer he studies the subject matter and so on.

These utilization rates unfortunately capture only the quantity dimension of the students application of his human capital. The need to isolate or incorporate the quality dimension of time expanded on a course introduces the third type of factor. For lack of a better name, we will call these elements efficiency variables. The argument here is straightforward--a student will get more out of a given hour of lecture if he thinks the lecture was "good"; he will get more out of a given hour reading the text if he thinks the text is "good."

We may then specify the functional form of our view as:

$$(1) \quad CA = f(H, UR, E)$$

where CA = cognitive achievement
H = human capital
UR = Utilization rate and
E = efficiency or opinions regarding elements employed within the educational process.

We also postulate that the relationship between each of the independent variables and the dependent variable is positive.

Even in this general form, one might argue that our view of the educational process is a simultaneous equation system with the values of the independent variables in fact dependent upon each other. It seems to us that this indeed may be the case. Students with lower human capital may compensate by increasing

utilization rates. Different opinions regarding elements employed within the educational process may be a function of the human capital of the user. Higher opinions of lecture or text may increase the frequency of lecture attendance or the number of hours spent with the text. Perhaps an even more fundamental question concerns the inclusion of outputs to explain outputs. The separate treatments of the affective and cognitive domains implies that the two areas are distinct and influenced by alternative approaches. The above specification on the other hand implies that a favorable outcome in the affective domain will help to achieve a favorable outcome in the cognitive domain; that is, favorable opinions about economics, the instructional elements, and shifts in opinions about economic problems and institutions are not only desirable in and of themselves but because they facilitate the attainment of cognitive objectives (and vice versa). Carrying this logic to the extreme would imply absolutely no tradeoff between cognitive and affective outcomes; to achieve one is to achieve the other. We of course do not accept this extreme argument but we do feel that affective outcomes may operate positively, up to a point, on cognitive output.

Having briefly sketched our view of the learning process we must now operationalize the model. In terms of human capital variables three are employed in all regressions. They are MSAT, VSAT, and rank in high school class (RHSC). The latter was included on the presumption that human capital may not be fully captured by MSAT and VSAT. Of course the three variables are not perfect measures of the stock of human capital at the time the students began the respective courses. In the case of the 223 students who are mainly sophomores these variables measure human capital as it existed some two years previously.

Finally, the major purpose of the regression analysis is to determine whether the differential cognitive results reported earlier could be attributed to the courses themselves. Thus we have included two dummy variables. The first is DV223 which was coded as a 1 for all 223 students and 0 for all others. The second is DV121 which was coded as a 1 for all 121 students and zero for all others. In this way we can compare the 101 students to each of the other groups by examining the sign and significance of the two dummy variables.

The regression version of the model thus appears as

$$CA = a + b_1(MSAT) + b_2(VSAT) + b_3(RHSC) + b_4(CE20) + b_5(CE9) + b_6(CE15) + b_7(CE16) + b_8(OPI) + b_9(DV223) + b_{10}(DV121)$$

All of the variables--excluding DV223 and DV121--are expected to operate positively on cognitive achievement. The elimination of pre TUCE scores as independent variables was justified on the basis of using the gap closing specification of the dependent variable.[20] Regressions were run on the Total TUCE as well as the three objective categories and the results are presented in Table VIII.[21]

These results offer insights in many areas but the two of concern here are the explanatory ability of the model and the impact of the courses themselves. With respect to the former we can begin by examining the R^2 values. We find that the model works fairly well in explaining total TUCE performance ($R^2 = 0.395$) but then unevenly among the objective categories. Here the model performed best in the simple application category ($R^2 = 0.425$) and worst in the recognition and understanding category ($R^2 = 0.101$).

---

[20]See Frank Gery "Is there a Ceiling Effect to the Test of Understanding College Economics?" Research Papers in Economic Education (New York: Joint Council on Economic Education, 1972) pp. 35-49.

[21]Additional regression analysis is contained in Appendix II.

