

Utah State University

DigitalCommons@USU

---

All Graduate Theses and Dissertations

Graduate Studies

---

5-1973

## Achievement of Students in Introductory Social Science Courses Offered by the College of Humanities, Arts, and Social Sciences at Utah State University

James Collins Lewis  
*Utah State University*

Follow this and additional works at: <https://digitalcommons.usu.edu/etd>



Part of the [Counseling Psychology Commons](#)

---

### Recommended Citation

Lewis, James Collins, "Achievement of Students in Introductory Social Science Courses Offered by the College of Humanities, Arts, and Social Sciences at Utah State University" (1973). *All Graduate Theses and Dissertations*. 3161.

<https://digitalcommons.usu.edu/etd/3161>

This Thesis is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Theses and Dissertations by an authorized administrator of DigitalCommons@USU. For more information, please contact [digitalcommons@usu.edu](mailto:digitalcommons@usu.edu).



ACHIEVEMENT OF STUDENTS IN INTRODUCTORY SOCIAL SCIENCE  
COURSES OFFERED BY THE COLLEGE OF HUMANITIES, ARTS,  
AND SOCIAL SCIENCES AT UTAH STATE UNIVERSITY

by

James Collins Lewis

A thesis submitted in partial fulfillment  
of the requirements for the degree

of

MASTER OF SCIENCE

in

Psychology

Approved:

UTAH STATE UNIVERSITY  
Logan, Utah

1973

378.7  
L587ac  
e.2

#### ACKNOWLEDGMENTS

I express my appreciation and gratitude to Dr. Keith Checketts, my major professor, for his patience, assistance, and encouragement in completing this study. I also thank my other two committee members, Dr. Marlan Nelson and Dr. Elwin C. Nielsen, for their helpful suggestions and support. I especially thank Dr. Dan E. Jones, Dr. Mary Anna Baden, and Dr. Robert C. Oaks, Jr., for allowing me to conduct the testing in their introductory classes, summer session, 1972. Also I would sincerely like to thank all those students who were so cooperative in participating in this research.

James Collins Lewis

## TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS .....	ii
LIST OF TABLES .....	v
LIST OF FIGURES .....	vi
ABSTRACT .....	vii
Chapter	
I. INTRODUCTION .....	1
II. REVIEW OF LITERATURE .....	5
III. STATEMENT OF PURPOSES, OBJECTIVES, AND HYPOTHESES .....	18
Objectives .....	18
Hypotheses .....	19
IV. METHODS AND PROCEDURES .....	21
Design .....	22
Data and Instrumentation .....	23
Analysis .....	26
V. RESULTS .....	27
Primary Hypothesis .....	27
Subsidiary Hypotheses .....	30
VI. SUMMARY AND CONCLUSIONS .....	35
Conclusions .....	37
Discussion .....	39
Recommendations .....	53
LITERATURE CITED .....	54

## TABLE OF CONTENTS (Continued)

	Page
APPENDIX .....	61
VITA .....	65

## LIST OF TABLES

Table	Page
1. Mean change scores of students in selected introductory social science courses offered by the College of Humanities, Arts, and Social Sciences at USU . . . . .	28
2. T-test values for absolute change scores of students in selected introductory social science courses offered by the College of Humanities, Arts, and Social Sciences at USU . . . . .	29
3. The ANOVA table (disproportionate frequency per cell with unweighted means) . . . . .	32
4. Table of mean change scores of students in courses tested with criterion-referenced tests at the end of the course (posttest No. 1) and two months after posttest No. 1 (posttest No. 2) . . . . .	33
5. Correlation and reliability data relating to the unreliability-invalidity dilemma . . . . .	40
6. Correlations between tests . . . . .	43
7. Index values for ceiling effect evaluation . . . . .	46

## LIST OF FIGURES

Figure	Page
1. Model of the factorial design showing mode of testing and the college offering the course . . . . .	23
2. Factorial design (2x2)-Type of test x college offering the course--using mean percentage gain scores . . . . .	31

## ABSTRACT

Achievement of Students in Introductory Social Science  
Courses Offered by the College of Humanities, Arts,  
and Social Sciences at Utah State University

by

James Collins Lewis, Master of Science

Utah State University, 1973

Major Professor: Dr. Keith T. Checketts

Department: Counseling Psychology

The evaluation of student achievement in college has only recently become an important area of research. The main focus of this study was to determine the achievement of students in introductory social science courses offered by the College of Humanities, Arts, and Social Sciences at Utah State University. Several other variables were considered throughout the study including (1) the achievement of students in introductory courses offered by other colleges than the College of Humanities, Arts, and Social Sciences, (2) the effects of using standardized and teacher-made tests in measuring achievement, and (3) the achievement gain/loss of students tested two months after the completion of the course.

The students in four introductory courses--Political Science 111, Sociology 101, Geology 101, and Economics 200--were tested. The students in Political Science 111 and Economics 200 were tested with standardized tests specific to the subject area. The students in Sociology 101 and Geology 101 were tested with



teacher-made tests. All students were given a pretest at the beginning of the course and a posttest at the conclusion of the course. The students in Sociology 101 and Geology 101 were tested a third time two months after the termination of the course to measure long term retention.

Null hypotheses were used in the study. The analysis of variance with the F-test of significance was employed at the .01 level of significance for the primary hypothesis and the .05 level for the subsidiary hypotheses. In conclusion, the data indicated that students in introductory social science courses offered by the College of Humanities, Arts, and Social Sciences--Political Science 111 and Sociology 101--showed an increase in knowledge after taking those courses. The data also showed increases in knowledge of those students taking introductory courses--Geology 101 and Economics 200--which were offered by colleges other than the College of Humanities, Arts, and Social Sciences.

The data in relation to the subsidiary hypotheses indicated that (1) students in introductory courses offered by other colleges at USU showed larger knowledge gains than the students in the introductory social science courses offered by the College of Humanities, Arts, and Social Sciences, (2) students in courses measured with teacher-made tests scored higher than the students in courses measured with standardized tests, and (3) an interaction existed between the method of testing and the college offering the course. Finally, the students in the two courses--Sociology 101 and Geology 101--who were tested two months after the completion of the course continued to show positive achievement gains.

## CHAPTER I

### INTRODUCTION

Demands for evaluation of colleges and universities have come from many sources--college presidents' offices, other college administrative offices, the faculty, the students, the state government, legislative bodies, education offices, pressure groups, courts, and the general public (McConnell, 1971; Dressel and Pratt, 1971).

There are many reasons for the demands for evaluation of colleges and universities. McConnell (1971) wrote:

Turmoil and disruption on the campuses; political action by students and faculty members; severe shrinkages in governmental, corporate, and individual incomes, coupled with rising taxes; and mounting distrust of higher education by the public are behind the increasing demand for colleges and universities to justify what they are doing and to disclose the effectiveness and efficiency of their operations. (p. 446)

Dressel and associates (1971) offered other reasons for the demand for evaluation such as "changing objectives and goals. . . , changing patterns of administrative organization and the roles of administrators, changes in academic programs, the changes and increasing complexity of institutional environments, social and physical, the demands of students, and new emphasis on cooperation and coordination of institutions" (p. 321). Manning (1969) stated that demand for evaluation is due to the promise evaluation holds for improving the educational process and making the educational system accountable.

The demand for evaluation of colleges and universities makes it imperative that data be provided to legitimize fund requests (Cronbach, 1963). "Colleges and universities must now seemingly daily provide justification of their financial existence with quantitative data." (Brich, 1971, p. 28) Data concerning the prediction of enrollments and assessment of physical plants are needed (Dressel and Pratt, 1971). Data about the costs of instruction are needed (Dressel and associates, 1971). Data about what happens to the student, the "value added," are also needed (Qualitative and Quantitative Performance and Achievement in Higher Education Report, 1971). All this data, but especially the latter, are needed to provide the public, the student, and others with confidence in the educational system (Lessinger, 1970).

Information about student academic growth is especially sketchy (Harris and Hurst, 1972). A review by Feldman and Newcomb (1969) of studies of college influence indicated that most studies have dealt primarily with changes in student attitudes, values, interests, and beliefs as a result of the college experience. Dressel and Pratt (1971) stated that practically no research has been directed at identification of the changes in student academic knowledge developed while attending college. "The typical approach to institutional evaluation is to focus on faculty and facilities, not to attempt to see what happens to students as a consequence of having been exposed to them." (Dyer, 1966, p. 463)

A common approach to institutional self-study not too unlike other types of college and university studies is to let departments review their own activities in terms of a general analysis and recommendations. These departmental reviews typically do not require analysis of student academic growth, however.

"Thus it happens regularly that, despite programs of institutional self-study, many units in the university change little or not at all." (Dressel and Dietrich, 1967, p. 25) It thus appears that though the public and others are demanding "product reliability" in terms of student capabilities, the research approaches have not been, and are not oriented to provide such data.

Utah State University (hereafter referred to as USU) is not unlike other colleges and universities. USU administrators also must develop and collect data with which to justify their institutional budgets. USU faculty research publications since 1959 and USU graduate theses and dissertations, however, offer little or no data to confirm the academic achievement of USU students. Some USU administrators have begun their own studies specifically aimed at obtaining results to meet the external and internal demands for evaluation.

The administrators of the College of Humanities, Arts, and Social Sciences have decided to investigate the academic growth of USU students in lower division, undergraduate introductory courses (hereafter referred to as "introductory courses"). There are several reasons for concentrating their initial efforts on the introductory courses. First, the lack of research funds has made it necessary to restrict the scope of the research. Second, the results from the General Examinations of the College Level Examination Program (CLEP) have shown that some students possess the knowledge of introductory courses offered by the College of Humanities, Arts, and Social Sciences before they enter that college (Christiansen, 1972). It is thus necessary to determine those introductory courses which add little to the pool of knowledge of USU students and to improve them or to replace them in the curricular offerings.

Thus far, the administrators of the College of Humanities, Arts, and Social Sciences have been able to initiate only one major study. That study (Christiansen, 1972), concerning learning in English Composition classes at USU, has generated some doubt that English Composition classes at USU contribute to the knowledge of the students taking those classes. The need is for additional research in other introductory courses at USU to provide administrators with current and ample data to confirm or disconfirm the importance of the contribution of these introductory courses to the academic growth of USU's students.

