R E P O R T   R E S U M E S

ED 018 400   24   SE 004 547
GUIDELINES FOR THE DEVELOPMENT OF A COURSE FOR SECONDARY SCHOOLS EMPHASIZING THE INTERACTION OF SCIENCE AND THE CULTURE OF MAN. FINAL REPORT.
BY- YAGER, ROBERT E.   COSSMAN, GEORGE W.
IOWA UNIV., IOWA CITY
REPORT NUMBER BR-6-8445   PUB DATE   JAN 66
CONTRACT OEC-3-7-068445-0374
EDRS PRICE MF-$0.75 HC-$6.44   159P.

DESCRIPTORS- *ACADEMIC ACHIEVEMENT, *CURRICULUM DEVELOPMENT, *EVALUATION, *SCIENTIFIC LITERACY, *SOCIOECONOMIC INFLUENCES, *SOCIAL STUDIES, CURRICULUM, CRITICAL THINKING, COURSE DESCRIPTIONS, GRADE 11, GRADE 12, INSTRUCTIONAL MATERIALS, SECONDARY SCHOOL SCIENCE, VALUES, FACTS ABOUT SCIENCE, TEST ON UNDERSTANDING SCIENCE, WATSON-GLASER CRITICAL THINKING APPRAISAL, STANFORD ADVANCED SCIENCE ACHIEVEMENT TEST, STUDY OF VALUES, SCIENCE OPINION SURVEY,

REPORTED IS THE EVALUATION OF AN EXPERIMENTAL INTERDISCIPLINARY COURSE ENTITLED "SCIENCE AND CULTURE" WHICH IS BEING DEVELOPED AT THE UNIVERSITY OF IOWA LABORATORY SCHOOL. IT IS DIRECTED AT BRIDGING EXISTING GAPS IN INTEGRATING THE RELATIONSHIPS BETWEEN SCIENCE AND TECHNOLOGY AND SOCIAL STUDIES IN THE SECONDARY SCHOOL CURRICULUM. THE OBJECTIVE OF THE STUDY WAS TO DETERMINE TO WHAT EXTENT, IF ANY, SEVERAL OF THE BASIC OBJECTIVES AROUND WHICH THE EXPERIMENTAL COURSE WAS STRUCTURED WERE ACHIEVED. SEVEN SPECIFIC HYPOTHESES, STATED IN THE NULL FORM, WERE INVESTIGATED. THE EXPERIMENTAL GROUP CONSISTED OF A HETEROGENEOUS GROUP OF 21 ELEVENTH AND TWELFTH GRADE STUDENTS IN THE UNIVERSITY OF IOWA LABORATORY SCHOOL. THE CONTROL GROUP WAS SELECTED FROM THE SAME SCHOOL AND SAME GRADE LEVELS ON A MATCHED PAIR BASIS. CRITERION TESTS WERE USED IN EVALUATING THE STATED HYPOTHESES. THE MATCHED-PAIRS SIGNED-RANKS TEST WAS USED AS THE STATISTICAL TEST OF SIGNIFICANCE. THE DATA INDICATED A SIGNIFICANTLY GREATER INCREASE FOR THE EXPERIMENTAL GROUP IN (1) UNDERSTANDING OF THE SCIENTIFIC PROCESS, (2) UNDERSTANDING OF SCIENTISTS AS AN OCCUPATIONAL GROUP, (3) UNDERSTANDING OF SCIENCE AS AN INSTITUTION AND ITS RELATIONSHIP TO OTHER INSTITUTIONS IN OUR SOCIETY, (4) CRITICAL THINKING ABILITY, (5) THE IMPORTANCE WHICH THEY PLACE ON THEORETICAL VALUES, AND (6) UNDERSTANDING OF THE CHARACTER OF SCIENTIFIC AND NON-SCIENTIFIC SEGMENTS WITHIN CULTURES AND KNOWLEDGE OF THE EVIDENCE FOR INTERACTION BETWEEN THEM. THE EXPERIMENTAL GROUP EVIDENCED NO GREATER INCREASE IN SUBSTANTIVE SCIENTIFIC KNOWLEDGE THAN THE CONTROL GROUP. IT WAS CONCLUDED THAT THE EXPERIMENTAL COURSE "SCIENCE AND CULTURE" IS EFFECTIVE IN PRODUCING A BROAD INCREASE IN THE SCIENTIFIC LITERACY OF SECONDARY SCHOOL STUDENTS SUCH AS THOSE WHICH COMPRISED THE EXPERIMENTAL GROUPS. APPENDED ARE THE (1) EXPERIMENTAL COURSE DESCRIPTIVE MATERIALS, (2) SPECIAL DIAGNOSTIC INSTRUMENTS, (3) COMPARATIVE STATISTICS FOR SAMPLE POPULATIONS, AND (4) "MY VIEW OF SCIENCE" ESSAYS. (DS)

BR-6-8445
P.A. 24

ED018400

FINAL REPORT

Project No. 6-8445-24
Contract No. OEC 3-7-068445--0374

GUIDELINES FOR THE DEVELOPMENT OF A COURSE FOR SECONDARY SCHOOLS EMPHASIZING THE INTERACTION OF SCIENCE AND THE CULTURE OF MAN

January 1966

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

SE004547

FINAL REPORT

Project No. 6-8445
Contract No. OEC 3-7-068445-0374

GUIDELINES FOR THE DEVELOPMENT OF A COURSE FOR SECONDARY SCHOOLS EMPHASIZING THE INTERACTION OF SCIENCE AND THE CULTURE OF MAN

January 1966

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

FINAL REPORT

Project No. 6-8445
Contract No. OEC 3-7-068445-0374

GUIDELINES FOR THE DEVELOPMENT OF A COURSE FOR SECONDARY SCHOOLS EMPHASIZING THE INTERACTION OF SCIENCE AND THE CULTURE OF MAN

Robert E. Yager

and

George W. Cossman

University of Iowa
Iowa City, Iowa

February 1968

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS USED IN TABLES

α	Level of Significance
CR	Critical Region
df	Degrees of Freedom
F	F Statistic
FAS I	Facts About Science, Understanding Science as Related to Society
FAS II	Facts About Science, Understanding of Scientists
N	Number of Subject Pairs
Sci. & Cult. Ach.	Iowa Science and Culture Study Achievement Test
SOS	Science Opinion Survey
SOV	Study of Values, Theoretical Scale
Stan. Ach.	Stanford Achievement Test - Advanced Science
t	t Statistic
T	Sum of Ranks, Wilcoxon Matched-Pairs Signed-Ranks Test
TOUS I	Test On Understanding Science, Understanding of Science as Related to Society
TOUS II	Test On Understanding Science, Understanding of Scientists
TOUS III	Test On Understanding Science, Understanding of the Scientific Process
Watson-Glaser	Watson-Glaser Critical Thinking Appraisal
X	Pretest Score
Y'	Predicted Posttest Score
z_r	Fisher Transformation of Product Moment Correlation Coefficient

PREFACE

The body of this final report consists of a Ph.D. dissertation completed by George W. Cossman under the direction of the investigator. Hence the body follows the guides prescribed by the Graduate College of the University of Iowa. The writing is that of Dr. Cossman as the instructor of the course, principal research worker, and principal writer of course materials. Approval for use of the dissertation as the final report was given in a letter to the investigator by Sue M. Brett. The project continues though seriously limited in terms of support funds and released time for the investigators to pursue the project at the desired level.

SUMMARY

An experimental course entitled "Science & Culture" was conceived and structured as an elective for eleventh and twelfth grade students in the Laboratory School at the University of Iowa. The course was an attempt at an interdisciplinary approach to solving the problem posed by failures of the regular courses in science and social studies as evidenced by a rather extensive testing program in the school. The specific failures are apparent from the hypotheses (stated in the null form) which were tested in this project. The hypotheses were:

1. The experimental course, "Science and Culture", does not produce significant increases in students' understanding of the scientific process as measured by the criterion instruments.

2. The experimental course, "Science and Culture", does not produce significant increases in students' understanding of scientists as an occupational group as measured by the criterion instruments.

3. The experimental course, "Science and Culture", does not produce significant increases in students' understanding of science as an institution and its relationship to other institutions in our society as measured by the criterion instruments.

4. The experimental course, "Science and Culture", does not produce significant increases in students' ability to think critically as measured by the criterion instrument.

5. The experimental course, "Science and Culture", does not produce significant increases in students' substantive scientific knowledge as measured by the criterion instrument.

6. The experimental course, "Science and Culture", does not produce significant increases in the importance which students place on theoretical values as measured by the criterion instrument.

7. The experimental course, "Science and Culture", does not significantly increase students' understanding of the character of scientific and non-scientific segments within cultures and their knowledge of the evidence for interaction between them as measured by the criterion instrument.

The following instruments were used as pre and post test measures to test the hypotheses. Many have subscores which approach one or more of the hypotheses. The tests were: Test on Understanding Science, Facts About Science, Watson-Glaser Critical Thinking Appraisal, Stanford Advanced Science Achievement Test, Study of Values, "Science Opinion Survey", and "Iowa Science and Culture Study Achievement Test". A rationale concerning the use and the limitations of each of these instruments is presented. A discussion concerning the statistical analyses is also included in the study. The non-parametric test of significance of Wilcoxon Matched-pairs Signed-ranks Test was used for testing the null hypotheses. It has been found that results of this test agree closely with the results of the parametric analysis of variance.