TABLE VIII
Regression Results - Total TUCE
and the Three Objective Categories

Dependent Variable		Constant	MCAT	VSAT	HSEC	UR	CS9	CS15	CS16	OPI	DV223	DV121	R^2	N	F Statistic
Total TUCE	b value	-0.07497	0.00867	0.00010	0.00070	0.01474	-0.06314	0.00290	0.03313	0.01058	0.13190	-0.10449	0.39538	226	14.12506
	beta weights		0.32775	0.04109	0.03757	0.04858	-0.15708	0.01119	0.11655	0.08949	0.30552	-0.22477			
	t-ratio		3.762*	0.677	0.646	0.843	2.568*	.178	1.722	1.637	3.496*	2.642*			
Perception and Understanding	b value	-0.12403	0.00075	-0.00040	0.00233	0.02562	-0.06576	0.00651	0.05645	0.00225	0.13795	0.02775	0.10131	226	2.43511
	beta weights		0.13879	-0.11804	0.09241	0.06200	-0.12006	0.01840	0.14592	0.01397	0.23456	0.04382			
	t-ratio		2.533*	1.602	1.303	0.883	1.610	0.240	1.767	0.209	2.202*	0.421			
Simple Application	b value	-0.17902	0.00079	0.00060	-0.00153	-0.00313	-0.04617	-0.00022	0.01805	0.00703	0.10277	-0.17036	0.42528	226	15.96347
	beta weights		0.14031	0.16223	-0.05493	-0.00638	-0.07663	-0.00057	0.04241	0.03967	0.31343	-0.24455			
	t-ratio		3.036*	2.744*	0.969	0.122	1.285	0.000	0.682	0.794	3.679*	2.948*			
Complex Application	b value	-0.56328	0.00946	0.00011	0.00177	0.04120	-0.06625	-0.00001	0.03372	0.02398	0.06927	-0.14165	0.19388	226	5.19506
	beta weights		0.11317	0.03089	0.06855	0.09745	-0.15395	-0.00003	0.08920	0.14556	0.11513	-0.21864			
	t-ratio		1.603	0.641	1.021	1.465	2.180*	0.000	1.289	2.306*	1.141	2.225*			

* Indicates significance at the .05 level

Another aspect of the model's efficacy is the sign and significance of the coefficients. Of the three human capital variables we find that MSAT is the most consistent--significant and positive in three of the four regressions. VSAT has one incorrect sign but is insignificant here and in the one instance it is significant, it has the proper sign. RHSC is consistently positive but insignificant.

UR represented the single utilization rate variable. It's performance is disappointing--insignificant in all four regression with one incorrect sign. As stated previously the wording of the question on the course evaluation form may explain the poor performance of this variable.

With the four efficiency variables, the results are also disappointing. The sign of the teacher evaluation is consistently negative and in two instances it is significant.[22] The text evaluation has two unexpected signs but in all four regressions is insignificant. The course evaluation is consistently positive but insignificant. OPI operates positively in all four regressions and is significant in the complex application regression. Clearly these results imply that the model needs further work and better devices to measure the various aspects of the educational process which it specifies.

Turning to the question of the effects of the course, the sign and significance of the dummy variables support our earlier assertions. In the total TUCE regression we find that DV223 is positive while DV121 is negative with both variables significant. The 223 students after controlling for the other independent variables obtain higher total TUCE scores than the 101 students but the 101 students are significantly better than the 121 students. On recognition and understanding questions 223 is again significantly better than

---

[22]For a fairly complete discussion of teacher evaluations see the Fall 1973 issue of the Journal of Economic Education.

101 but there is no difference between 101 and 121. With simple applications 223 is significantly better and 121 significantly worse than 101. We had hoped that the 223 - 101 comparison would not be significant in this category. On the complex application we do obtain our expected results--no significant difference between 223 and 101 and 101 is significantly better than 121. Simply the results of the regression analysis do confirm the earlier reported results, again disconfirming only one of our a priori expectations.

Conclusions

In the teaching of any course the instructor must identify his objectives, develop a strategy for obtaining these objectives, and determine the extent to which the objectives have been obtained. The same is true of a department which offers alternative introductory courses. The identification of objectives treated in this paper include both cognitive and affective outcomes for three different courses as well as the expected tradeoffs. The strategies for obtaining these different outcomes was developed by examining the structure of each course. The TUCE was used to determine the cognitive outcomes while three different measures were used to assess affective outcomes. In general, the results indicated that our objectives in each course were being attained; courses indeed can be designed to achieve different outcomes.