The problem is, then, that despite the research done by the College of Humanities, Arts, and Social Sciences concerning the learning which occurs in English Composition classes at USU, there is still a lack of information concerning the learning which occurs in other introductory courses offered by the College of Humanities, Arts, and Social Sciences at USU.

## CHAPTER II

### REVIEW OF LITERATURE

The research of this study may best be classified as an administrative study of a college within a university. This study was conducted for the College of Humanities, Arts, and Social Sciences to obtain an exact measurement of student learning in selected introductory courses (specifically introductory social science) offered by that college. As an administrative study conducted for specific purposes, the research measures (1) only academic gains of students, (2) only selected introductory courses offered by certain colleges at USU, and (3) only student growth as measured by non-departmental related research personnel.

This research is thus different or unique in many ways from the other types of research that have been done at the university. Unlike the departmental studies it does not rely on faculty recommendations (Dressel, 1968). Unlike institutional self-study it is concerned only with examining a small segment of the university. Unlike educational research it is concerned with identifying only gains/losses in student growth rather than pinpointing the curricular or instructional methodology correlated to the change. Unlike institutional research it makes no effort to compare student gains in courses at USU with students in similar courses at other universities (Dressel and associates, 1971).

Before a review of research related to administrative studies of academic achievement in selected introductory courses offered by the College of Humanities, Arts, and Social Sciences at USU is presented, a critical examination of the factors involved in this study needs to be presented.

The effect of college education has been a topic of interest for many years. The "value added" concept (Task Force Report, 1971) has been recently coined to represent the changes that occur in students as a result of their four years in college. "Value added" is a comprehensive term. It refers to "changes in intellectual abilities, changes in factual knowledge, changes in emotional sensitivity and changes in values" (Barton, 1959, p. 14). The research of this thesis will concentrate specifically on that aspect of "value added" related to changes in factual knowledge.

The acquisition of factual knowledge, referred to also as achievement, knowledge acquisition, academic growth, and change, is important in the educational process because it provides the informational basis for other intellectual abilities and skills (Barton, 1959; Proceedings, 1969; Bloom, 1959). Though the processes of measuring change in factual knowledge as a result of the college experience has been inadequate in the past (Cronbach, 1963; Astin, 1971), the pretest-posttest technique is recommended to measure the overt behavioral changes in performance (Buehe, 1969; Glaser, 1968; Roberts, 1969; Forehand, 1971; Dyer, 1970; Milton and Shoben, 1968; Westland, 1971).

The difference in pretest-posttest means (i. e. mean change scores) on tests of academic achievement for groups of students is necessary in the evaluation of the teaching and curriculum in institutions of higher education (Ford, 1971;

Cronbach, 1963). This evaluation of both teaching and curriculum is essentially the evaluation of the college's input into the educational process (Birch, 1970). Though the mean change score does not specify the variable (i. e. teaching or curriculum) which is responsible for mean loss or gain in student scores, it does provide an index for judging a college's effectiveness and contribution to the student's acquisition of factual knowledge (Cronbach, 1963; Dyer, 1966; Manning, 1969; Roberts, 1969).

In order to obtain an accurate indication of a college's effectiveness and contribution to the student's knowledge several factors must be considered (1) the type of learning measured, (2) the type of test used to measure the change, and (3) the type of mean change score used to quantify the change.

There are four types of learning which can be measured to investigate the acquisition of factual knowledge--recall, retention, transfer, and relearning (Kagan and Havemann, 1968). Though all are desirable approaches to the measurement of learning, evaluation using recognition is (1) the approach with which the student is most familiar, (2) the approach which most easily can be objectivized, and (3) the approach which can be used to measure both short and long term memory, the latter being a valuable indication of learning as distinguished from mere memorization (Postman, 1967).

The testing instruments are crucial in determining what is measured. In measuring the gains in factual knowledge two types of achievement tests are applicable--the standardized achievement test and the criterion-referenced achievement test.



The standardized test with its advantages of objectivity and standardization may measure too broad a spectrum of the subject area (Cronbach, 1963; Palalino, 1970). The criterion-referenced test (CRT) on the other hand, though an excellent test to use to determine how well a curriculum is achieving its objectives (Cronbach, 1963), has the drawback of questionable objectivity (reliability and validity) and standardization (scoring and administration).

The type of mean change scores used to quantify the change is also important to the accurate measurement of academic growth or change. There are four major types of mean change scores--crude or raw gain, true gain (Lord, 1956), residual gain (Cronbach and Furby, 1970), and percentage gain (Menlo and Johnson, 1971). Of the methods, the percentage gain score appears to have the most advantages. Percentage gain scores are more highly correlated to post-test scores in a positive direction. The range of variation of scores in the distribution of percentage gain scores is small. The use of percentage gain appears to have a normalizing effect on the distribution of scores. The percentage gain scores don't appear to penalize the high starters. Percentage gain scores provide a common base on which to compare scores with different bases (Menlo and Johnson, 1971).

Regardless of the type of mean score used to measure gain, however, several problems still remain a part of the measurement of change. First, the tests used may not measure the entire range of achievement in the course. In other words, the test may not have been difficult enough to measure the upper limits of knowledge. This is the "ceiling effect." Second, the student who earned a high score on the pretest will earn somewhat of a lower score on the

posttest, whereas the students who earned a low pretest score will receive a somewhat higher score on the posttest. This is known as the "regression effect" and occurs partially because of the errors of measurement due to a poorly constructed test and the guessing element and because the pretest and posttest are partially correlated. Third, the very elements that make a test sound psychometrically (1) highly reliable scores, and (2) highly correlated test-retest scores obtained within a year or two of each other, are not the elements which make a test good to measure change. Fourth, tests which are not factorially pure may reflect different levels of ability for different students. Fifth, but not critical when using the percentage gain scores, the gain scores generally assume equal intervals at all points in the test (Borg, and Gall, 1971).

Three other factors affect the accuracy of change scores. First, the focus of instruction can concentrate on the test material to such an extent that the test is taught during the course of the instruction. Second, the differential motivation of the students can be an uncontrolled extraneous variable. Third, the content of the test may be more closely parallel to the abilities of average and below average college students--"the content effect." Even with the faults inherent in using change scores, however, sophisticated researchers still use them for lack of better measurement procedures (Cronbach and Furby, 1970).

With regard to the various approaches available with which to measure achievement at the college level, the research indicates that some approaches may facilitate the accuracy of the measurement. This researcher thus has chosen to delimit the measurement of student achievement to (1) measuring the acquisition of factual knowledge, (2) using mean change scores, (3) using the

percentage gain scores to quantify the change scores, (4) testing recognition as opposed to recall or transfer, and (5) using both standard achievement and criterion-referenced tests.

Having considered these delimiting factors, a review of the research related to administrative studies of academic achievement in selected introductory courses offered by the College of Humanities, Arts, and Social Sciences at USU is now presented.

There have been different types of studies investigating achievement in college. Few of these studies apply directly to achievement gains in specific college courses, however.

Studies about general intellectual or scholastic aptitude change (Connel, 1954; Livesay, 1939; Louise, 1947; Shue, 1948; Silvery, 1951), though indicating intellectual development during college, do not relate directly to questions about specific knowledge acquisition due to a specific course or a series of courses.

Studies relating various indices of college quality to student learning and intellectual development (Astin, 1968; Rock, Centra, and Linn, 1970; Rock and Centra, 1971; Rock, Baird, and Linn, 1972) have used measures of final achievement (i.e. final test score) rather than measures of achievement gain or loss (i.e. pretest-posttest difference).

Many studies of achievement gains after one, two, or four years of college (Lenning, 1968; Harris and Hurst, 1972; Harris and Booth, 1969; Beanblossom, 1969) have reported achievement gains in broadly defined areas of study (i.e. humanities, social science, mathematics, and natural science), but not in specific courses.

Only in studies related to a comparison of teaching methods in specific college courses (primarily introductory courses) is any information found related to achievement gains. Even then, many studies comparing teaching methods (Emery and Enger, 1972; Nelson, 1971; VanPelt and Austin, 1971; Burman and Flaherty, 1968; and, Lareau, 1971) have placed more emphasis on comparing the gains due to the teaching methods used than on overall gains due to taking the course as such. Some, however, have produced findings relevant to the impact of introductory courses.

Tovrea (1971) compared the effects of two methods of teaching general biology. He reported (1) that students showed no significant gains in achievement on critical thinking skills at the completion of the course regardless of the method, and (2) that students showed significant mean gain scores in knowledge of biology under both methods of instruction.

Futz (1970) reported no gains in freshman English classes at a community college in the Boston area in 1969. Futz compared a non-directive teaching method and a traditional teaching method. The measurement of achievement of both groups was conducted before and after the course of instruction with two 20-minute and one 40-minute essays. These essays were rated by five experienced raters on a scale of 1-3. The results were (1) that there was no measurable change in the writing achievement of either group, and (2) that there was no decrease in writing ability in the non-directive group. Futz (1970) concluded that student gains were negligible under either teaching method.

Sharon (1971) stated that three studies had found significant achievement gains by college students after taking a college course or a group of college

courses. In all three cases the General Examination (GE) of the College Level Examination (CLEP) were used as the testing instruments.

Two other researchers, Grover (1970) and Norton (1967), have offered their comments on performance gains in specific college courses, but have failed to present any research data to verify their conclusions.

At Utah State University (USU) two studies have been reported which relate directly to achievement gains in introductory courses.

First, Erickson (1969) published a thesis on the "Effects of a College Mathematics Sequence Upon the Attitudes and Achievement in Mathematics of Prospective Elementary School Teachers." While in the Mathematics 21 (Mathematics for Elementary Teachers) course, 126 students took pretests and posttests, consisting of the Stanford Achievement Test Series. "Of those students who took both the pretest and the posttest in mathematical achievement, only four showed no increase on the posttest over the pretest. No one had the same posttest score as pretest score. This would perhaps indicate that almost everyone learned something in the course." (Erickson, 1969, p. 25)

The second study was conducted by Christiansen (1972). Christiansen tested a total of 110 entering freshman at USU, falling in five groups (24 students in each group). The groups were as follows:

The Advanced Placement Group: This group consisted of 1971-72 USU freshmen students who received a score of 3, 4, or 5 on the Advanced Placement Test in English Composition.