From the specific data presented, it is evident that the group of twenty-one eleventh and twelfth grade students exposed to the "Science and Culture Treatment" showed substantial mean gains in scores on a variety of evaluative instruments. The data also show that a group of effectively matched subjects who were not exposed to the treatment failed to achieve such mean gains. The results of the main analysis have indicated that this observed difference in growth could have been obtained by chance in only one of one hundred replications of the experimental comparison. Thus, since logical and empirical evidence was presented to exclude other systematic differences as explanations, the greater growth exhibited by the experimental subjects can, with considerable confidence, be attributed to the effect of the treatment. Since neither of the comparison groups showed a significant growth in one of the dependent variables studied, the treatment must be considered ineffectual for that variable.

Translated into specific terms and placed in relationship to the seven hypotheses presented in the statement of the problem, these outcomes indicate that subjects who were students in the experimental course, "Science and Culture", evidenced a significantly greater increase in:

1. understanding of the scientific process, as measured by the Test On Understanding Science and the "Science Opinion Survey", than subjects who were not.

2. understanding of scientists as an occupational group, as measured by the Facts About Science Test and the Test On Understanding Science, than subjects who were not.

3. understanding of science as an institution and its relationship to other institutions in our society, as measured by the Facts About Science Test and the Test On Understanding Science, than subjects who were not.

4. critical thinking ability, as measured by the Watson-Glaser Critical Thinking Appraisal, than subjects who were not.

5. the importance which they place upon theoretical values, as measured by the Study of Values, than subjects who were not.

6. understanding of the character of scientific and nonscientific segments within cultures and knowledge of the evidence for interaction between them, as measured by the "Iowa Science and Culture Achievement Test", than subjects who were not.

But they also indicate that subjects who were students in the experimental course evidenced no greater increase in substantive scientific knowledge, as measured by the Stanford Advanced Science Achievement Test, than subjects who were not.

These positive results are corroborated at several points by evidence obtained from the "open-ended" techniques used in the study, the "Descriptive Words Test' and the "My View of Science" essays. In isolation such data would be of questionable significance since the symbols of learning cannot be equated to learning itself. However, when placed on the background provided by the criterion test data, it is indicative of more than the fact that students have learned to produce the accepted verbal symbols. There is good reason for contending that they understand the meaning of these symbols as they relate to science. The "My View of Science" essays are especially valuable in that they serve the additional function of reflecting student attitudes toward science. In a number of cases they provide evidence of a fact which was clear to the course instructors from the outset: some of the experimental subjects selected the "Science and Culture" course because of a desire to fulfill a science requirement while avoiding a formal science course. This fact has obvious significance for interpreting the outcome of the study. It should be kept clearly in mind when the question of bias due to "self-selection" is considered. Moreover, an examination of the pretreatment-posttreatment essays will reveal that the treatment was able to overcome a considerable anti-science attitude in several cases. The fact that no significant difference in growth was observed on the Stanford Science Achievement Test is not surprising for two reasons. First, the "Science and Culture" course was not designed to increase students' knowledge of science facts. Secondly, the intrasessional history of control group subjects included a mean number of semesters of science courses which was almost four times that for experimental subjects.

In conclusion, the findings of this study rather clearly demonstrate that the experimental course, "Science and Culture", is effective in producing a broad increase in the scientific literacy of secondary school students such as those which comprise the experimental group. The substantial significance of the results obtained with an analysis employing nonparametric statistics justifies placing considerable confidence in this assertion. Since the need for fostering scientific literacy and attitudes is a matter of major educational concern, the potential value of such a course is evident. Although additional research is necessary before any conclusions can be drawn concerning the effectiveness of these course materials with other learners and other teachers, there seems to be reasonable cause for optimism.

THE EFFECTS OF A COURSE IN SCIENCE AND CULTURE

DESIGNED FOR SECONDARY SCHOOL STUDENTS

by

George Wilford Cossman

A thesis submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in
Science Education, College of Education, in the
Graduate College of The University of Iowa

August, 1967

Thesis supervisor: Professor Robert E. Yager

CHAPTER I

INTRODUCTION

Introduction to the Problem

In an address given at the twenty-fifth anniversary (1966) dinner of the Cooperative Committee on Science Teaching, the well known University of Chicago psychologist and sociologist, Robert Havighurst, reviewed developments in the recent history of science education. During his comments Professor Havighurst attempted to look to the frontier of science education and to speculate concerning the problems which it will be seeking to solve in the future. He stated that his greatest concern was with the problem of the "uncommitted" adolescents, "the rather sensitive, intelligent young person who hasn't got enough faith in our society -- or enough faith in himself -- to commit himself to get in there and make a career for himself, do the things that other people do, graduate from college and go right to work, or take a graduate degree". Concerning the solution to this problem Havighurst stated, "I'm inclined to think that with a different kind of rather positive teaching about the universe -- man and the universe -- we might help these boys and girls see themselves as part of something that has a long history that they can carry on -- a desire to further goals and get themselves really committed in the process of carrying on in society". He went ahead to define the problem of the relation of man to the universe as the conflict between the scientific and the humanistic traditions, the split into what

Sir Charles Snow (1964) calls "two cultures". And having thus clarified matters, he stated that the educational problem is one of bridging the gap, presenting "a way of looking at man and the universe that boys and girls can take hold of and that can give them a positive approach to life". In his conclusion Havighurst stated,

> We are likely to see a number of attempts in the next few years, both at the high school and college levels, to make an integration of what we know scientifically about man in the universe with something that is more intuitive and has come down through literature and philosophy . . . . So while it does seem a little peculiar, I realize, to talk about science teaching as an effort to help young people commit themselves to carrying on the human adventure, I have been impressed that people at all times in talking about science teaching have said that it may be important in helping people see themselves in history and to find their place and find their task. So I would think that we may find, in ten years, some very interesting experimental courses probably at the senior high school level, and perhaps the college level, that use science in this attempt to combine somehow the vision of the humanist with the vision of the scientist.

Professor Havighurst has provided science educators with a considerable challenge in placing upon them the responsibility for halting the growing separation between science and other fields of learning, especially the humanities. There is very little, if any, evidence to indicate that a solution to this problem has been formulated and seriously pursued upoto now by any group within the structure of education. Despite the furor which arose after Sir Charles Snow's 1959 Rede Lecture and the growing tide of pleas for action from such men as Conant, Cohen, Sarton, Butterfield, Bronowski and Burrt, matters today are not very different from what they were before 1959 (Holton, 1965, p. vii). It is true that responsibility for what is perhaps a related task has been accepted by both science and social studies educators: the preparation of students to function effectively as citizens in a society based upon and committed to science. Specifically, social studies educators see the development of the knowledge and understanding necessary for social competence within our science oriented and science affected society as one of their primary goals (Cummings, 1957). Science educators now accept the fostering of an understanding of scientists and the scientific process and the encouragement of critical thinking as among their principle responsibilities (Rogers, 1960, p. 19). However, both groups have found that their independent efforts to meet these objectives have fallen short of the mark.

In the fall of 1965 the investigator, along with Dr. Robert Yager, University of Iowa Professor of Science Education and Dr. Doyle Casteel, University of Iowa Associate Professor of Social Studies Education, began considering the possibility of preparing materials for a secondary school course which would attempt to overcome the shortcomings of past efforts to achieve the above mentioned goals and at the same time bridge the "gap" separating the sciences and the liberal arts. After considerable study it was decided that the feasability of such a course was sufficient to justify preparing and testing a preliminary version. Thus several months before Professor Havighurst delivered his remarks, the groundwork was being laid for a course of the kind which he proposed would appear in ten years.

Because of a background in both the liberal arts and sciences and previous experience in preparing and teaching materials for multi-disciplinary courses, the investigator, with the help of Professors Yager and Casteel, subsequently began producing materials for the proposed course. During the 1966-67 academic year the first three units of the course, entitled "Science and Culture", were taught to a heterogenous group of twenty-one eleventh and twelfth grade students in the University of Iowa Laboratory School. A brief description of the course as well as specimen of the reading material from Unit Three will be found in Appendix A.

Statement of the Problem

The purpose of the research to be reported here was to determine to what extent, if any, several of the basic objectives around which the experimental course was structured were achieved. Specifically the following hypotheses, stated in the null form, were investigated:

1. The experimental course, "Science and Culture", does not produce significant increases in students' understanding of the scientific process as measured by the criterion instruments.

2. The experimental course, "Science and Culture", does not produce significant increases in students' understanding of scientists as an occupational group as measured by the criterion instruments.

3. The experimental course, "Science and Culture", does not produce significant increases in students' understanding of science as an institution and its relationship to other institutions in our society as measured by the criterion instruments.

4. The experimental course, "Science and Culture", does not produce significant increases in students' ability to think critically as measured by the criterion instrument.

5. The experimental course, "Science and Culture", does not produce significant increases in students' substantive scientific knowledge as measured by the criterion instrument.

6. The experimental course, "Science and Culture", does not produce significant increases in the importance which students place on theoretical values as measured by the criterion instrument.

7. The experimental course, "Science and Culture", does not significantly increase students' understanding of the character of scientific and non-scientific segments within cultures and their knowledge of the evidence for interaction between them as measured by the criterion instrument.