We hope that this paper will stimulate others to undertake evaluations of their own courses. In these activities, as indicated by our own difficulties, special attention should be given the TUCE. In particular the dual disaggregation of the TUCE provides extensive analytical capabilities. Simply if one is teaching an analytical course he should pay particular attention to the objective category results but also be careful that the content coverage of his course

parallels that of the TUCE. It also appears that the entire affective domain is an area in which much research needs to be done. Finally, it also appears that individuals must be careful in specifying what they believe to be the educational process taking place in their courses before proceeding to empirical analysis. Hopefully this paper has provided some insights in these latter areas which others may build upon.

Appendix I

Performance by Content Categories

As was noted in the body of the paper 223 students perform best in six of the seven content categories and are second in the remaining category. The 121 students are best in one category, second in another and on the other five are in third place. The primary factor which seems to affect these rankings is the concentration which an individual course places on a given topic. For example, little time was spent on growth in 101 or 223 while 121 does spend some time on this topic. Thus the superior performance of 121 students in this content category can be explained by concentration alone. Also most of the conventional macroeconomics taught in 121 is presented within the context of GNP or income determination and thus the better performance of these students with respect to 101 students here is also understandable. In both of these content categories--determination of GNP and determinants of economic growth--there are no complex application questions. This is one dimension of why 121 students do better in these areas but exhibit relatively poor performance on complex application questions. In the same vein it is interesting to note that more than half of the 121 correct complex applications occurred in the area of policies for stabilization and growth. Thus there is a rather clear linkage between the objective and content categories of TUCE.

The data presented in Table III also seem to reflect a ceiling effect. For example, the maximum average attainment on any content category was about 86% (macroeconomic accounting) for the 101 and 223 students. This implies that in any given content area, it may be difficult to get an average attainment greater than 86%. This specific result may be biased upwards, however, since there are only three questions in this content category and it also has the

highest percentage of recognition and understanding questions--66% as compared to 50% for the next highest category. If we look past this for the moment, the 223 students attained 76% on the scarcity category of six questions but this was the only other score which was above 71%. Indeed twelve of the remaining scores are between 55% and 71% while the other six are all above 45%. Thus we may be dealing with a rather narrow range of final achievement for the topical areas stressed in any given course.

In light of this difficulty, the data may be analyzed in an alternative way which, hopefully, deals more directly with the goals of teaching; that is, the percent improvement in the content categories. This is defined as a ratio of the difference between posttest and pretest scores to pretest score and the results are presented in Table A.I. Here we find much more variation: in three cases we find the scores on the posttest more than double those of the pretest and with the exception of one case in which there was absolutely no improvement, the minimum gain is 30%. These bounds probably indicate the general range of effects for a given course. The results again indicate that the superior performance of the 223 students was the result of an across-the-board improvement. In the three cases in which the 223 students are not superior only one appears to be a major difference--monetary--in which the 101 students practically doubled the improvement of the 223 students. Again only in the growth category do 121 students appear as superior although they also continue to perform better than the 101 in the GNP determination category.

It has also been suggested both because of the limited number of questions in certain content categories and of the possible overlap of concepts between categories that the questions be aggregated into four categories. The categories are listed in Table A.II as well as the results first for percent correct

TABLE A.I.

Percentage Improvement in TUCE by Content Categories

Category	101	121	223
Scarcity, functioning of economic systems, base element of supply and demand	41.7	30.4	47.1
Macroeconomic accounting	45.3	32.5	71.5
Determination of GNP (income expenditure theory)	86.1	128.1	151.4
Money, banking, and monetary theory	126.4	44.5	66.7
Government fiscal policies	38.4	34.9	79.0
Determinants of economic growth	0	45.8	37.5
Policies for stabilization and growth	76.8	45.9	65.5

on the TUCE posttest and percentage improvement. With respect to percent correct we find 223 superior in all four categories but with little variation in these percentages between categories and courses. The lowest is 49% correct while the highest is 76%. In terms of percentage improvement 223 is superior in three of the four categories with more variation. The range here is from 33% improvement to 105% improvement. These results are not particularly different from those using seven content categories and thus because of the greater detail provided by the seven categories it is to be preferred.