The CLEP Credit Group: This group included 1971-72 USU freshmen students who received credit for freshman English by scoring 450 or more points on the CLEP General Examination (GE) in English Composition.

The CLEP Waiver Group: This group contained 1971-72 USU freshmen students who waived the freshman English requirement by scoring from 390-449 points on the CLEP General Examination in English Composition.

The A-B Grade Group: This group was 1970-71 USU freshman students who received credit for three quarters of freshman English and whose freshman English grade point average (GPA) was either a "A" or a "B".

The C-D Grade Group: This group was composed of 1970-71 USU freshman students who received credit for three quarters of freshman English and whose freshman English GPA was either a "C" or a "D".

All five groups of students took the CLEP subject examination in English Composition, which consists of a 90-minute objective part and a 90-minute essay part, at the end of their freshman year at USU. The objective part of the test was machine scored and designed so that the individual could receive a score of 20 to 80 points. As for the essay part, three members of the USU English Department scored each essay on a scale of 1 to 12 points. The grades of the three essay graders were summed providing a range of scores from 3 to 36 points. The mean and standard deviations (S. D.) of each group are presented in the appendix.

Differences between each pair of group means was calculated. On the objective section of the test, the results show significant differences ( $p < .01$ )

(1) between the Advanced Placement Group and the C-D Grade Group, (2) between the Advanced Placement Group and the CLEP Waiver Group, (3) between the Advanced Placement Group and the A-B Grade Group, (4) between the CLEP Credit Group and the C-D Grade Group and, a significant difference ( $p < .05$ ), (6) between the A-B Grade Group and the CLEP Waiver Group (tables in the appendix). Actually, the groups can be categorized as those groups taking English composition classes at USU and those groups not taking English composition class at USU. The mean objective test score of those groups required to take English composition was higher ( $p < .01$ ) than the mean objective test score of those students not required to take English composition.

The results seem to show that (1) students who did not take English composition scored higher on the objective test at the end of three terms than did the USU students who took English composition, and (2) the students who took English composition did not score any different on the essay part than those students who were not required to take English composition. The implied conclusion is that English composition had failed to increase the knowledge of those students taking the course. Caution needs to be taken in accepting such a conclusion, however, because the groups already were not comparable by ability and/or educational background. No attempt was made to determine the initial mean scores on an objective exam pretest for the different groups thus providing a base from which to determine change.

Despite the caution which needs to be taken in accepting the conclusion that English composition is failing to increase the knowledge of those students taking the course, this study has created enough doubt in the minds of USU

administrators to question the achievement of students in other introductory courses offered by the College of Humanities, Arts, and Social Sciences at USU.

The research conducted over the years at USU, however, provides USU administrators with little information of the achievement of students in other introductory courses offered by the College of Humanities, Arts, and Social Sciences. The research has been primarily concerned with the comparison of teaching methods (Lauer, 1972; Lyon and Marston, 1972; White, 1969; Willems, 1970) or the prediction of college achievement via some variable(s) (Giles, 1965; Day, 1939; Bateson, 1949; LaPray, 1962; Johnson, 1951; Lunt, 1956; Hendricks, 1967; Tyrrel, 1958; Pickett, 1958). As was stated in the introduction, a review of USU faculty research publications since 1959 offers no data relating to achievement gains in college introductory courses at USU.

In summary, several observations can be made. First, there has been a limited amount of research investigating achievement gains of students in specific introductory courses offered by the College of Humanities, Arts, and Social Sciences. This review has uncovered only two studies at USU concerning gains of students in introductory courses. There is a need for more research along these lines.

Second, the research that does exist relating to achievement in introductory courses does not provide a consistent picture of the achievement in college courses from which one could predict achievement gains in those courses. Lenning (1968) suggests achievement gains exist after one and two years of college. Harris and Hurst (1972) agree that significant gains in knowledge are



made in the first year of college. Harris and Booth (1969) found gains in humanities, natural science, and history, but not in math and English after one year in college. Beanblossom (1969) found gains in humanities, natural science, and history, but not in math and English after one year in college. Beanblossom (1969) suggested that greater knowledge exists after two years of liberal arts courses. Tovrea (1971) found significant mean gains in knowledge of biology. Putz (1970) reported that achievement gains in freshman English classes after one year were negligible. Sharon (1971) reported gains in college courses.

Third, the research indicates that achievement gains might be more probable in certain subject areas than others. There are indications that English and social science courses might not be as prone to display gains as courses in natural science, math, and humanities. Lenning (1968) said English scores were least sensitive to student growth. English and mathematics were the only two subjects that showed a negative mean score change. Harris and Hurst (1972) found social science to demonstrate the least gain. Harris and Booth (1969) reported small gains in social science but not in English. Beanblossom (1969) reported hardly any gains in the social Sciences. Jewell (1970) and Putz (1970) report non-significant gains in English composition classes.

Thus, knowing that there is (1) a limited amount of research investigating achievement gains of students taking specific introductory courses offered by the College of Humanities, Arts, and Social Sciences, (2) an inconsistency in the research data concerning the achievement gains of students in introductory college courses offered by the College of Humanities, Arts, and Social Sciences,

and (3) an indication that achievement gains in English and social science introductory courses might contribute least to the knowledge gains of the students, then investigation of the achievement of students at USU in introductory social science and English courses offered by the College of Humanities, Arts, and Social Sciences seems warranted. However, due to the recent demands placed on the English Department as a result of the Christiansen study (1972) the investigation of achievement of students in English composition classes was deemed inadvisable, thus limiting the scope of this investigation to the achievement of students in introductory social science courses.

CHAPTER III  
STATEMENT OF PURPOSES, OBJECTIVES,  
AND HYPOTHESES

The need for information concerning achievement of students taking specific introductory courses, specifically social science courses offered by the College of Humanities, Arts, and Social Sciences at USU, has been presented in the introduction and the review of literature. In the introduction, the need for administrators to account for the request of funds was presented. In the review of literature, the absence of appropriate research concerning achievement in selected introductory courses was presented.

The research indicated two course content areas, social science and English, in the course offering of the College of Humanities, Arts, and Social Sciences which displayed little or no evidence of contribution to the achievement gains of students, but only the social science area was chosen for investigation. The purpose of this study was thus to investigate the achievement that occurs in selected social science introductory courses offered by the College of Humanities, Arts, and Social Sciences at USU.

Objectives

The primary objective of this research was to determine the achievement of students taking introductory social science courses offered by the College of Humanities, Arts, and Social Sciences.

In an attempt to add further clarity to the results, several other factors were investigated. This investigation of other factors provided several subsidiary objectives in the research which were: (1) to determine the achievement of students in the introductory social science courses offered by the College of Humanities, Arts, and Social Science as compared to the achievement of students in introductory courses not offered by the College of Humanities, Arts, and Social Sciences, (2) to determine the effect of testing with standardized tests and criterion-referenced tests on the measurement of achievement gains, (3) to determine the long term retention of students in courses tested with criterion referenced tests, and (4) to determine the interaction between the mode of testing and the college offering the course.

### Hypotheses

The hypothesis in regard to each objective was:

#### Primary hypotheses

1. There is no difference between the pretest and posttest scores of students taking introductory social science courses offered by the College of Humanities, Arts, and Social Sciences at USU.

#### Subsidiary hypotheses

2. There is no difference in achievement between students taking introductory social science courses offered by the College of Humanities, Arts, and Social Science and the achievement of students taking introductory courses offered by the other colleges at USU.

3. There is no difference in achievement measured by standardized tests and achievement measured by criterion-referenced tests.

4. There is no interaction between the mode of testing and the college offering the course.

5. There is no difference between achievement of students tested with criterion-referenced tests at the beginning of the course (i. e. Sociology and Geology) and the achievement of those same students tested with the same criterion-referenced test two months after the course termination.

The primary hypothesis is tested at the  $p < .01$  level of significance and the subsidiary hypotheses are tested at the  $p < .05$  level of significance. These levels of significance were chosen because of the expected large difference in pretest and posttest results for the introductory social science courses tested (primary hypothesis) and the expected smaller differences that would occur in analyzing the subsidiary hypothesis.

## CHAPTER IV

### METHODS AND PROCEDURES

The population of this study consisted of all students taking introductory social science courses offered by the College of Humanities, Arts, and Social Sciences at USU. The College of Humanities, Arts, and Social Sciences at USU includes ten departments. Of these ten departments, three could be accurately classified as offering social science courses--those courses pertaining to human society, its classes, and individual relationships, past and/or present (American Heritage Dictionary, 1971; Dressel, 1968). These three departments are (1) History and Geography, (2) Political Science, (3) and Sociology, Social Work, and Anthropology. Each of these departments offers its basic lower division undergraduate introductory courses at either the 100 or 200 number level. An essential element of these courses is that they usually are offered every term. There are presently 25 courses so numbered and offered by these three departments. In the areas of political science, sociology, social work and anthropology, one basic introductory course is listed. In the areas of history and geography, 10 basic introductory courses are listed, however, only three of these are categorized as the initial introductory social science course in the series. Thus, of the 25 introductory social science courses offered by these three departments, seven can actually be classified as basic prerequisite first term

introductory social science courses. These courses are the ones referred to in this study.

The sample consisted of intact classes of USU students in two of the seven introductory courses, in the population--Political Science 111 (American State and Local Government and Politics), and Sociology 101 (Introductory Sociology).

### Design

Besides these two introductory social science classes two other basic prerequisite first term introductory classes offered by colleges other than the College of Humanities, Arts, and Social Sciences--Geology 101 and Economics 200--were chosen to test. With this particular array of courses, not only could achievement be examined in the two introductory social science courses, but a comparison could be made between the achievement of students in the social science classes offered by the College of Humanities, Arts, and Social Sciences and the achievement of students in introductory courses offered by the other colleges.