CHAPTER II

REVIEW OF RELATED LITERATURE

In reality, to speak of research of the kind undertaken in the present study is a misnomer. This statement is based upon the fact that previously there has been no attempt made to bring about a broad improvement in "scientific literacy" among a group of secondary school students. Thus to research the outcome of such an attempt has been an obvious impossibility. If "scientific literacy" is resolved into its several components, research literature related to several of these components can be found, however.

Pella, O'Hearn and Gale have suggested that a "scientifically literate individual presently is characterized as one with an understanding of the (a) basic concepts of science, (b) nature of science, (c) ethics that control the scientist in his work, (d) interrelationships of science and society, (e) interrelationships of science and the humanities and (f) differences between science and technology" (Pella, O'Hearn, and Gale, 1966, p. 199). Although somewhat vague when isolated from their elaboration, these categories of scientific literacy closely parallel the types of understanding which the "Science and Culture" course is designed to foster. The major point of disjunction lies in the fact that the course is not specifically intended to increase students' fund of scientific knowledge. This task is performed adequately by traditional courses in science. It is also true that it attempts to not only enhance understanding of the process of science but also seeks to encourage several of the attitudes involved, specifically those of critical thinking and theoretical orientation. From among the above areas of behavioral objectives only two could be found represented in educational research literature, that of fostering an understanding of science and scientists and that of increasing critical thinking ability. The research involved was almost entirely limited to that in science education.

There are a number of different conceptions of the type of behaviors which constitute critical thinking and a considerable fraction of the literature is devoted to research attempting to identify the attributes involved. The work of Ennis (1962), Furst (1950), and Guilford (1950, 1951) is representative. Although research regarding the relationship between critical thinking, creative thinking, intelligence, and judgement is lacking there seems to be reasonable agreement that critical thinking involves a composite of attributes including the use of correct logic; the use of scientific methods (emphasis on evidence and hypothesis testing); and the urge to be cautious, inquisitive, and systemmatic.

Two major approaches are advocated for fostering critical thinking. The first involves merely creating a type of environment conducive to it. The results of the Eight Year Study (Houle, 1943) and the research of Dressel and Mayhew (1954) provide evidence for the success of this approach in large scale situations. Both show that institutions with a greater degree of student self-determination, flexibility of curriculum, and freedom from authoritarian control of behavior have a significantly better record in increasing critical thinking. Studies on a smaller scale have shown that questioning and critical behavior are less apt to occur in formalized, highly structured situations (Carpenter, 1956).

The second approach is more direct and involves at least some type of planned effort to "teach" critical thinking. In some cases this may amount to actual instruction in the rules of logic and their application. Henderson (1958) relates a study in which secondary school students were tutored in the principles of logic and given practice in their use. He reports that the experimental students showed greater gains on measures of critical thinking than did their controls. Dressel and Mayhew (1954), on the other hand, note that institutions having a special course in critical thinking do not stand apart from those which do not have such courses. In one study they report that students who completed a course in logic showed no significant growth in critical thinking ability. It seems appropriate to question the possibility of a different relationship between training and testing in these two situations. According to Dressel (1960) evidence favors direct teaching over indirect teaching of scientific methods and attitudes. The question would seem to be one of what degree and kind of "directness" is needed. Several studies (Zapf, 1938, Wessell, 1941) have provided evidence to support the expected conclusion regarding the effect of simply teaching traditional subject matter by traditional methods. Even excellent teaching, if done to merely teach conclusions, accomplishes nothing as far as fostering careful thinking or any other scientific attitudes. The fact that several investigators (Strauss, 1931, Caldwell, 1931, and Dewitt, 1959) have been able to report positive correlations between subject matter achievement and the acquisition of habits of critical thinking is most apt to be due to a correlation between the two behavioral variables and a third unidentified variable.

The research regarding actual experiences which are conducive to the development of critical or creative thinking does not consistently point to the same conclusions. There does seem to be agreement concerning the value of discussion methods in stimulating active thinking and subsequent growth in critical thinking (Bloom, 1953). When discussion methods are employed students are placed in a position which encourages thinking creatively and allows them to judge other's thinking as well as have their own judged (McGeoch and Irion, 1952). However, on the whole, research literature in the area of creative thinking points most clearly to one conclusion; the need for more research (Washton, 1966).

In contrast to the normative research done regarding critical thinking, at least one genuine innovation for increasing understanding of science and scientists has been reported and evaluated in the literature. This is especially interesting in view of the fact that "rational thinking ability" was clearly recommended as an educational objective in the 1944 report of the Educational Policies Commission, while understanding science was not. In 1960 a set of history of science case studies designed for use by secondary school students was published (Klopfer, 1960). The case study approach to teaching an understanding of science and scientists had been introduced at Harvard by Conant in 1948 (Conant, 1951). In 1963 Klopfer and Cooley reported the results of an extensive study designed to determine the effectiveness of the History of Science Cases for High Schools (HOSC) in changing student understanding of science and scientists (Klopfer and Cooley, 1963, pp. 33-47).

The criterion instrument used in the study, the Test On Understanding Science or TOUS, was developed by the two investigators for such evaluative use (Cooley and Klopfer, 1963, pp. 73-80). Klopfer and Cooley reported that their forty experimental classes showed a mean (adjusted) growth of 5.09 raw score points on this test, while the control classes showed a growth of only 2.10 points. Since the difference was highly significant the investigators concluded that the HOSC method was an effective adjunct to ordinary instruction in secondary school science courses.

In the period since the appearance of the TOUS the literature has become rich in reports of research involving its use. Despite having been constructed for secondary school students, it has been used with teachers, administrators, teacher preparation students and college science professors. Some of this research seems to have been well directed. But a large percentage involves a "Here's something else to try!" attitude. In the following paragraphs a small, but representative, sample of studies from the former category is outlined.

Kleinman (1965) has reported using the TOUS to measure the understanding achieved by junior high school students in classes taught by teachers using "critical thinking questions" as opposed to teachers who do not use them. The results show a significantly higher score is obtained by students in the classes taught in the "critical thinking pattern".

Sorensen (1966) compared the TOUS scores of a high school group studying biology in a laboratory-centered program to those of a group in a lecture-centered program in the Salt Lake City School System. He reports that the pretest-posttest gain scores were significant in each case and that those for the "lab-treatment" group were significantly greater than those for the "lecture-treatment" group.

Crumb (1965) used the TOUS to demonstrate that secondary school physics students show significant gains in their understanding of science over the school year. He concluded that the magnitude of gain was sensitive to the teaching method used in the course.

Not all of the studies using the TOUS have given positive results. Trent (1965) has reported that there is no significant difference in the understanding of science gained by students studying PSSC physics and those in courses using some more traditional kind of materials. Snider (1966) found that there was no significant difference between the TOUS scores of physics students taught by teachers who differ in their rating on the Flanders teacher-pupil interaction scale. The mean pretest-posttest gain scores on the TOUS were reported as being essentially zero in all classes.

Even though the number of studies using the TOUS is several times the number reported here, their general nature should now be clear. More important than the number of studies reported in the literature is the fact that the HOSC project is the only significant innovation designed to increase student score gains on the TOUS. It should be noted, however, that the history of science case study approach has now been abandoned at Harvard as failing to accomplish the goals set for it (Allen, 1966).

CHAPTER III

EXPERIMENTAL PROCEDURE

Instrumentation

In order to make possible the necessary tests of significance for the seven hypotheses presented in the statement of the problem, data must be collected in an equal number of behavioral areas. In all except one case (substantive knowledge) the behavioral variable involved is one which has been largely ignored in educational research until quite recently. The consequence is that available measuring instruments are few in number and typically still in the experimental stages. Under the circumstances, an attempt was made to identify several instruments which could be used to provide data for testing each of the hypotheses. Unfortunately this effort proved successful for only three of the six hypotheses concerned. However, two of the tests used as sole sources of data are very likely ample for the purpose. The instrument used to evaluate critical thinking ability is one which has gained rather wide acceptance since its publication and possesses considerable content validity. In the case of the "importance placed on theoretical values" variable it was possible to make use of a widely accepted test by administering the entire set of items and selecting those upon which the desired measurement index could be based. Since no published test presently available was appropriate as a source of data for the behavioral variable concerned in the final hypothesis, a special test was written for the purpose. Since the items in this instrument are taken directly from the collection of understandings which it is intended to measure, the score which a respondent receives is an appropriate index of his level of understanding so long as we assume that those items are representative of that collection and are properly constructed. Beyond such "logical validity" it can, and will be, shown that results on this test correlate well with those obtained on similar tests.

A summary of the behavioral characteristics which are dependent variables in this study and the instruments used to obtain index values descriptive of these characteristics will be found in Table 1. In the following paragraphs a few brief comments will be made concerning the instruments involved.

The Test on Understanding Science first became available in its final form during 1961 (Cooley and Klopfer, 1961). Previous to that time no published test of its kind was available. The sixty four-choice items on the test are designed to sample adolescents' understanding of science in three specific areas. Keys are available for obtaining scores in each of these areas. The first of the three pertains to understanding of the scientific enterprise (Subscore I). The questions involved concern such matters as the human element in science, communication among scientists, the international character of science, and the interaction of science and society. The second area of understanding pertains to scientists as an occupational group (Subscore II). The test items deal with abilities needed by scientists, generalizations about scientists as people, and institutional pressures on scientists. The remaining portion of the test measures understanding of the methods and aims of science. Specifically, the questions involved probe such matters as the nature of theories and models, generalities about scientific method, tactics and strategy in science, accumulation and falsification and the relationship between science and technology.