Table A. II.

Results Using Four Content Categories

Percent Correct on TUCE Posttest

Category	101	121	223
1. Scarcity, functioning of economics systems, base elements of supply and demand	71%	66%	76%
2. Macroeconomic accounting and determination of GNP (income-expenditure theory)	63%	58%	71%
3. Money, banking, and monetary policy and government fiscal policies	63%	49%	70%
4. Determinants of economic growth and policies for stabilization and growth	54%	57%	62%

Percent Improvement on TUCE

Category	101	121	223
1. Scarcity, functioning of economic systems, base elements of supply and demand	42%	30%	47%
2. Macroeconomic accounting and determination of GNP (income-expenditure theory)	60%	66%	105%
3. Money, banking, and monetary policy and government fiscal policies	81%	40%	72%
4. Determinants of economic growth and policies for stabilization and growth	53%	33%	57%

Appendix B

Further Regression Analysis

One question which might arise with respect to the results presented in Table VIII concerns the sensitivity of conclusions concerning the three courses to alternative definitions of the dependent variable. Two alternative specifications of the dependent variable were tried--absolute score on the TUCE posttest and percent improvement (ratio of the difference between post less pre to pre). In both instances the TUCE pretest was included as an independent variable. The results indicate that no such sensitivity exists--the sign and significance of the dummy variables remains as reported in Table VIII. As a point of interest, R^2's are maximized by using percent improvement as the dependent variable and including the TUCE pretest as an independent variable--typical increase on the order of .2 (an increase in explained variation of 20%).

We also addressed the sex question which might be classified as a human capital variable. The problem here was to control simultaneously for sex and course. Thus five dummy variables were employed M223 (males in 223=1), F223 (females in 223=1), M121 (males in 121=1), F121 (females in 121=1), and F101 (females in 101=1). The comparison group is males in 101. The results of these regressions for the total TUCE and the three objective categories are presented in Table B.I. Again the gap closing version of the dependent variable was employed.

Looking first at the 101 male versus female comparison we find no significance difference on the overall TUCE and two of the three objective categories. On recognition and understanding question 101 females are

TABLE B.1.

Regression Results - Total TUCE
and the Three Objective Categories
Allowing for Sex

Dependent Variable		Constant	MSAT	VSAT	RHSC	UR	CE9	CE15	CE16
Total TUCE	B VALUE	-0.16818	0.00064	0.00019	0.00150	0.01641	-0.05519	0.00191	0.02571
	BETA		0.21811	0.06793	0.06383	0.05385	-0.13722	0.00735	0.09051
			(12.975)	(1.703)	(1.933)	(0.925)	(5.306)	(0.015)	(1.869)
	t ratio		3.602*	1.304	1.390	0.961	2.303*	0.122	1.367
Recognition and Understanding	b value	-0.14115	0.00066	-0.00033	0.00344	0.02245	-0.05900	0.00606	0.05066
	beta		0.16525	-0.09820	0.13628	0.05415	-0.10782	0.01716	0.13189
			(4.839)	(1.758)	(3.545)	(0.608)	(2.129)	(0.051)	(2.547)
	t ratio		2.199*	1.325	1.882	0.779	1.459	0.225	1.595
Simple Application	B VALUE	-0.32585	0.00079	0.00070	-0.00096	-0.00038	-0.04125	-0.00053	0.01327
	BETA		0.18121	0.18886	-0.03459	-0.00084	-0.06853	-0.00136	0.03122
			(9.093)	(10.162)	(0.357)	(0.000)	(1.344)	(0.001)	(0.226)
	t ratio		3.015*	3.187*	0.597	0.000	1.159	0.031	0.475
Complex Application	B VALUE	-0.17290	0.00044	0.00022	0.00263	0.04617	-0.07341	-0.00231	0.02147
	beta		0.10819	0.06291	0.10184	0.10887	-0.13112	-0.00638	0.05430
			(2.306)	(0.802)	(2.201)	(2.730)	(3.500)	(0.008)	(0.486)
	t ratio		1.518	0.895	1.483	1.652	1.870	0.089	0.697

* Indicates significance at the .05 level.