The four classes were also distinguished by the method used to test each class. Two classes were tested with standardized tests--Economics 200 and Political Science 111. The other two classes--Sociology 101 and Geology 101--were tested with criterion-referenced (teacher-made) tests. This grouping of differential test composition allowed for the analysis of the effects of test type on the achievement gains of the students. The design conforms to the 2x2 factorial design below:

COLLEGE OFFERING THE COURSE	MODE OF TESTING	
	STANDARDIZED TESTS	CRITERION-REFER- ENCED TESTS
Selected introductory social science courses offered by the College of Humanities, Arts, and Social Sciences	POLITICAL SCIENCE	SOCIOLOGY
Selected introductory courses offered by other colleges at USU	ECONOMICS	GEOLOGY

Figure 1. Model of the factorial design showing mode of testing and the college offering the course

#### Data and Instrumentation

Selection of tests. The selection of the standardized tests was partially predetermined. In the case of Economics 200, Lyon and Marston (1972) used the Test of Understanding in College Economics prepared by the Committee for a College-Level Test of Economic Understanding of the Joint Council on Economic Education. The students were given the "B" version of the test at the beginning of spring term and the "A" version (parallel form) at the end of spring term 1972.

In the case of Political Science 111, the Principles of Democracy Test produced by Science Research Associates was used. The test was designed for use with grades 9-12 but because (1) the test was available via the counseling department at no cost, (2) other alternate political science exams demonstrated



no better qualifications, and (3) the test met the basic administration time requirements (1 hour maximum), it was used.

The choice of criterion-referenced (teacher-made) tests for Sociology 101 and Geology 101 was determined largely by the lack of acceptable standardized tests in these areas of study at the college level. The professors of both Sociology 101 and Geology 101 were asked to participate in compiling a test for purposes of measuring achievement gain. The test was to measure the student's comprehension of the essential elements of the course. The only other requirement was that the tests last no longer than one hour. Dr. Robert Oaks (Geology 101) prepared a 50 question test requiring the students to select the right answer from a choice of answers and place that answer on a separate answer sheet. Dr. Mary Anna Baden (Sociology 101) prepared a 50 question multiple-choice test from a compilation of questions drawn from various sociology textbooks. Political Science 111, Geology 101, and Sociology 101 used the same test for the pretest and posttest in their respective courses.

Administration of the tests. The Economics 200 tests, pretest (Form B) and posttest (Form A), were administered spring term, 1972, by Dr. Glen Marston. The Political Science 111 test was administered by this researcher at the beginning and end of the first session of summer school, 1972. The Geology 101 and Sociology 101 tests were administered by the respective instructors of those classes both at the beginning and the end of the first session of summer school, 1972.

The data for Economics 200 was obtained from the spring research data compiled by Lyon and Marston (1972) rather than testing students in Economics

200 during the first session of summer school, 1972. The first, and most important, was that the dean of the College of Humanities, Arts, and Social Science was anxious to obtain data on the achievement in these four courses by the end of the first session of summer school, 1972. Economics 200 terminated only at the end of the second session of summer session, 1972, therefore making the data impossible to gather in time. Equally as important, spring quarter data for Economics 200 was used because (1) the achievement data for Economics 200 was available for a large sample group, (2) Economics 200 was a full term (10 week course) in the summer just as it was in the spring, and (3) Lyon and Marston (1972) had used a well known reputable test to measure achievement. These advantages were enough for this researcher to accept the data from spring term Economics 200.

The students in the two courses that were tested with teacher-made tests, Sociology 101 and Geology 101, were tested a third time at the beginning of Fall Term, 1972. This additional testing was to ascertain the retention of the course knowledge. This procedure insured that the student scores represented more than regurgitation of learned facts acquired in the short summer session (5 weeks). The delayed testing also prevented any possible effects of instructor teaching of the test. The latter was not the case in the other two courses as the instructors had no knowledge of the questions on the standardized tests. Thus, the test results at the end of the course were assumed to represent valid measures of achievement (i.e. unaffected by short term memory and teacher keying of course content to the test questions).

### Analysis

The analysis of the achievement gains/losses in each class was conducted by using the t-test of significance with raw score gains and losses. A second analysis of the achievement gains in each class was conducted by using the t-test of significance with percentage gain scores.

The analysis of the subsidiary hypotheses required more detailed procedures. The first three subsidiary hypotheses were analyzed in a 2x2 factorial design using analysis of variance with the F-test of significance.

The fourth subsidiary hypothesis was analyzed with the use of the t-test of significance. In this case the data included calculations of gains with post-test scores at the end of the first summer session, 1972, and again with the post-test scores at the beginning of the fall term, 1972.

## CHAPTER V

## RESULTS

The results of this study are presented in five sections related to the five hypotheses of this thesis.

Primary Hypothesis

Hypothesis No. 1: This hypothesis stated that there was no difference between the pretest and posttest scores of the students taking selected introductory social science courses offered by the College of Humanities, Arts, and Social Sciences at USU. The results (mean change scores) are presented in Table 1 with raw scores and percentage scores.

These results showed a difference in the mean achievement change from pretest to posttest. Whereas the simple raw score mean change indicated less of a change in Political Science than Sociology, the mean percentage change showed less change in Sociology than in Political Science. This result was due to the special characteristic of using percentage change scores as the percentage change score did not penalize the high starters as did the simple raw gain score. The mean pretest scores of the students in Political Science were much higher ( $\bar{X} = 45.5$  out of 65 scorable points or 70 per cent correct) than the mean pretest scores of the students in Sociology ( $\bar{X} = 27.3$  out of 50 scorable points or 55 per cent correct). Thus the raw change scores tended to reduce the magnitude

Table 1. Mean change scores of students in selected introductory social science courses offered by the College of Humanities, Arts, and Social Sciences at USU

COURSE	MEAN RAW SCORE		MEAN CHANGE SCORE	
	PRETEST	POSTTEST #1*	RAW GAIN	PERCENTAGE** GAIN
Political Science (PS 111)	45.5	52.8	+7.3	+36.6
Sociology (Soc 101)	27.3	35.0	+7.7	+32.3

\* Posttests will be distinguished as #1 and #2 to distinguish between the posttest given at the conclusion of the course (posttest #1) and the posttest given approximately two months after the conclusion of the course (posttest #2).

\*\* Percentage gain scores are calculated according to the formulas provided by Menlo and Johnson (1971) below:

When the final score for an individual exceeded his initial score, the percentage gain score was defined as:

$$P.G. = \frac{(x_2 - x_1)}{R_P} 100$$

where  $x_2$  = final score  
 $x_1$  = initial score  
 $R_P$  = highest possible score minus initial score

When the final score was less than the initial score, percentage gain was defined as:

$$P.G. = \frac{(x_2 - x_1)}{R_n} 100$$

$x_2$  and  $x_1$  same as above  
 $R_n$  = initial minus lowest possible score

of the change in the case of the high starters which increased the error due to the ceiling effect. The percentage change minimized the error due to the ceiling effects by maximizing the scores of the original high starters. In this case the percentage gain scores gave a more accurate indication of change from the pretest to the posttest.

The absolute differences in change were analyzed with the t-test to determine if the differences were statistically significant. The data are presented in Table 2 below.

Table 2. T-test values for absolute change scores of students in selected introductory social science courses offered by the College of Humanities, Arts, and Social Sciences at USU.

COURSE	RAW CHANGE SCORE VALUES		PERCENTAGE GAIN VALUES	
	MEAN CHANGE	t-test value	MEAN CHANGE	t-test value
Political Science	+7.3	6.7*	+36.6	8.2*
Sociology	+7.7	5.1**	+32.3	5.3**

\*  $16^F .99 = 2.921$

\*\*  $8^F .99 = 3.355$

Table 2 shows that the absolute mean changes for Political Science and Sociology, whether calculated with raw change scores or simple percentage change scores were significant. This means that there was one chance in a 100 that the difference obtained between the pretest and posttest might have been due

to chance alone. Thus, the data indicated that the students increased in knowledge of the subjects political science and sociology during the period in which they took the course. The primary null hypothesis was rejected as there was a difference between the pretest and posttest scores of students taking selected introductory social science courses offered by the College of Humanities, Arts, and Social Sciences at USU.

#### Subsidiary Hypotheses

Hypothesis No. 2: This hypothesis stated that there was no difference in achievement between students taking introductory social science courses offered by the College of Humanities, Arts, and Social Science and the achievement of students taking introductory courses offered by other colleges at USU.

A 2x2 factorial design containing percentage change scores for the four classes tested was used to determine the validity of this hypothesis. The same design was used in the analysis of hypothesis No. 3 proposing no difference in achievement measured by standardized tests and achievement measured by criterion-referenced tests, and hypothesis No. 4 stating no interaction between the mode of testing and the college offering the course. The pretest and posttest No. 1 (the posttest given at the conclusion of the course) achievement scores were used to determine the percentage change scores in the four classes. The 2x2 factorial design is presented below in Figure No. 2.

Since each class had a disproportionate number of students the two way ANOVA with disproportionate frequency per cell method of unweighted means was used to determine the statistical significance of the difference between the

COLLEGE OFFERING THE COURSE	TYPE OF TEST		TOTAL
	STANDARDIZED TEST	CRITERION- REFERENCED TEST	
College of Humanities, Arts, and Social Sciences	PS 111 +36.6	SOC 101 +32.3	+68.9
Other colleges	ECON 200 +25.0	GEO 101 +58.1	+83.1
TOTAL	+61.6	+90.4	+152.0

Figure 2. Factorial design (2x2) - Type of test x college offering the course-- using mean percentage gain scores

mean percentage change scores of students taking social science courses from the College of Humanities, Arts, and Social Science and the mean percentage change scores of students taking introductory courses from other colleges at USU. The ANOVA F-value for the difference between the rows (displayed in Table 3 below) was  $F = 13.62$  and was significant, therefore, the null hypothesis was rejected. It should be noted that the extremely high mean percentage change value in geology appeared to make the F-value for the difference between the sum of the rows significant and was the source of much of the variance.