TABLE 1

RESUME OF BEHAVIORAL VARIABLES AND ASSOCIATED INSTRUMENTATION

Hypothesis	Behavioral Variable	Measurement Index
1	Understanding of the Scientific Process	Test on Understanding Science (Subscore III) "Science Opinion Survey"
2	Understanding of Scientists	Test on Understanding Science (Subscore II) Facts About Science (Subscore II)
3	Understanding of Science as Related to Society	Test on Understanding Science (Subscore I) Facts About Science (Subscore I)
4	Critical Thinking Ability	Watson-Glaser Critical Thinking Appraisal
5	Substantive Scientific Knowledge	Stanford Advanced Science Achievement Test
6	Importance of Theoretical Values	Study of Values (Theoretical Scale)
7	Understanding of Science as Related to Culture	"Iowa Science and Culture Study Achievement Test"

Within several months after the experimental course was underway the investigator became aware of a test on understanding the nature of science which had been prepared by Professor Merritt E. Kimball for use with adults (Kimball, 1965). Professor Kimball consulted a considerable volume of literature in the area of the philosophy of science and from it constructed a model of the scientific process. He then prepared a pool of two hundred short statements about the nature of the scientific process based upon the assertions in the model. This pool of statements was then selectively refined through the use of a panel of consultants until a group of twenty-nine statements remained. These statements became the stimulus items for the test. The respondent is asked to agree or disagree with each of the items. Scoring is based upon the total number of responses which represent correspondence with the model. Kimball has used his "opinion survey" to compare the understanding of the scientific process possessed by qualified science teachers, practicing scientists, and philosophy majors. The results which he obtained provide a basis of reference for the scores obtained by subjects in the present study. A specimen of the "Science Opinion Survey" is presented in Appendix B.

Facts About Science, despite its rather misleading title, is designed to measure the understanding which students have of the role of science in contemporary society and the characteristics of persons in the scientific profession. Separate subscores are provided for in the scoring keys. According to the author "the test is meant to indicate how the student perceives the implications of these facts and techniques {that are of concern in science courses} for society and for his life" (Stice, 1958, p. 1). Thus the focus is on understanding of science as an institution, how it differs from other institutions, how it affects society and daily life and how scientists as a group differ from people in general. One of the two subscores (Subscore I) reflects the student's level of understanding regarding the first three areas and the other (Subscore II) his understanding of the fourth.

Since the term "critical thinking" has been variously defined and is therefore apt to be misinterpreted, it is important to make clear what definition is implicit in the Watson-Glaser Critical Thinking Appraisal. According to the authors of this instrument, critical thinking is a "composite of attitudes, knowledge, and skills" (Watson and Glaser, 1964, p. 10). They include in this composite the ability to recognize the existence of problems, the insistence upon evidence to support assertions, knowledge of the character of valid inferences and generalizations and the role of weighing evidence in arriving at them, and the ability to skillfully employ and apply all of the preceding. In measuring the degree to which an individual possesses the characteristics included in this composite, the Critical Thinking Appraisal calls for responses to both "neutral" and "loaded" topics. These responses provide the opportunity to exhibit the characteristics enumerated above with respect to problems, statements, arguments and the interpretation of data. The authors state that the stimulus situations are similar to "those which a citizen in a democracy might encounter in his daily life as he works, reads newspaper or magazine articles, hears speeches, participates in discussions on various issues, etc." (Watson and Glaser, 1964, p. 2).

Within the battery of tests used, the Stanford Advanced Science Achievement Test is by far the most orthodox. It represents one of a collection of commonly used instruments designed to measure subject matter achievement. The test is non-specific for the various subject matter divisions within science and therefore provides a general index of the student's factual knowledge with respect to the entire discipline.

In 1928 Spranger proposed that every person has a dominant value, a unifying philosophy of life, which structures his personality (Spranger, 1928). He suggested six basic value types: the theoretical, the economic, the aesthetic, the social, the political, and the religious. Spranger's theory later led to the development of a paper and pencil inventory designed to determine the relative importance of each of these six values to the individual. The inventory, designed by Allport, Vernon, and Lindzey, contains a series of items which call for choices between competing statements or require choices to be ranked in order of preference (Allport, Vernon, and Lindzey, 1951). The respondent reveals the relative importance he attaches to each of the six values in the way he chooses between responses relevant to the six value areas. The higher an individual "scores" in the theoretical area the greater is the importance that individual places on discovering the truth, seeking to observe and reason, and finding order and pattern in things. By selecting the items which relate to the theoretical area and totaling the response weights which are assigned, it is possible to obtain what can be described as a "theoretical value index". The response items which are involved in deriving this index will be found reproduced in Appendix B.

The "Iowa Science and Culture Study Achievement Test" was devised specifically for the purpose of collecting data to test the last hypothesis presented in the statement of the problem. The fifty four-choice items on the test sample the student's understanding of the value systems which characterize both scientific and non-scientific activities and institutions, as well as his understanding of the reciprocal relationship between these two areas. The items deal with generalizations rather than minutiae and cover a considerable variety of topics within the above framework. A high score on the test reflects an advanced level of sophistication in perceiving the common ground between intellectual and cultural history and the history of science. A specimen of the test is presented in Appendix B.

It is apparent that the "Science and Culture Achievement Test" is a "broad spectrum" instrument which to some degree, while unique in its historical orientation and in sampling understanding of the interaction between the scientific and non-scientific facets of culture, subsumes the testing objectives of the Facts About Science Test, the "Science Opinion Survey", and the Test on Understanding Science. Since it is ad hoc in origin and not the product of experienced test designers, there may be some hesitation to accept its legitimacy on the basis of logical validity alone. Thus it seems appropriate to report the correlation between scores on the "Science and Culture Test" and the above mentioned published tests. Questions may also arise concerning the correlations between raw scores on the pairs of tests used to collect data on the first three behavioral variables listed in Table 1. For these reasons and others, intercorrelations between posttest scores on all ten criterion tests are reported in Table 2. The entire population of forty-two subjects was used in computing the coefficients of correlation in the table.

An inspection of Table 2 will reveal that the correlations between raw scores on the "Science and Culture Achievement Test" and other published tests are substantial. As could have been anticipated, the "theoretical value index" does not correlate significantly with any of the other test scores. It will also be noted that the correlations between the tests paired together in Table 1 are quite high: .57 for the "Science Opinion Survey" and Subscore III on The Test on Understanding Science, .76 for Subscore II on Facts About Science and Subscore II on The Test on Understanding Science, and .74 for Subscore I on Facts About Science and Subscore I on The Test on Understanding Science. These correlations appear sufficient to justify the intended use of both scores in each pair as indices to describe the specified behavior variable.

TABLE 2

INTERCORRELATION MATRIX FOR POSTTEST SCORES ON ALL CRITERION INSTRUMENTS

	FAS I	FAS II	TOUS I	TOUS II	TOUS III	Wat.-Glaser	Stan. Ach.	Sci.&Cult.Ach.	SOS	SOV
FAS I		.71	.74	.59	.59	.63	.46	.68	.37	.28
FAS II	.71		.57	.76	.70	.54	.40	.65	.10	-.07
TOUS I	.74	.57		.70	.76	.48	.45	.61	.39	.31
TOUS II	.59	.76	.70		.68	.60	.32	.70	.50	.25
TOUS III	.59	.70	.76	.68		.57	.53	.72	.57	.37
Watson-Glaser	.63	.54	.48	.60	.57		.33	.57	.33	.18
Stan. Ach.	.46	.40	.45	.32	.53	.33		.43	.08	.10
Sci.&Cult.Ach.	.68	.65	.61	.70	.72	.57	.43		.58	.27
SOS	.37	.10	.39	.50	.57	.33	.08	.58		.25
SOV	.28	-.07	.31	.25	.37	.18	.10	.27	.25	

Design of the Experiment

Once the hypotheses to be evaluated have been selected and instruments capable of measuring the pertinent variables have been chosen, the problem of selecting a proper experimental design must be dealt with. The objective is obviously one of collecting all of the data necessary to provide a basis for the best possible evaluation of the hypotheses. If it were not for the fact that the experimental "materials" happen to be so complex and heterogeneous and the practical limits of manipulation so real, this problem would be no more involved in the behavioral sciences than elsewhere. However, things being as they are, it is seldom that a "model design" with all its advantages can be readily carried over into practice. To the extent that the investigator's design achieves correspondence with the ideal design for his experiment, he achieves the maximum possible precision, validity and generalizability. At those places where he fails and cannot take appropriate measures to counter the failure, he surrenders some degree of one or more of these qualities. Without knowledge of these failures it is impossible to know what inferences can be properly drawn from an experiment and what degree of confidence can be placed in them. For this reason an attempt will be made to briefly outline the "ideal" paralleled in the design used in this study, the particular strengths of that design and the nature and significance of departures from the parallel.