TABLE B.1.
(continued)
Regression Result - Total TUCE
and the Three Objective Categories
Allowing for Sex

OPI	M223	F223	M121	F121	F101	R^2	F Ratio	N
0.01221	0.13285	-0.07906	-0.10819	-0.15270	-0.09272	0.442	12.93691	225
0.10331	0.30979	-0.07942	-0.22059	-0.15338	-0.07498			
(3.763)	(10.169)	(1.529)	(6.193)	(5.599)	(1.587)			
1.939	3.189*	1.236	2.488*	2.366*	1.259			
0.00555	0.09037	-0.08041	-0.02469	-0.03886	-0.29917	0.1418	2.69501	225
0.03452	0.15489	-0.05937	-0.03700	-0.02869	-0.17784			
(0.273)	(1.652)	(0.555)	(0.113)	(0.127)	(5.800)			
0.522	1.285	0.744	0.336	0.356	2.408*			
0.0078	0.25289	0.02932	-0.12649	-0.17106	0.07042	0.45078	13.3846	225
0.04402	0.39403	0.01968	-0.17223	-0.11482	0.03805			
(0.693)	(16.703)	(0.095)	(3.838)	(3.186)	(0.415)			
0.832	4.086*	0.308	1.959*	1.784	0.644			
0.02518	0.06047	-0.19011	-0.15355	-0.22796	-0.09078	0.22803	4.8169	225
0.15308	0.10130	-0.13719	-0.22491	-0.16450	-0.05274			
(5.965)	(0.785)	(3.296)	(4.651)	(4.652)	(0.567)			
2.442*	0.886	1.815	2.156*	2.156*	0.752			

significantly below their 101 male counterparts. Both 121 males and females are significantly below 101 males on the overall TUCE. For the 121 males this is also true on two of the three objective categories while only true in one category for the 121 females. We find no significant difference between 223 females and 101 males while 223 males are significantly better than the 101 males in two of the four regressions. In summary then there is some slight evidence of a sex factor in 101 and 121--to the advantage of males in 101 and to the advantage of females in 121--and stronger evidence in 223--to the advantage of males.

As a final comment on the regression analysis there is the question of multicollinearity between variables. The correlation coefficients are reproduced in Table B.2. for the regressions presented in Table VI. Other than the expected high correlation between the two dummy variables, the remaining correlation coefficients do not suggest any high degree of mullicollinearity. The variables exhibiting the highest collinearity are three of the efficiency variables but even here the highest correlation coefficient is only about .5.

TABLE B.2.

Correlation Coefficients for Independent Variables

	MSAT	VSAT	RHSC	UR	CE9	CE15	CE16	OPI	D 223	D 101
MSAT	1.00000	0.42081	0.32598	-0.18389	-0.07000	-0.03134	-0.01344	-0.05920	0.11311	-0.13845
USAT	0.42081	1.00000	0.27025	-0.12329	-0.12385	-0.07432	0.00037	0.01080	0.18997	-0.20132
RHSC	0.32598	0.27025	1.00000	-0.02348	0.01337	0.10891	-0.01838	-0.08126	0.11522	-0.17058
UR	-0.18389	-0.12329	-0.02348	1.00000	0.12558	0.11836	0.13144	-0.00096	-0.31572	0.15120
CE9	-0.07000	-0.12385	0.01337	0.12558	1.00000	0.27324	0.46632	-0.10062	-0.19027	-0.12240
CE15	-0.03134	-0.07432	0.10891	0.11836	0.27324	1.00000	0.49692	-0.07668	-0.03833	-0.04424
CE16	-0.01344	0.00037	-0.01838	0.13144	0.46632	0.49692	1.00000	-0.10464	-0.15688	0.10220
OPI	-0.05920	0.01080	-0.08126	-0.00096	-0.10062	-0.07668	-0.10464	1.00000	0.00548	-0.11789
DU223	0.11311	0.18997	0.11522	-0.31572	-0.19027	-0.03833	-0.15688	0.00548	1.00000	-0.75819
DU121	-0.13845	-0.20132	-0.17058	0.15120	0.12240	-0.04424	0.10220	-0.11789	-0.75819	1.00000