Hypothesis No. 3: This hypothesis is stated that there was no difference between the achievement measured by the standardized tests and the achievement measured by the criterion-referenced tests. Table 3 shows that the mean difference between columns provides a value of  $F = 11.28$ . This value was significant. The results show that there are five changes in 100 that the difference between



Table 3. The ANOVA table (disproportionate frequency per cell with unweighted means)

SV	DF	SS	MS	F
Between Columns	1	207.36	207.36	11.28*
Between Rows	1	205.41	250.41	13.62*
Interaction	1	149.69	149.69	8.14*
Error	81	22,060.27	MSE (ADJ) 18.38	

$$*F_{.95}(1, 81) = 4.00$$

achievement of students measured with a standardized test and the achievement of students measured with a criterion-referenced test will occur by chance alone. The null hypothesis must therefore be rejected. Again, the extremely high mean percentage change score in geology appeared to be the source of much of the variance between the columns.

Hypothesis No. 4: This hypothesis stated that there was no interaction between the mode of testing and the college offering the course. Table 3 shows the interaction F-value to be 8.14. The results indicate that 5 out of 100 times there will be an interaction between the college offering the course and the type of test used to measure achievement by chance alone. The Scheffe comparison test was used to investigate the observed sources of interaction. Using the comparison of single means only, the interaction between the mean percentage

change score of students in Economics 200 and Geology 101 provided a confidence interval which did not cover zero. Thus, the interaction was significant.

Hypothesis No. 5: This hypothesis stated that there was no difference between the achievement of students in geology and sociology tested with a criterion-referenced test at the beginning of the course and the achievement of these same students tested with the same device two months after the end of that course.

The absolute change data (both raw score change and percentage change) for the two courses tested with criterion-referenced tests is presented in Table 4 below:

Table 4. Table of mean change scores of students in courses tested with criterion-referenced tests at the end of the course (posttest No. 1) and two months after posttest No. 1 (posttest No. 2)

COURSE	MEAN CHANGE SCORES			
	RAW CHANGE		PERCENTAGE CHANGE	
	POSTTEST	POSTTEST	POSTTEST	POSTTEST
	NO. 1	NO. 2	NO. 1	NO. 2
	-PRETEST	-PRETEST	-PRETEST	-PRETEST
GEOLOGY	+22*	+14.5*	+58.1*	+39.0*
SOCIOLOGY	+ 7.7*	+ 6.0*	+32.3*	+24.6*

\*( $10^F .95 = 2.228$  ;  $8^F .95 = 2.306$ )

The comparison of pretest-posttest No. 1 and pretest-posttest No. 2 change scores showed that the absolute size of the change score became smaller when the posttest No. 2-pretest score was used. The significant factor to note was that even after two months from the completion of the course, the achievement change measured from the pretest was still significant for both classes regardless of the change score method used. These results indicated that the probability is 5 out of 100 that the difference between mean achievement scores on the pretest and posttest No. 2 will have occurred by chance.

CHAPTER VI  
SUMMARY AND CONCLUSIONS

The purpose of this study, conducted for the College of Humanities, Arts and Social Sciences at USU, was to investigate the achievement that occurs in selected introductory social science courses offered by that college. Two social science introductory courses were selected from the list of courses offered by the College of Humanities, Arts, and Social Science. The students in one course, Political Science 111, were tested with a nationally standardized political science test. The students in the other course, Sociology 101, were tested with a teacher-made test constructed by the professor in that course. The students in both classes were pretested at the beginning of the first session of summer school, 1972, and again at the end of the first session of summer school, 1972.

Two other classes were chosen which were offered by colleges at USU other than the College of Humanities, Arts and Social Sciences. These two courses acted as a control measure with which to compare the achievement of the students taking the social science courses offered by the College of Humanities, Arts, and Social Sciences. The students in one course, Economics 200, were tested with a nationally standardized test, the Test of Understanding in College Economics. The students in the other course, Geology 101, were given a teacher-made test constructed by the professor of that course. The

students in the Geology 101 course were tested at the beginning and end of the course during the first session of summer school, 1972. The students in Economics 200 were tested at the beginning and end of the course, spring term, 1972.

Two courses using the teacher-made exams, Geology 101 and Sociology 101, were tested a third time (posttest No. 2) two months after the completion of the course. The additional testing was performed to measure the long term retention in these two courses and the amount of teacher quing of students to the test questions.

The study tested five hypotheses:

1. There is no difference between the pretest and posttest scores of students taking selected introductory social science courses offered by the College of Humanities, Arts, and Social Sciences at USU.
2. There is no difference in achievement between students taking introductory social science courses offered by the College of Humanities, Arts, and Social Sciences and the achievement of students taking introductory courses offered by the other colleges at USU.
3. There is no difference in achievement measured by standardized tests and achievement measured by criterion-referenced tests.
4. There is no interaction between the mode of testing and the college offering the course.
5. There is no difference between achievement of students tested with criterion-referenced tests at the beginning of a course and the achievement of

those same students tested with the same criterion-referenced test two months after the courses termination.

The data permitted a rejection of all five hypotheses. First, the data showed that students in selected introductory social science courses offered by the College of Humanities, Arts, and Social Sciences did gain in knowledge. Second, the students in the selected introductory social science courses offered by the College of Humanities, Arts, and Social Sciences did not show as much knowledge gain as those students tested in selected introductory courses offered by other colleges at USU. Third, students who were in courses measured with the criterion-referenced tests scored higher than those students who were in courses measured with the standardized tests. Fourth, there was an interaction between the method of testing and the college offering the course. Last, the students in geology and sociology courses tested with the criterion-referenced test two months after the completion of the course showed achievement gains.

### Conclusions

The data indicates that students in the introductory social science classes tested do acquire knowledge by taking the courses. The knowledge gained is more than might be expected by chance alone. The general introductory courses in political science and sociology are not duplications of high school political science and sociology courses. These findings can be assumed for the other five introductory social science courses offered by the College of Humanities, Arts, and Social Sciences too, due to the similarity of the courses.

This can be further substantiated by the positive achievement gains made by students taking similar introductory courses offered by colleges other than the College of Humanities, Arts, and Social Sciences. The reliability of the above conclusions is largely substantiated by the similarity in achievement obtained from the students' performance on the tests given two months after the completion of the course. The students in the two courses tested continued to show positive achievement gains. Though the gains were not as large as the gains at the completion of the courses, they were still more than would be expected by chance alone. Thus even with the use of the criterion-referenced tests, the gains in achievement were evident.

Though it was found that (1) the students in introductory courses offered by other colleges at USU showed larger knowledge gains than the students in the introductory social science courses, (2) the students in courses measured by criterion-referenced tests scored higher than the students in courses measured with standardized tests, and (3) the interaction between the method of testing and the college offering the course existed--this researcher would advocate further investigation before these findings were accepted due to the large gain shown in Geology 101 class. This major deviation in gain score from the other three courses appears to bias the use of the ANOVA statistic and casts doubt as to the accuracy of the conclusions drawn from the findings.

### Discussion

In discussing the results and conclusions of this research, a review of the internal and external invalidity factors related to this research design (Campbell and Stanley, 1963) is deemed important. Without such a review it would be impossible to determine if change had occurred and if the change could be generalized to the population considered. Before the internal and external invalidity factors related to this design are examined, however, the concept of measuring change (rather than concentrating on final scores adjusted by pretest scores) seems warranted.

Carl Berierter (Lord, 1963) describes three persisting dilemmas in the measurement of change. The three are entitled (1) the overcorrection-undercorrection dilemma, (2) the physicalism-subjectivism dilemma, and (3) the unreliability in validity dilemma. Only the latter dilemma seems to pose a serious question to the use of change scores in this research.

The unreliability-invalidity dilemma "stems from the fact that high reliability of change scores usually requires low test-retest correlations, with the implications that in such a case the test may not measure the same thing on the two occasions and the change scores will therefore be meaningless" (Berierter, 1963, p. 20).

The dilemma becomes more clear when noticing the relationship of the variables in the reliability formula (below) for the difference in scores on tests  $x$  and  $y$ . In this equation  $r_{x-y}$  is the reliability of the difference between  $x$  and  $y$ ,  $r_{xy}$  is the intercorrelation between the two tests, and  $\bar{r}$  is one-half the sum



of the reliabilities of tests x and y (Gulliksen, 1950):

$$r_{x-y} = \frac{\bar{r} - r_{xy}}{1 - r_{xy}}$$

A critical review of the reliability and correlation ratios will show that the dilemma is obvious in this study also.

Table 5. Correlation and reliability data relating to the unreliability-  
invalidity dilemma

COURSES	Correlations Pretest/ Posttest No. 1	Reliability (KR 20) Average of Pretest/ Posttest No. 1	Reliability of the mean difference of Pretest/Posttest No.1
POLITICAL SCIENCE 111	.66	.58	.24
SOCIOLOGY 101	.70	.74	.13
GEOLOGY 101	.60*	.65**	.12**
ECONOMICS 200	.47	***	***
*	Correlation of pretest/posttest No. 2		
**	Reliabilities of pretest/posttest No. 2		
***	Test data not available with which to calculate reliabilities		

Berietter approaches the dilemma by (1) providing a technique to reduce the correlation while maintaining a relatively high reliability, (2) suggesting that attributes be conceptualized independently of any instruments supposed to measure it (Jordan, 1960), and (3) suggesting the use of change scores only

indicators of change without specifying the variable changed or the variable causing the change.

Even with Berieter's discussion of the dilemma and his proposals to the remedy the dilemma many researchers (Buehe, 1969; Glaser, 1968; Roberts, 1969; Forehand, 1971; Dyer, 1970; Milton and Shoben, 1968; Westland, 1971; Cronbach, 1963; Dyer, 1966; Manning, 1969; Ford, 1971) still use change scores with which to measure variables in the academic setting. Of course, the fact that these researchers used the change score method of quantifying data does not make them any less suspect to the dilemma posed.