Following the recommendations of current methodological literature, the pretest-posttest control group design with matched pairs has been employed as closely as possible. The essential procedural steps in this design are: (1) Initial Selection -- a random sample of subjects is selected from the available population, (2) Matching -- these subjects are measured on a response variable or variables thought to correlate with the criterion variable(s) and are then paired as closely as possible in terms of their performance, (3) Assignment -- one member of each matched pair is randomly assigned to the experimental group, the other to the control group, (4) Pretesting -- both groups are measured on the criterion variable(s), (5) Treatment -- the experimental group is exposed to the treatment whose effects are to be studied, while the control group receives no exposure, (6) Posttesting -- both control and experimental groups are measured on the criterion variables again.

As far as precision is concerned there are three common methods used to bring about an increase in its level: matching, the use of pretest-posttest growth scores as criterion measures, and covariance adjustment. It is already apparent that the above design incorporates the first of these three methods. Assuming that the necessary conditions for employing an analysis of covariance can be satisfied, both of the other two methods can be used as well. For reasons which will be made clear later, it was not possible to satisfy these conditions in the present study. However, the use of growth scores was quite feasible. Thus although not possible to achieve the ideal in terms of precision, it was possible to come rather close to that ideal.

Turning now to the question of achieving maximum validity, Campbell and Stanley (1963, p. 175) have suggested that there are eight different extraneous variables which, if not controlled, jeopardize the validity of experiments by producing effects that become confounded with the treatment effect. These eight are: (1) Intrasessional History -- the specific activities which occur along with the treatment during the experimental period, (2) Maturation -- the changes which occur as a result of the mere passage of time, (3) Testing -- the effect which being exposed to a pretest has upon the results obtained in posttesting, (4) Instrumentation -- the changes which occur as a result of changes in instrumental calibration or interpretation between the beginning and end of the experiment, (5) Statistical Regression -- in which groups with extreme scores regress toward more average scores, (6) Selection Bias -- which operates through differential selection and assignment of subjects to produce initial inequality on the criterion variable(s), (7) Experimental Mortality -- the selective removal of subjects from the comparison groups, (8) Selection Interaction -- the presence of differential selection and assignment of subjects resulting in systemmatic between groups differences in maturation, intrasessional history, motivation, etc.

The pretest-posttest control group design provides a high degree of control over all eight of these variables. The main effects of maturation, testing, and instrumentation can clearly be eliminated as possible sources of treatment as opposed to no-treatment group differences, since both groups are equally subject to their influence. In general, the same can be said for intrasessional history, but the situation is a bit more subtle here. As far as a treatment versus no treatment design is concerned, the problem can be best understood by thinking of the possibility that the treatment group has been exposed to some experience associated with the treatment and capable of influencing posttest results but not properly considered part of it. An absurd example would be a discussion of the answers to the questions which appeared on one or more of the pretests. In actual matter of fact, cases of serious confounding from intrasessional history are rather hard to imagine in educational experiments involving criterial variables with considerable "inertia", variables which ordinarily do not undergo change unless a rather significant "force" comes to bear. In such experiments there is little chance that an influence with enough "impulse" to significantly affect the criterion variables could escape detection. The possibility of regression effects is easily rejected since matching assures that neither group occupies an extreme position on the scale of criterion scores. Thus if regression effects do exist both groups regress equally. Differential mortality presents no particular validity problem since mortality need not be differential. When for some reason complete data is not available for a subject in either group, that subject as well as the second member of the matched pair can be eliminated. Of course, this loss of data is undesirable for statistical reasons and may be undesirable in terms of altering the representativeness of the groups as well. But such would be equally true with other designs, assuming some logical rationale could be found for selecting a person to be eliminated from the second group of subjects. As far as the two selection variables are concerned, the first, selection bias, is ruled out to the extent that matching has equated the two groups with respect to the criterion variables. Furthermore, the use of growth scores rather than posttest scores as criterion variables provides control over whatever differences remain after matching. The combination of matching and growth score comparison removes an objection which is sometimes made to the use of growth scores alone: the fact that equal gains on different segments of the measurement scale are not necessarily equally significant. When the correlations between the matching variables and the criterion variables are reasonably high, that is, when it would have been possible to predict pretest scores with relative accuracy, matched pairs will tend to achieve nearly equal pretest scores and hence their gains will fall on approximately the same portions of the score continuum. The extent to which selection interaction has been controlled is a function of the degree to which randomization has accomplished its intended purpose. Since subjects are randomly assigned to the control and treatment groups after they have been placed in matched pairs, the liklihood of systematically placing subjects of a particular kind in one group is extremely small.

Beyond the fact that it does not employ covariance adjustment the actual design employed in this study differs from the ideal design at two points: in the method used to initially select subjects (procedural step 1) and in the method used to assign subjects to groups (procedural step 3). These differences were made necessary by the practical limitations inherent in the experimental situation. Since it is not possible to randomly select students and demand that they register for a particular course, it was necessary to allow "self selection" (within the limits established by requiring junior or senior standing) to determine the makeup of the experimental group. In order to achieve matching it was then necessary to select subjects for the control group who were as much like the experimental subjects with respect to the matching criteria as possible. Since data on these variables were readily available in the files kept by the school guidance department, this was easily accomplished.

The significance of departure from the first procedural step in the ideal design, in which all subjects are selected randomly from a population, is actually related as much to the question of generalizability as it is to that of validity. Clarity will be served best if both sides of the problem are dealt with together. The major purpose of random selection from a large population is to achieve representativeness, that is, to make possible the generalizing of the conclusions of the experiment from the sample to the parent population. In reality, it makes little sense in the present situation to speak of random sampling from a <u>real</u> population in order to achieve representativeness for that population. The population and the sample are identical: the junior and senior students in a particular secondary school who seek to enroll in a particular course. However, it can be contended that this group represents a random sample from a <u>hypothetical</u> parent population composed of "all students who will in the future enroll in the 'Science and Culture' course offered by the University of Iowa Laboratory School" -- assuming general conditions remain stable. The important matter of generalization to a real parent population such as "senior high school students in Iowa secondary schools" is largely left to be explored by future experiments.

The problem of validity raised by the method used in the initial selection of subjects is that, under the circumstances, the experimental and control groups do not, in a strict sense, come from the same population. The implications of this fact can best be considered in combination with a discussion of the second departure from the ideal design, the failure to randomly assign subjects to groups. This follows from the fact that both departures relate to the same question of altered validity.

The relationship between actual and ideal design can best be seen if a new ideal design is considered which is closely related to the one already discussed and equivalent to it when very large samples are involved.

In this design matching is not used as an adjunct to randomization. This is made possible by the fact that if we independently select two large samples from the same population by random procedures it is reasonable to assume that the two samples are equivalent. If we now temporarily substitute this procedure for the random selection - matching - random assignment procedure, the overall departure of the actual procedure from the ideal procedure is reasonably clear. In the ideal procedure subjects for both groups are randomly selected from a given population. In the actual procedure subjects for the experimental group were "self selected", i.e., came from a subpopulation of "seekers". Subjects for the control group were selected by the investigator on the basis of their desirability as controls and were "recruited". It should not be concluded, however, that this latter group necessarily comes from a subpopulation of "nonseekers". Some of its members will return to school in the fall as "seekers". The essential issue here is whether or not confounding with the selection-interaction variable is apt to occur. The level of control over the other seven extraneous variables is not placed in question. Specifically, the question is whether or not a superior performance on the part of the experimental group could be explained as a result of differential selection having produced a systemmatic difference between the two groups with respect to the specific influence of maturation, history, testing, or some other such variable. Assuming that the matching process has been effective, the matching variables and all other variables closely related to them are, of course, eliminated from the list of possibilities. This being the case, a consideration of the nature and effectiveness of matching is pertinent.

The control and experimental groups were matched on the following variables: sex, grade level, cumulative grade point average, intelligence as measured by the Henman Nelson (form A), and current or previous registration in senior high school chemistry. The reason for choosing these particular matching criteria is evident except perhaps for the last in the series. Since the laboratory school at the University of Iowa has a three track science curriculum with senior high school chemistry as the first course on the upper track, whether or not a student is or has been enrolled in chemistry is an index of whether or not he is on that upper track. All students are required to complete six semesters of science before the tenth grade. Thus matching students on the chemistry criterion equates them, to some degree, on involvement with formal science courses. It does not, of course, match them in terms of science courses taken during the time of the experiment.

The matching achieved on sex and the chemistry criteria was perfect. In the case of grade level it was necessary on five occasions to pair an eleventh grade female in the experimental group with a twelfth grade female in the control group. However, there seems to be no reason to believe that the experimental group is favored by such a "mismatch". If indeed there is any significant advantage involved, it would be more reasonable to say that it lies on the side of the control group. The extent to which the two groups were matched on intelligence and cumulative grade point average is indicated by the between group product - moment correlation coefficients for these two variables. This coefficient was .914 for grade point average and .493 for intelligence quotient. Although grade point average in science courses was not used as a matching criterion, the correlation coefficient obtained for this variable was .828. Additional comparative statistics will be found in Appendix C.