The major thrust of this research was to determine if change occurred in knowledge acquisition and in what direction that change occurred. In this respect, the approach agreed with one of Berieter's approaches (No. 3) to the dilemma of using change scores. No attempt was made to identify the independent variables functionally related to the change other than grouping them under the title of the "effects of college introductory classes." The specificity of the tests to a large degree limited the measurement of change to the effects of the introduction classes and ruled out the acquisition of more general information which could be gathered outside the classroom situation. As to the labeling of the change as "achievement" simply because the same objective test was used requiring factual knowledge responses, the Berieter argument poses some questions. It is possible that the score on the first test represented differences mainly in prior acquaintance with the course and that the final test score represented learning ability in the subject.

Thus, with one qualification, the labeling of change as "achievement" rather than verbal aptitude or some other factor, the dilemma of unreliability and invalidity can be resolved in this administrative study because of the impressive past precedent of other researchers' use of the technique and the special purposes to which change scores are used in this particular study, primarily to measure change and its direction as measured by tests specific to the course.

As previously mentioned, designs are distinguished in terms of internal validity.

Experiments are more or less internally valid depending on how well extraneous variables have been controlled by the researcher. If extraneous variables are not controlled in the experiment, we cannot know whether observed changes in the experimental group are due to the experimental treatment or to an extraneous variable. (Borg and Gall, 1971, p. 365)

In this research the dependent variable, though in question because of the comments of Berierter, is "achievement" whereas the independent variable is the introductory college course (specifically the introductory social science courses offered by the College of Humanities, Arts and Social Sciences). The eight factors relevant to internal validity will be discussed in relation to the possible influence of other variables besides the introductory college course.

A factor important to assessing the measured change is the effects of regression. This factor occurs most in test-retest situations with initially low and high performance groups. It is doubted that the extremely high and low scores represent "true" scores and that on the second testing of the students that these groups will regress toward the mean (Borg and Gall, 1971). Though

the effect is usually discussed in terms of the error of measurement, "it is more generally a function of the degree of correlation; the lower the correlation the greater the regression toward the mean" (Campbell and Stanley, 1963, p. 11). In relation to these two positions the regression effect does not appear to have much bearing in this study as (1) an average cross section of college students were tested, and (2) the correlations between the tests (below) were not exceptionally low.

Table 6. Correlations between tests

COURSES	Pretest/ Posttest No. 1	Pretest/ Posttest No. 2	Posttest No. 1/ Posttest No. 2
Economics 200	.47		
Political Science 111	.66		
Geology 101	.51	.60	.72
Sociology 101	.70	.70	.96

A second factor, history, or the specific events occurring between the first and second measurements in addition to the experimental variable can be dismissed as the time interval between the tests was too short (approximately five weeks) for any other variable to have much effect. Only in the case of political science did the increased news coverage of the 1972 presidential

campaign seem a powerful external variable which could have influenced achievement. If history had been a critical variable the delayed testing of geology and sociology students would have shown an increase from the scores of the posttest No. 1 but this did not occur.

Maturation, a third factor, does not seem to play a part in this study as the time interval between the tests again was too short. All pretests and posttests (No. 1) were given in the normal class period and required no extra student time compared to the regular class activities. Maturation would be more likely to occur in a study taking several years to determine change. The maturation factor did occur to some degree in the way the instructor presented the tests to the students. Some professors used much more encouragement during the final tests than they did in the pretest. This could account for the higher performance on the final test, a maturation factor that was not controlled in the procedure.

The factors of instrumentation and mortality appear to have little impact on the data. In regard to instrumentation there were no changes in the calibrations, observers, tests, or scorers while obtaining measurements. All tests were objective in nature thus voiding any subjectivity to the grading. In regard to mortality, all class statistics were calculated using only those students who had taken both pretest and posttest. In some cases this meant some students had to be removed from the calculations but this allowed for comparability of such factors as correlation coefficients, reliability measurements, and mean scores from one test situation to another within a subject

area. The intact group was used in preference to the random selection of volunteers not only because of the benefits it allowed for generalizability but because it was expected that there would be less mortality using intact than random volunteers (Beanblossom, 1969).

The effects of pretests on the performance of a posttest was uncontrolled in this experiment. The entire class was given the exam rather than randomly assigning a pretest to only half the students and giving the final test to all the students. Thus, the pretest effects cannot be separated from the change scores.

Selection has always been a crucial element when considering internal validity. Selection bias occurred in using intact groups of students enrolled for the courses. The composition of students in these intact groups was not as representative of the students taking introductory courses at USU previous to 1972-1973 because of the new general education requirements making certain core classes not mandatory. Students at USU now are allowed to select their introductory courses in certain subjects thus bypassing the traditional core course areas. The result is a more specialized, motivated student in certain areas. This fact was clearly the case as the professors in political science, sociology and geology all indicated considerable reductions in enrollments in their classes over previous summer sessions. Though this fact supports the validity (generalizability) it hinders the internal validity by selecting a group of dedicated, motivated students. Comparability between classes may be impossible, too, as different subject areas may attract a different caliber of student as so often happens in the upper division courses. No control was established

for this variability due to differences in abilities of students taking each course and may have a direct influence on achievement.

Several factors relevant to internal validity can interact to provide threats to the effects of the independent variable. The interaction of testing and selection introduces the concept of ceiling effect.

The ceiling effect means that the range of difficulty of the test items is limited; therefore, the test does not measure the entire range of achievement possible on the dimension being measured. (Borg and Gall, 1971, p. 407)

A rule of thumb to test for ceiling effect (Diederick, 1956) is to determine if the highest final score is 90 per cent correct and the average goes up only from 50 to 60 per cent correct. A chart is provided below to indicate how the four evaluation instruments in this study compared to such a rule.

Table 7. Index values for ceiling effect evaluation

TEST	Highest correct on final test per cent of total possible	Pretest $\bar{X}$ correct as per cent possible	Pretest No. 1 correct as per cent of possible	Posttest No. 2 correct as per cent of possible
POLITICAL SCIENCE 111	95%	70%	81%	
SOCIOLOGY 101	84%	55%	70%	67%
GEOLOGY 101	86%	24%	68%	53%
ECONOMICS 200	64%	24%	41%	

Only in the case of the political science test, Principles of Democracy Test, Form A, does the ceiling effect seem to be of any consequence. This can be directly attributable to the tests intended use at grades 9-12 and not at the college level. The use of the percentage gain score to an extent will weight gains in the upper limits more heavily and will provide some adjustment for the ceiling effect on this test. It should be mentioned in relation to percentage gain that although the percentage gain scores do not differ drastically from the raw gain scores, the percentage gain scores provided a valuable method for comparing the results of the different classes by normalizing the distribution of scores to some extent and adjusting for the ceiling effect.

In conclusion, there are several factors of internal validity which were not controlled and may account for some of the variability in the achievement changes measured. These factors include maturation, testing, and selection.

External validity, or representativeness, is the other important factor that must be discussed in order to determine the carryover which can be expected between the experiment and the related field situations.

As previously stated, of the 25 introductory social science courses offered by these three departments, seven can actually be classified as basic pre-requisite first term introductory social science courses. Therefore, the two classes chosen from these seven represent slightly more than 25 per cent of the courses that could be tested and consequently a rather large proportion of the population being studied.



Other factors enter into the discussion of the representativeness of the sample. These may also add or detract from the sample's generalizability to the seven basic introduction courses in social science and even to the 25 lower division social science courses offered by the College of Humanities, Arts, and Social Sciences. For instance, the courses which succeed the initial introductory course and which are not part of a sequence of course (i. e. numbered sequentially 101, 102, 103) seem to develop more complex ideas than the initial course. The increased complexity may or may not result in larger or smaller gains/losses in achievement. The results of this study, thus, do not specify what the achievement might be in 100 or 200 level non-sequenced introductory social science courses following the initial sequenced introductory courses.

Another factor affecting the generalizability from the two course sample used in this study to the other introductory social science courses is that the testing occurred in the 5-week summer session. This was a shortened time period as compared to the normal 10-12 week long term. The lack of comparability thus results in two areas: (1) the potential incomparability between students in the summer session and the 9-12 week term in the fall, winter, or spring terms, and (2) the potential incomparability between the learning that occurs in the shortened term and the regular length term.

In relation to this first area of concern, the students attending the summer session and enrolled in the introductory courses appear to come from the same population as those students attending USU during the fall, winter,

and spring terms. Of the 12 students taking Sociology 101 and tested, three left town at the end of the summer quarter, while the other nine students returned to the campus for further study in the fall, 1973. Of the 13 students taking Geology 101, all students were present for the fall quarter though two refused to take the test in the fall. These two examples point out that the students taking the introductory courses are not the graduate level population which is typically associated with the transient group of students which commonly frequent the summer school sessions.

No data was collected on which to compare the subjects of this particular study on the basis of IQ, GPA, major or sex, with students who attended the 9 month session of school (fall, winter and spring terms). There are some speculations that summer school undergraduates differ in their motivation and basic ability. The summer session might include those highly motivated students who want to get through college in three years or those students who were unable to pass, complete or take the course in the normal 9 month session. It seems to this researcher that the performance of the above named groups would offset (counter balance) each other when using the overall class mean achievement score.

A second factor related to a shortened term is the effect it might have on achievement. Mangham (1971)

. . . revealed that neither the high ability group nor the low ability group showed a significant difference in its academic performance under the semester and quarter systems. . . there was evidence that the majority of students in each group preferred the quarter system. It appears that they favored the quarter system because they liked the increased pace and greater pressure which

they experienced under this system. Apparently it seemed less drawn out and less boring. (p. 3296)

Another study (Monk and Stalling , 1971) reports :

. . . that moderate variations in test frequency do not significantly affect the college students' learning. . . the student does not need many cues to discern the type of learning required by a course, and that if his learning is being measured by "hourly" and final examinations, frequent quizzes are of little importance to his learning once he knows what will be demanded of him. (p. 187)

These two studies indicate that the shortened term may have little negative effect on the performance of the students in college, but quite the contrary, it may have a positive effect.

Two additional comments in relation to the interaction of selection and the experimental treatment as a factor of internal validity are needed. First, the discussion thus far should make it clear that the achievement gains in the two social science courses sampled cannot be used to infer achievement in non-social science courses (i.e. Introduction to Novel, Enjoying Music, or Introduction to Short Story) offered by the College of Humanities, Arts, and Social Science. Second, the two social science courses may serve as an adequate sample to indicate achievement in social science courses offered by departments in colleges other than the College of Humanities, Arts, and Social Sciences (i.e. Economics and Psychology). This assumption is somewhat substantiated by the achievement results in Economics 200 shown in this research. Students in economics showed achievement gains after one term of study. Even though the gains were not as great as the gains in the two social science courses offered by the College of Humanities, Arts, and Social Sciences, they were significant.

A second factor of external validity is the interaction of testing and the dependent variable. This factor was uncontrolled in this experiment. Typically a pretest would not be given to students taking a course at college though there are exceptions to this rule. Considering the former to be the case in most instances, the effects due to the pretest make generalizability to the normal classroom situation impossible.

A third factor, is the reactive effects of experimental arrangements. The built in biases of using intact groups and professors who were a bit skeptical of the use of the data provided some uncontrolled effects of the experimental treatment. Several professors prefaced the pretest with derision yet they spent considerable time encouraging the students to do well on the post-test. The students knew that they were part of an experiment that was in addition to their normal requirements. The fact that no remuneration (i. e. money or grades) was given for student participation and that only 6 of 48 students in geology, sociology, and political science dropped out for one reason or the other, appears to evidence that the students were seriously concerned with the experimentation. This situation approximates the type of participation one might expect in a normal testing situation. If extra credit for testing would have been given, the situation would have been a closer approximation to the normal classroom.

The fourth factor of external validity, the multiple treatment interference, occurs only in experimental designs exposing the subject to more than one experimental treatment, and thus is not relevant to this study.

In summary, the findings of this research can be accepted within the framework of the limitations mentioned. The limitations of using change scores is not as serious a limitation in this study as it might be in other studies using change scores. Using change scores to determine change and the direction of the change is the proper use of change scores. Though labeling the change "achievement" can be questioned, the specificity of the test in relation to the course content and the factual, objective nature of the tests seemed to obviate the classification of test behavior as "achievement."

The primary limitation in terms of evaluating change appeared to be the failure to control for the effects of the pretest--using the same test for pretest and posttest. Other than this one factor, the use of percentage gain scores and the specificity of the tests nullified the effects of the other factors of internal invalidity pretest-posttest design to measure change.

The amount of generality of the findings seems limited to the achievement of students in the seven introductory social science classes offered by the College of Humanities, Arts, and Social Sciences. Generality of achievement gains to other introductory courses offered by colleges other than the College of Humanities, Arts, and Social Sciences seems warranted though that objective was not the primary purpose of this research. The generality to students taking these introductory courses in other terms seems accurate despite the use of summer session students and the shortened term (5 weeks). The use of pretest and the variability in the motivational techniques of the teachers in preparing the students to take the tests appear to be the two limiting factors to generality in the college classroom situation. Both could

have been controlled and will be discussed in the recommendation section below.

Within the limits mentioned above, this researcher concludes that the data of this study suggests that the introductory social science courses offered by the College of Humanities, Arts, and Social Science do increase the achievement of students taking such courses. Recommendations for further studies are presented below.

#### Recommendations

1. The same study could be replicated using a test composed of a sampling of questions from each of the four tests in this study.
2. The same study could be replicated using a design which allows the pretesting effect to be isolated.
3. The same study could be replicated using students attending these introductory social science courses during the normal 9-week session.
4. The same study could be replicated using social science courses offered by other colleges with which to compare the social science courses offered by the College of Humanities, Arts, and Social Sciences.

## LITERATURE CITED

- Astin, Alexander W. 1968. Undergraduate achievement and institutional excellence. *Science* 161(3824):661-668.
- \_\_\_\_\_. 1971. Is higher education really better education? American Council of Education, Washington, D.C.
- Barton, Allen H. 1959. Studying the effects of college education. The Edward W. Hazen Foundation, New Haven, Connecticut.
- Bateson, Russel B. 1949. A study of the 1947 American Council of Education Psychological Examination and its usefulness in predicting the grades of Utah State Agricultural College freshman. M.S. Thesis, Utah State University, Logan, Utah.
- Beanblossom, Gary F. 1969. The use of CLEP scores in evaluating liberal arts curriculum. University of Washington, Seattle, Washington.
- Birch, Anthony. 1970. Institutional research is feedback link between administration and management. *College and University Business* 49(5):28-31.
- Bloom, Benjamin S. 1969. Some theoretical issues relating to educational evaluation. National Society for the Study of Education. Yearbook 68(2):26-50.
- Borg, Walter R. and Meredith D. Gall. 1971. Educational research an introduction. David McKay Company, Inc., New York.
- Buethel, Chris. 1969. Testing, testing...1...2...3. *The Clearing House* 43(9):536-538.
- Burman, Ronald S. and Flaherty, Lorraine. 1968. The 1967 freshman english program at Wisconsin State University at LaCrosse. Wisconsin State University, LaCrosse, Wisconsin.
- Campbell, Donald T. and Julian Stanley. 1963. Experimental and quasi-experimental designs for research. Rand McNally and Company, Chicago.

- Christiansen, Mark. 1972. USU study of the validity of awarding credit by examination in English composition. Unpublished Ph. D Dissertation, Utah State University, Logan, Utah.
- Cronbach, Lee J. 1963. Course improvement through evaluation. *Teacher College Record* 64(8), 672-683.
- \_\_\_\_\_ and Lita Furby. 1970. How we should measure "change"--or should we? *Psychological Bulletin* 74(1):68-80.
- Day, James Francis. 1939. Some factors related to student achievement of entering freshman at the Utah State Agricultural College. M.S. Thesis, Utah State University, Logan, Utah.
- Dressel, Paul L. 1968. College and university curriculum. McCutchan Publishing Corporation, Berkeley, California.
- \_\_\_\_\_ and associates. 1971. Institutional research in the university: a handbook. Jossey-Bass, San Francisco.
- \_\_\_\_\_ and John E. Dietrich. 1967. Departmental review and self-study. *The Journal of Higher Education* 38(1):25-37.
- \_\_\_\_\_ and Sally B. Pratt. 1971. The world of higher education. Jossey-Bass, Inc., San Francisco.
- Dyer, Henry. 1970. Can we measure the performance of educational systems? *The Bulletin of the National Association of Secondary School Principals* 54(346):96-105.
- \_\_\_\_\_, 1966. Can institutional research lead to a science of institutions? *Educational Record* 47(4):452-466.
- Emery, David E. and Thomas P. 1972. Computer gaming and learning in an introductory economics course. *The Journal of Economic Education* 3(2):77-85.
- Erickson, Barney L. 1969. Effects of a college mathematics sequence upon the attitudes and achievement in mathematics of prospective elementary school teachers. M.Ed. Plan B report, Utah State University, Logan, Utah.
- Ford, Patrick J. and Mayhew, Lewis B. 1971. Changing the curriculum. Jossey-Bass Inc., Publishers, San Francisco.
- Forehand, Carlisle A. 1971. An evaluation system for curriculum innovation. *Teachers College Record* 72(4):577-591.



- Giles, LaVerl C. 1965. Factors related to first year college success in a selected group of scholarship recipients. M.S. Thesis, Utah State University, Logan, Utah.
- Glaser, Robert. 1968. Ten untenable assumptions of college instruction. *Educational Record* 49(2):154-159.
- Grover, Burton L. 1970. Course evaluation on the local level. *Educational Leadership* 27(6):591-596.
- Gulliksen, Harold. 1950. *Theory of mental tests*. John Wiley and Sons, Inc., New York.
- Harris, J. and E. Booth. 1969. An analysis of the performing of Georgia College freshman and sophomores on the College Level Examinations. Institute of Higher Education. University of Georgia, Athens, Georgia.
- Harris, John and Patricia Hurst. 1972. Does college assure academic growth? *Change Magazine* 4(1):8+.
- Hendricks, James Vance. 1967. The Iowa Tests of Educational Development as predictors of academic success at Utah State University. M.S. Thesis, Utah State University, Logan, Utah.
- Johnson, Walter Clarence. 1951. High school grades and college aptitude tests as indices of college achievement and continuation at Utah State Agricultural College. M.S. Thesis, Utah State University, Logan, Utah.
- Kagan, Jerome and Earnest Havemann. 1968. *Psychology an introduction*. Harcourt, Bruce and World, Inc., New York.
- LaPray, Anthony. 1962. Predicting success of general registration students. M.S. Thesis, Utah State University, Logan, Utah.
- Lareau, Edward Henry, Jr. 1971. Comparison of two methods of teaching expository composition and evaluation of a testing instrument. Ed.D. dissertation, LeHigh University, Bethlehem, Pennsylvania, 119 p. (Original not seen; abstracted in *Dissertation Abstracts International* 32(5):2437-2438. November 1971)
- Lauer, William C. 1972. Evaluating effectiveness of using business education department prepared videotapes in the teaching of intermediate type-writing at Utah State University. M.S. Thesis, Utah State University, Logan, Utah.