The above evidence would seem to be sufficient to exclude the matching criteria from the collection of variables which may have been differentially distributed in the selection process. Several others can be eliminated for equally cogent reasons. If the list of extraneous variables previously discussed is examined as a guide to sources of confounding, simple logic is sufficient to remove selection - maturation, selection - instrumentation, selection - mortality, and selection - testing interaction as explanations for a difference in the degree of shift from pretest to posttest scores. Even in the presence of reasons for believing that differential selection has taken place for the extent or rate of maturation during the experimental period, there would be no logical basis for using it to explain differences in criterion performance. This follows from the simple fact that none of the criterion variables tend to undergo change with maturation. Selection - instrumentation interaction is eliminated since the psychometric instruments employed are not subject to changes in calibration and there is no relationship between the basis upon which a subject is selected and the "reading" of the instruments used to evaluate his performance. Selection - mortality interaction can be excluded even though it may be true that a "seeker" is less likely to withdraw from a course than a conscripted student is to refuse to take a battery of posttests. Here the exclusion is made possible by the fact that both members of the matched pair are removed from consideration when one of the pair withdraws. Finally, although the basis for interaction is again present when the testing variable is considered, the likelihood of serious confounding is not great. Indeed, there is reason for believing that the experimental group would be "sensitized" by pretesting more than would the control group, since they would be apt to see the testing as related to a course which they were about to take. Had the time which elapsed between pretesting and posttesting been reasonably short it would seem plausible that this difference in sensitization could differentially affect posttest performance. However, it is improbable that such is true when the two testing periods are separated by nine months. Furthermore, both groups were advised that their performance on the tests would have no bearing on course grades or any other such individual evaluation.

Two variables from the list suggested by Campbell and Stanley remain to be considered: intrasessional history and statistical regression. The second of these can be dealt with rather easily since the prerequisite for selection - regression interaction is not satisfied. In order for the selection process to produce differential regression one of the groups must be differentially selected in terms of extreme scores on the criterion variables. The presence of significant multiple correlations between the matching variables and pretest scores is ample proof that this eventuality did not arise in the selection process. Reference to Table 3 will clearly indicate that such proof is available for all of the criterion tests except one. It is not particularly surprising to find that the combination of intelligence quotient and grade point average is not a good predictor for pretest "score" on a value scale. However, there is an alternative source of evidence to indicate whether either group occupied an extreme relative position on this scale. This evidence is provided by a statistical test of the null hypothesis for difference in group mean "scores". Although the nature of the statistical test used is more appropriately presented later, it can be reported here that the results indicate the hypothesis of no difference cannot be rejected.

TABLE 3

MULTIPLE CORRELATION COEFFICIENTS, STANDARD ERRORS OF ESTIMATE AND F-TESTS OF SIGNIFICANCE FOR CORRELATION OF INTELLIGENCE QUOTIENTS AND CUMULATIVE GRADE POINT AVERAGE WITH PRETEST SCORES

Criterion Test	Correlation Coefficient	F-Value	Standard Error of Estimate
FAS I	.50	6.35	3.5
FAS II	.39	3.41	3.9
TOUS I	.50	6.59	2.1
TOUS II	.55	8.58	2.2
TOUS III	.63	13.01	2.8
Stan. Ach.	.39	6.19	3.54
Watson-Glaser	.67	8.52	15.8
Sci.&Cult.Ach.	.55	4.03	8.28
SOV	.21	0.87	6.81

$F_{.99} = 5.18$ df = 2, 39

$F_{.95} = 3.23$

Generally speaking, there is a wide variety of ways in which the selection differences that distinguish the experimental and control groups could give rise to differences in their intrasessional histories. In considering these possibilities several very important facts must be kept in mind if the intention is to explain away any superior performance shown by the experimental group. First, the differences in history must be reasonably constant in the sense of producing positive changes in the criterion variables where the experimental subjects are concerned and negative changes or no change at all in the case of the control subjects. Secondly, these differences must be of a kind that bear directly on the specific kinds of knowledge or understanding involved in the criterion variables and must be potent enough to overcome the inertia which, as we have stated previously, these variables possess. It follows that a considerable fraction of the subjects in the experimental group must experience a common and potent positive influence of the proper kind or must be individually subject to a number of different influences of this kind while few, if any, members of the control group are subject to the same or similar influences. It seems reasonable that the only really probable circumstance fitting this description would be the participation of a significant part of the experimental group in another course (or courses) which pursued some of the same objectives that the experimental course was designed to achieve. Within the laboratory school curriculum there are few if any such courses. The only realistic possibilities would be science courses. A study of the course schedule cards for the subjects involved in the study failed to reveal the necessary pattern. In fact, members of the control group were found to have completed a total of twenty-two semesters of science during the school year while members of the experimental group completed only six.

Beyond the suggestions given by Campbell and Stanley, there is one additional possibility of selection interaction which needs to be considered; that of selection - motivation interaction. Specifically, it is reasonable to suggest that a group of "seekers" is more highly motivated while responding to the test instruments used in the study than is a group doing so because they were conscripted. Although this situation may at first seem to present a serious problem, it can be dealt with rather easily. This follows from the fact that through the use of pretest - posttest differences or growth scores as criterion measures an overall between groups difference in motivation is compensated for. Given a member of the experimental group and a member of the control group who are equal in all respects except motivation, the experimental subject's score will be higher than that of the control subject by an amount "m". Assuming the motivation differential remains constant, on the occasion of posttesting the experimental subject's score will again be higher than that of his "equal" by at least an amount "m". However, the difference this time will theoretically consist of "m" plus some amount "t" resulting from the fact that the treatment has affected the score achieved by the experimental subject. When the two criterion (growth) scores are compared by taking the difference between them this difference will be "t", the contribution from differential motivation having been cancelled.

In addition to having now illuminated the precision, generalizability, and validity which are provided for in the design used, the previous discussion has accomplished one additional purpose. The description of the operations which were performed in obtaining the comparison groups and the elaboration of what can be considered reasons for equating the results achieved to those achieved by randomization have provided support for the contention that the comparison groups and the experimental framework used in this study are such that the comparisons involved are not affected by factors other than the independent variable. This being the case, the following statistical statement is justified in the case of each criterion variable -- assuming that the experimental treatment is totally ineffectual, the differences between the pretest - posttest growth scores of matched subjects are symmetrically distributed around a mean of zero (i.e., have a mean value of zero). The statistical tests of significance which will be used to evaluate the hypotheses presented in the statement of the problem will be based upon this statement.

Collection of Data

Since the investigator was fortunate enough to have the full cooperation of the laboratory school administration, it was possible to conduct both pretesting and posttesting in a systemmatic and organized fashion. Both control and experimental groups were given the necessary battery of tests under group testing conditions during essentially the same periods of time. In the case of pretesting this was the first week of the fall semester and for posttesting the final week of the spring semester. Within those periods of time, testing occasions were essentially randomly distributed and there is no reason to suspect that sources of bias such as time of day, day of week, proximity to special activities, etc., were incurred. All tests were administered under equally favorable working conditions and with faithful attention to the manual of directions provided by the test publisher. The same sorts of remarks regarding the purposes of the testing program were made to both the experimental and the control subjects. An equal effort was made to enlist their interest and cooperation. It was made clear to each group that the individual's best performance was desired but that no threat was involved. Subjects were advised that all results would be kept confidential, but would be made available to the student himself if he so desired. In view of the fact that testing of the kind in question is accepted as routine by students in the laboratory school there is no reason to suspect lack of serious effort on the part of either experimental or control subjects.

The scoring of all standardized tests was performed with scoring keys obtained from the publishers. The "Science Opinion Survey" and "Iowa Science and Culture Study Achievement Test" were scored in compliance with keys provided by the authors. After scoring had been completed, each subject's raw scores were recorded in a summary table. I. B. M. data cards were then prepared from this table and carefully verified against it later.

In addition to collecting data by traditional testing techniques, two less highly structured methods were used. Here the intention was not to obtain numerical data to be used in statistical tests of significance, but to gather information which could serve to supplement or verify the information obtained with formal tests.

In the first of these more "open-ended" techniques students were given a sheet of paper headed with the directions -- "Supply a list of words which serve to describe each of the following". Listed below these directions were the words "science", "culture", "experiment", "reasoning", and "art", with a space provided in each case for the student's response. This technique was used with both the experimental and control groups nine weeks after the opening of classes and again during the posttesting period.

The second technique was somewhat less structured than the "Descriptive Words Test" described above. In this case students were simply requested to write a short essay entitled "My View of Science". They were asked to make their comments without reservation and advised that their essays would not be used for grading purposes but were intended to "satisfy the curiosity of the instructor". To further insure candid responses, the subjects were free to submit their essay without identification. This procedure was used with the experimental group only and was carried out during the first and last weeks of the course.

Choice of the Method of Analysis

Although certain statistical techniques presently find wide use in educational research, the applicability of these techniques to any and all data is not thereby guaranteed. Each investigator must choose a method of statistical analysis which is most appropriate and advantageous for treating his particular data. The confidence which can be placed in the conclusions he draws depends upon the care which has been taken in making the choice and carrying out the selected statistical procedures. If the conditions of the experiment fit the conditions specified by normal curve theory, traditional or, more properly, parametric statistics can be appropriately employed. Under these circumstances the advantage of a high level of power is enjoyed -- the liklihood of rejecting false hypotheses is maximized. However, these conditions are not easily satisfied and it is seldom true that the degree to which they have been satisfied can be known with certainty. As a consequence, there is a broad range of uncertainty concerning the validity of the statistical analysis. Moreover, the actual probability of committing Type I errors is apt to be quite different from the stated level of significance (α). The magnitude of the difference depends upon the conditions violated and the seriousness of the violation. Obviously a flagrant violation removes all meaning from the conclusions reached in the analysis. In addition to avoiding the possible need for seriously qualifying the conclusions drawn from a statistical analysis of data, there is a second reason for carefully examining the feasibility of employing parametric tests of significance and methods of increasing precision. In the process of the examination information may be obtained which indicates the need for more detailed study of the treatment effect once the main analysis has established its presence. For these reasons the following section presents an investigation of the pertinent data parameters.