- Lenning, Oscar, Leo A. Munday and James E. Maxey. 1968. Student educational growth during the first two years of college. American Educational Research Association, Washington, D.C.
- Lessinger, Leon M. 1970. Accountability in public education. In Invitational Conference Testing Problems. Educational Testing Service, Berkeley, California.
- Livesay, Thayne Miller. 1939. Does test intelligence increase at the college level? *Journal of Educational Psychology* 30(1):63-68.
- Lord, Frederic M. 1956. The measurement of growth. *Educational and Psychological Measurement* 16(4):421-437.
- \_\_\_\_\_. 1963. Elementary models for measuring change. In Harris, C. W. 1963. Problems in measuring change. The University of Wisconsin Press, Madison, Wisconsin.
- Louise, Sister M. Florence. 1947. Mental growth and development at the college level. *Journal of Educational Psychology* 38(2):65-83.
- Lunt, David A. 1956. The achievement of college freshman students in selected areas of political science in relation to courses taken in high school, college grade-point average, and general mental ability. M.S. Thesis, Utah State University, Logan, Utah.
- Lyon, Kenneth S. and Glen F. Marston. 1972. Learning and attitude change in students subjected to a national income simulation game. Unpublished paper, Utah State University, Logan, Utah.
- Mangham, Clarence Walter. 1970. A study of the relationship between student academic ability and student performance under the quarter and semester systems in a junior college. Ph.D. dissertation, University of California, Berkeley, California 204 p. (Original not seen; abstracted in *Dissertation Abstracts International* 31(7):3296. January 1971.)
- Manning, Winton H. 1969. The functions and uses of educational measurement. In Invitational Conference Testing Problems. Educational Testing Service, Berkeley, California.
- McConnell, T. R. 1934. Changes in scores on the Psychological Examination of the American Council of Education from freshman to senior year. *Journal of Educational Psychology* 25(1):66-69.
- \_\_\_\_\_. 1971. Accountability and autonomy. *Journal of Higher Education* 42(6):446-466.

- Menlo, Allen and M. Clemens Johnson. 1971. The use of percentage gain as a means toward the assessment of individual achievement. *California Journal of Educational Research* 22(5):193-201.
- Milton, Ohmer and Edward Joseph Shoben. 1968. *Learning and the professors*. Ohio University Press, Athens, Ohio.
- Monk, Janic J. and William M. Stallings. 1971. Another look at the relationship between frequency of testing and learning. *Science Education* 55(2):183-188.
- Nelson, Dennis Lee. 1970. The effect of specially state instructional objectives on the achievement of collegiate undergraduate economics students. Ph.D. dissertation, University of Minnesota, Minneapolis, Minnesota, 131 p. (Original not seen; abstracted from *Dissertation Abstracts International* 32(1):99. July 1971)
- Norton, Daniel P. 1967. The potency of educational treatments. In Grover, Burton L. 1970. Course evaluation on the local level. *Educational Leadership* 27(6):591-596.
- Patalino, Marianne. 1970. Rationale and use of content-relevant achievement tests for the evaluation of instructional programs. University of California, Los Angeles, California.
- Pickett, Louis M. 1958. The general aptitude test battery as a predictor of college success. M.S. Thesis, Utah State University, Logan, Utah.
- Postman, Leo. 1968. Comments on professor Gagne's paper entitled "Instructional variables and learning outcomes." University of California, Los Angeles, California.
- Proceedings of the 1969 invitational conference on testing problems. 1969. Educational testing service, Princeton, New Jersey.
- Putz, Joan M. 1970. When the teacher stops teaching-an experiment with freshman English. *College English* 32(1):50-57.
- Qualitative and quantitative performance and achievement on higher education report of Task Force IV. 1971. Connecticut Commission for Higher Education. Hartford, Connecticut.
- Roberts, Dennis M. 1969. Measurement of academic learning. *Educational Forum* 33(2):207-211.

- Rock, Donald A., J. A. Centra and R. L. Linn. 1970. Relationships between college characteristics and student achievement. *American Educational Research Journal* 7(1):109-121.
- \_\_\_\_\_ and \_\_\_\_\_. 1971. College environment and student academic achievement. *American Educational Research Journal* 8(4):623-634.
- \_\_\_\_\_, Leonard L. Baird, and Robert L. Linn. 1972. Interaction between college effects and student aptitudes. *American Educational Research Journal* 9(1):149-161.
- Sharon, Amiel T. 1971. Measurement of college achievement by the College Level Examination Program. *Educational and Psychological Measurement* 31(2):477-484.
- Shuey, Audrey Mary. 1948. Improvement in scores on the American Council Psychological Examination from freshman to senior year. *Journal of Educational Psychology* 39(7):417-426.
- Silvey, H. M. 1951. Changes in test scores after two years in college. *Educational and Psychological Measurement* 11(3):494-502.
- Tovrea, Verner Basil. 1970. A comparison of the effects of two methods of teaching on achievement of selected goals and objectives for general biology in a community college. Ed.D. dissertation, University of Northern Colorado, Greeley, Colorado, 193 p. (Original not seen; abstracted in *Dissertation Abstracts International* 31(12):6433-6434. June 1971)
- Tyrrel, Jerry H. 1958. Predicting graduation probabilities of Utah State University freshman engineering students. M.S. Thesis, Utah State University, Logan, Utah.
- Van Pelt, W. Austin. 1971. Instructional methodology and student subcultures, a study of achievement in Introduction to Sociology courses. Ph.D. dissertation, University of Denver, Denver, Colorado, 120 p. (Original not seen; abstracted in *Dissertation Abstracts International* 31(7):3307. January 1971.)
- Westland, Gordon. 1971. Selection, performance, and assessment. *Universities Quarterly* 25(3):344-352.
- White, Charles C. 1969. The use of programmed tests for remedial mathematics instruction in college. Ph.D. dissertation, Utah State University, Logan, Utah.

Willems, Alvin E. 1970. A comparison of two forms of cognitive questioning applied to the learning of selected technological information and skills. Ed.D dissertation, Utah State University, Logan, Utah.

## APPENDIX

USU STUDY OF THE VALIDITY OF AWARDING CREDIT BY  
EXAMINATION IN ENGLISH COMPOSITION

A random selection of 24 students from each of the five groups designated below took part in this study:

1. Adv. Place. 1971-72 USU freshman students who received a score of 3, 4, or 5 on the Advanced Placement test in English Composition.
2. CLEP Cr. 1971-72 USU freshmen students who received credit for freshmen English by scoring 450 or more points on the CLEP General Examination in English Composition.
3. CLEP Waiver 1971-72 USU freshmen students who waived the Freshmen English requirement by scoring from 390-449 points on the CLEP General Examination in English Composition.
4. A-B (F.E.) 1970-71 USU freshmen students who received credit for three quarters of Freshmen English, and whose Freshmen English grade-point-average was from a "B" to an "A".
5. C-D (F.E.) 1970-71 USU freshmen students who received credit for three quarters of Freshmen English, and whose Freshmen English grade-point-average was from a "D" to a "C".

All five groups of students took the CLEP subject examination in English Composition, which consists of a 90 minute objective part and a 90 minute essay part.

The objective part of the test is machine scored and designed so that an individual can receive a score of from 20 to 80 points.

Each of the essays was graded by three members of the USU English Department. An individual could score from 1 to 12 points ("F" to "A") on one essay. In this analysis, the grades of the three essay graders were summed giving each individual one essay score which could range from 3 to 36 points.

The figures below show the mean and standard deviation (S. D.) each group received on each part of the examination:

	OBJECTIVE			ESSAY	
	MEAN	S. D.		MEAN	S. D.
Adv. Place.	55.38	7.35	Adv. Place.	17.17	6.74
CLEP Cr.	51.83	8.83	CLEP Cr.	13.88	5.90
A-B (F. E.)	48.12	8.11	A-B (F. E.)	13.42	5.99
CLEP Waiver	41.62	6.36	C-D (F. E.)	11.25	4.23
C-D (F. E.)	39.92	5.93	CLEP Waiver	10.88	5.09

Differences between all of the group means are shown below. Those which are significant are marked with asterisks and are also specified individually. One asterisk indicates that we can say with a probability of .95 that the means of the populations from which the samples were drawn are different. Two asterisks indicates that we can say with a probability of .99 that the means of the populations from which the samples were drawn are different.

	$\bar{X}$	$\bar{X}_i - \bar{X}_5$	OBJECTIVE		
			$\bar{X}_i - \bar{X}_4$	$\bar{X}_i - \bar{X}_3$	$\bar{X}_i - \bar{X}_2$
1. Adv. Place.	55.38	15.46**	13.76**	7.26**	3.55
2. CLEP Cr.	51.83	11.91**	10.21**	3.71	
3. A-B (F. E.)	48.12	8.20**	6.50*		
4. CLEP WAIVER	41.62	1.70			
5. C-D (F. E.)	39.92				

  

	$\bar{X}$	$\bar{X}_i - \bar{X}_5$	ESSAY		
			$\bar{X}_i - \bar{X}_4$	$\bar{X}_i - \bar{X}_3$	$\bar{X}_i - \bar{X}_2$
1. Adv. Place.	17.17	6.29**	5.92**	3.75	3.29
2. CLEP Cr.	13.88	3.00	2.63	.46	
3. A-B (F. E.)	13.42	2.34	2.17		
4. C-D (F. E.)	11.25	.37			
5. CLEP WAIVER	10.88				



The following comparisons were significant:

OBJECTIVE		ESSAY	
Adv. Place. vs C-D(F. E.)	15.46**	Adv. Place. vs CLEP Waiver	6.29**
Adv. Place. vs CLEP Waiver	13.76**	Adv. Place. vs C-D (F. E.)	5.92**
Adv. Place. vs A-B(F. E.)	7.26**	* .95 probability (Tab. Value = 4.59)	
CLEP Cr. vs C-D (F. E.)	11.91**	** .99 probability (Tab. Value = 5.51)	
A-B (F. E.) vs C-D (F. E.)	8.20**	The correlation (r) between the objective and essay parts of the test was .56.	
A-B (F. E.) vs CLEP Waiver	6.50*		
* .95 probability (Tab. Value =5.84)		No significant difference on others measured between groups.	
** .99 probability (Tab. Value =7.02)			

## VITA

James Collins Lewis

Candidate for the Degree of

Master of Science

Thesis: Achievement of Students in Introductory Social Science Courses offered by the College of Humanities, Arts, and Social Sciences at Utah State University

Major Field: Counseling Psychology

## Biographical Information:

Personal Data: Born at Ashland, Oregon, January 21, 1943; son of Herb and Nell Lewis.

Education: Attended elementary school in Ashland, Oregon; graduated from Ashland Senior High School in 1961; received the Bachelor of Science degree from the University of Oregon with a major in Business Administration in 1965; completed the requirements for Master of Science degree, specializing in Counseling Psychology, at Utah State University in 1973.

Professional Experience: 1972-1973, resident counselor for Hillside School in Logan, Utah.