The test of significance most commonly used in the analysis of data obtained in an experiment designed such as the present one employs a pretest-posttest score difference or gain score as the criterion variable. Whether or not there is a significant difference between the treatment and control groups on these gain scores is then determined with the t-test. In the special case of only two conditions (treatment and no treatment) the use of the F-test would yield identical results. According to Campbell and Stanley this is not, however, the preferred approach (Campbell and Stanley, 1963, p. 193). Instead, they recommend the use of an analysis of covariance with the pretest scores as the covariate. In either case similar mathematical conditions must be met. However, before entertaining the question of whether or not valid use of an analysis of covariance can be made, a previous question must be answered. Will there be any advantage to using this method of analysis? The answer to this question depends upon whether or not a reasonably high correlation exists between the criterion variables and the proposed covariates. If no such correlation exists the analysis of covariance is unable to provide the increased precision which otherwise recommends its use. Moreover, the possibility of using the regression of posttest scores on pretest scores in any more detailed study of the treatment effect (pursuant to the main analysis) is removed. The Pearson product-moment correlations of pretest scores on posttest scores for the combined control and experimental groups are presented in Table 4. Since no pretest was administered for the "Science Opinion Survey", no correlation coefficient can be given in this case.

TABLE 4

PRODUCT-MOMENT CORRELATION COEFFICIENTS BETWEEN PRETEST AND POSTTEST SCORES FOR THE ENTIRE EXPERIMENTAL POPULATION

Criterion Test	Correlation Coefficient
FAS I	.55
FAS II	.59
TOUS I	.57
TOUS II	.69
TOUS III	.62
Stan. Ach.	.75
Watson-Glaser	.57
Sci.&Cult.Ach.	.63
SOV	.82

It is apparent from Table 4 that the correlation prerequisite for a gainful use of the analysis of covariance can be met. Hence the possibility of satisfying the requirements for its valid application will be examined.

An examination of several commonly used texts in experimental design and analysis reveals that the most critical requirements for the analysis of covariance can be summarized as follows (Walker and Lev, 1953; Tate, 1955; Lindquist, 1953):

1. The populations involved in the experiment must be unbiased samples from the same parent population.

2. The distribution of criterion measures for each of these populations must be normal.

3. The variance of these distributions must be the same for all populations.

4. The regression of the criterion variable on the covariate or predictor variable must be the same for all populations.

Little needs to be said regarding the first requirement since it has been discussed previously in a slightly different connection. This requirement really incorporates two conditions, one to insure representativeness and another to protect against lack of equivalence between groups. Both of these conditions were discussed earlier with the decision being that, while the second condition can be met, the first cannot. But the first condition is of no consequence to the justifiability of using the analysis of covariance. Furthermore, so long as the inferences from the analysis are restricted to the sample under study or to a hypothetical parent population, no violation of the principles of good research has been made.

There is no reliable means for determining whether or not a population of measures is normally distributed. Although a test using chi square can be performed, it is virtually pointless when the number of cases involved is small. According to Walker and Lev it is usually sufficient to examine the grouped frequency distribution of scores and make an intuitive decision as to whether or not the distribution appears reasonably normal (Walker and Lev, 1953, p. 201). Ray suggests that departures from normality in terms of skewness (asymmetry) and kurtosis (flatness or peakedness) can be evaluated with the use of calculations involving standard or "z" scores[1] (Ray, 1960, pp. 71-72,). It is Walker and Lev's opinion that such measures tend to be unreliable when used with small samples, however (Walker and Lev, 1953, p. 183). In view of these opinions, the decision was made to use the IBM 7044 computer to obtain frequency distributions, histograms, and indicies of skewness and kurtosis for the entire collection of criterion score distributions. This mass of information was then carefully examined to detect any consistent indications of serious departures from normality. Although the results were somewhat inconclusive, no clearly pronounced violations of the normality requirement appeared. In view of its volume and somewhat tenuous nature, the graphical and numerical information used in the distribution analysis will not be presented. Since the authors of all of the statistical references consulted in the course of the study agree in claiming that some departure from normality does not invalidate parametric statistics, the results of the above analysis do not seem sufficient to reject their use.

Unlike the situation regarding the normality requirement, it is

---

[1]The measure of skewness is given by $\Sigma z^3/N$ and that for kurtosis by $\Sigma z^4/N$.

possible to perform a statistical test for the homogeneity of variances stipulated by requirement "3". In the case of related samples a t-value is computed using the following equation (Walker and Lev, 1953, p. 190):

$$t = \frac{(s_2^2 - s_1^2)\ (N - 2)^{1/2}}{2s_1s_2\ \ (1 - r_{12}^2)^{1/2}}$$

s_1^2 = sample variance for first group of scores

s_2^2 = sample variance for second group of scores

s_1 = sample standard deviation for first group of scores

s_2 = sample standard deviation for second group of scores

r_{12} = correlation between the two sets of scores

The results of having performed this calculation for the control and experimental group posttest scores are presented in Table 5.

In view of the fact that the test of homogeneity of variance is quite sensitive to non-normality, the information in Table 5 is in reality indicative of whether departures from the conditions stated in either or both requirements "2" and "3" exist. Thus when it is necessary to reject the null hypothesis, one cannot be certain which of the two conditions have been violated or if in fact both have been. But, of course, knowledge of which of the three circumstances exists is immaterial. Since in this particular situation the consequences of accepting the null hypothesis when it is false (Type II error) are more serious than rejecting it when it is true, a level of significance (α) of .05 has been chosen. This choice is supported further by the fact that the computed t-values include the errors introduced by the coefficient of correlation term. Since correlation coefficients gained from small samples are apt to be somewhat unreliable, it may well be that the t-values have been falsely attenuated in some cases.

TABLE 5

t-VALUES AND TESTS OF SIGNIFICANCE FOR HOMOGENEITY OF VARIANCE BETWEEN CONTROL AND EXPERIMENTAL GROUP POSTTEST SCORE DISTRIBUTIONS

Criterion Test	t-Value
FAS I	1.84
FAS II	2.29
TOUS I	2.1[illegible]
TOUS II	2.31
TOUS III	0.48
Stan. Ach.	0.67
Watson-Glaser	2.06
Sci.&Cult.Ach.	2.09
SOV	1.88

df = N - 2 = 19 CR: $-1.73 \geq t \geq 1.73$ $\alpha = .05$

Table 5 indicates that in all but two cases it would be impossible to assume that the conditions of normality and homogeneous variance exist. This conclusion suggests that the notion of making use of an analysis of covariance should now be held with very strong reservations. However, according to Lindquist homogeneity of regression is the most critical requirement to be satisfied in practice (Lindquist, 1953, p. 330). Furthermore, the discovery of heterogeneous variances indicates that, if a significant main treatment effect is operating, there is a need for investigating its uniformity. Knowledge of the relationship between the group regression lines of posttest scores on pretest scores assists in this investigation.

According to Tate (1955, p. 522) there is no really convenient test for the homogeneity of regression. This being the case, he recommends that the investigator first form the regression equations for the total group under study and for each comparison group. An inspection of these equations is then used as the basis for judging whether or not the assumption of homogeneous regression is tenable. The equations for the lines of regression of posttest scores on pretest scores are presented in Table 6. It will be recalled that the term representing the slope of the line is the coefficient of regression. These coefficients were obtained from the pretest versus posttest Pearson product-moment correlation coefficients (Blommers and Lindquist, 1960, p. 420). As before, the lack of pretest scores for the "Science Opinion Survey" explains the absence of that set of equations in the table.

TABLE 6

EQUATIONS FOR THE TOTAL AND WITHIN GROUPS REGRESSION OF POSTTEST SCORES ON PRETEST SCORES

Criterion Test	Group	Equation
FAS I	Total	$Y'_T = .60\ (X_T - 28.1) + 28.3$
	Experimental	$Y'_E = .42\ (X_E - 28.1) + 30.0$
	Control	$Y'_C = .71\ (X_C - 28.1) + 26.6$
FAS II	Total	$Y'_T = .81\ (X_T - 29.6) + 29.7$
	Experimental	$Y'_E = .78\ (X_E - 29.5) + 31.0$
	Control	$Y'_C = .84\ (X_C - 29.7) + 28.5$

TABLE 6-- Continued

Criterion Test	Group	Equation
TOUS I	Total	$Y'_T = .74\ (X_T - 13.18) + 13.52$
	Experimental	$Y'_E = .70\ (X_E - 13.0) + 14.3$
	Control	$Y'_C = .94\ (X_C - 14.0) + 12.7$
TOUS II	Total	$Y'_T = .86\ (X_T - 12.6) + 12.8$
	Experimental	$Y'_E = .68\ (X_E - 13.0) + 14.3$
	Control	$Y'_C = .82\ (X_C - 12.2) + 11.4$
TOUS III	Total	$Y'_T = .70\ (X_T - 14.0) + 15.1$
	Experimental	$Y'_E = .90\ (X_E - 13.7) + 16.7$
	Control	$Y'_C = .66\ (X_C - 14.3) + 13.6$
Stan. Ach.	Total	$Y'_T = .54\ (X_T - 47.1) + 47.8$
	Experimental	$Y'_E = .67\ (X_E - 47.3) + 47.7$
	Control	$Y'_C = .49\ (X_C - 47.0) + 48.0$
Watson-Glaser	Total	$Y'_T = .59\ (X_T - 72.1) + 76.5$
	Experimental	$Y'_E = .58\ (X_E - 73.0) + 81.1$
	Control	$Y'_C = .57\ (X_C - 71.3) + 71.9$
Sci.&Cult.Ach.	Total	$Y'_T = .86\ (X_T - 22.5) + 25.8$
	Experimental	$Y'_E = .65\ (X_E - 23.2) + 29.7$
	Control	$Y'_C = .89\ (X_C - 21.9) + 21.9$
SOV	Total	$Y'_T = .88\ (X_T - 33.7) + 36.0$
	Experimental	$Y'_E = .76\ (X_E - 34.7) + 38.9$
	Control	$Y'_C = .87\ (X_C - 32.8) + 33.0$

An inspection of Table 6 reveals that for some of the criterion variables the condition of parallel regression lines is clearly met. However, in several cases the opposite seems to be true. Under the circumstances it would seem wise to abandon the method of covariance analysis as being not generally applicable. Although its use with several of the criterion variables might be defensible, the combined evidence points to a serious risk of invalidity. The existence of evidence for heterogeneous variance and regression remains important apart from this decision and will be dealt with in more detail in Chapter IV.

Fortunately, the rejection of parametric statistics does not create a particular problem since it is possible to employ tests of significance which are not based upon normal curve theory. In particular, we refer to so-called "distribution free" or nonparametric statistics. Nonparametric statistical tests are, by definition, tests whose model does not specify conditions concerning the parameters of the population distribution under study. Thus in using them one need not deal with questions of normality, equal variances, etc. Furthermore the probability of committing Type I error is equal to or less than the chosen level of significance. With a normal curve theory test, on the other hand, the chances of committing an error of the first kind may be greater or less than the chosen level of significance, depending upon the form of the population distribution. There is generally no way of knowing the actual control which is exerted over falsely rejecting a hypothesis of no difference. Of course, the advantages of nonparametric statistics do not come without some cost. In return for relative freedom from restrictive conditions and uncertain control over Type I errors, one must yield power. In most cases nonparametric tests of significance are less apt to reject hypotheses of no difference when they are false. This, perhaps more than any other reason, explains why nonparametric tests have failed to gain widespread use. However, when there is reason to believe that parametric conditions cannot be met, the investigator is obliged to turn to nonparametric analysis. Having done so he can be assured of testing his null hypotheses with the stringency he claims and can be certain that those which he rejects were indeed false.

It is doubtless true that the use of an analysis of covariance with pretest scores as covariates in studies employing matched pairs is recommended for purposes of increasing experimental precision. The increase in precision which actually occurs depends, of course, upon the degree to which pretest scores within matched pairs are made equal by the matching procedure. If indeed a high degree of equality has been achieved, the gain in precision is slight and a test of significance applied to the difference between the growth scores of control and experimental subjects is, considering the imprecision in covariance adjustments, capable of equal precision. At several points in the previous discussion occasions arose for indicating that in the present study the evidence points toward a successful "leveling" on pretest scores. Additional evidence will be presented in the following chapter. But for the present we shall take the position that a comparison of growth scores provides a precise method of analysis and therefore is an adequate method to be used. This being the case, the null hypothesis presented in the statement of the problem can now be stated in a form which better fits the actual design and method of analysis employed in the study. Since there seems to be no necessity for repeating all seven hypotheses, suffice it to say that the form is as follows: There is no significant difference between the pretest-posttest growth scores achieved by experimental and control subjects on a given criterion variable, as measured by the criterion instrument.

The non-parametric test of significance which will be used for testing the null hypotheses is the Wilcoxon Matched-pairs Signed-ranks Test. According to Ray, empirical evidence has indicated that the results obtained with this test agree closely with the results of the parametric analysis of variance (Ray, 1960, p. 80). A brief discussion of the Wilcoxon Test will introduce the following chapter.

CHAPTER IV

ANALYSIS OF THE DATA

Main Analysis

Since a number of statements have been made previously regarding the extent to which matching was achieved on the criterion variables, it will be appropriate to begin the analysis of data with a test of significance for the difference between control and experimental subjects' pretest scores. This will have the further advantage of introducing the use of nonparametric statistics. Siegel (1956) has provided an excellent reference for those who desire background information.

The null hypothesis to be tested is that there is no difference between the mean of the control group pretest scores and the mean of the experimental group pretest scores. Or, stated in terms more to the point and more appropriate for the test of significance to be used; the differences between the pretest scores of subjects in matched pairs are symmetrically distributed around a mean value of zero (i.e., have a mean of zero). This hypothesis will be tested in the case of each criterion variable except "Science Opinion Survey" scores. As has been stated earlier, pretesting with this instrument was not possible. The Wilcoxon Matched-pairs Signed-ranks Test will be the test of significance used.

Although the signed-ranks test is straightforward in nature, those accustomed to the more commonly used parametric tests may find it foreign. A somewhat simpler but closely related nonparametric test of significance, the Sign Test, can serve to make the nature of the signed-ranks test clear. If indeed there is a mean difference of zero between the pretest scores of matched subjects, the probability that a given control subject's score will exceed that of the matched experimental subject is equal to the probability that the reverse is true, both probabilities having the value .5. In applying the sign test the differences between the matched subjects' scores are found (always subtracting in the same order) and the number of cases wherein the difference is positive and the number wherein it is negative are tabulated. When the difference is zero the pair is discarded from the analysis. If the null hypothesis is true, about fifty percent of the differences should be positive and fifty percent negative, for the same reason that a tossed (unbiased) coin should land on either face approximately fifty percent of the time. The test of significance consists of determining whether or not the actual outcome deviates too far from what would be expected under the laws of probability; of determining whether or not too few differences of one sign occur. When a prediction as to which sign will occur less frequently is involved, a one-ended test is used and the expectation is that "too few" differences of that sign will occur. What is meant by "too few" depends of course on the chosen level of significance (α). In a two-ended test the critical region consists of all outcomes whose associated level of probability is equal to or less than $\alpha/2$. In a one-ended test the value is simply α. The probability associated with a given outcome can be computed by using the general formula for the binomial probability distribution or it can be read from statistical tables (Meredith, 1967, p. 281).

The Wilcoxon Matched-pairs Signed-ranks Test is a logical extension of the Sign Test. It makes use of not only the direction of differences between the scores of matched subjects, but their magnitude as well. Larger differences are assigned greater weight in determining the value of the test statistic through a process of ranking. It is important to note that in converting differences to ranks allowance is made for ordinary measurement errors. The criterion instrument is assumed to be precise enough to permit decisions of which of two sets of matched subjects differ more in their performance. The magnitude of the separation between the two differences is not introduced into the calculation. It is for reasons such as this that nonparametric tests tend to be less powerful than their parametric counterparts. However, there is cause for questioning whether the margin of power which parametric tests gain by assuming precise measurement is always well founded.

In carrying out the signed ranks test, score differences within matched pairs are determined as before. They are then ranked in order of absolute magnitude with the greatest difference being assigned the largest rank. Cases of zero difference are discarded. Each rank is assigned a positive or negative sign depending upon the direction of the difference to which it corresponds. The sum of the ranks of each sign is then taken. Under the null hypothesis these two sums of ranks should be approximately equal. But if either sum is too small, rejection of the null hypothesis is indicated. The actual test of significance involves operations fundamentally like those involved in the Sign Test. The value of either sum of ranks, positive or negative,can be used as the test statistic (T). Depending upon the direction in which the difference in scores was taken, each is an index of the degree to which one comparison group's scores tended to exceed the other's. Ordinarily the value of the smaller sum is selected for computational simplicity. The critical region is then defined by all T-values whose probability of occurence under the null hypothesis is less than $\alpha/2$ for a two-ended test and less than α for a one-ended test. When no prediction has been made regarding which set of scores is most apt to exceed the other(and a two-ended test is used), it is immaterial whether the T-value is the sum of positive ranks or negative ranks. However, when a prediction has been made (and a one-ended test is used), the null hypothesis is rejected only if the T-value which falls in the critical region is the sum of the ranks whose sign is opposite to the one which reflects a larger score in the set favored by the prediction. Tables of T-values and their corresponding probabilities for a variety of sample sizes can be found in several sources (Siegel, p. 254; Meredith, pp. 264-65).

Table 7 presents pretest means, T-values, sample sizes and tests of significance for the difference between the pretest scores obtained by matched pairs on the criterion tests. A double-ended test has been used since no prediction is being made as to the probable direction of departure from the null hypothesis. The variation in sample sizes is explained by the discarding of cases with tied scores.

The results in Table 7 indicate that none of the hypotheses of no difference between pretest scores on the criterion tests can be rejected at ordinary levels of significance. These results corroborate the previous evidence indicating a successful choice of matching criteria and also support the claim that subjects were well matched in terms of these variables. Any statements made previously and qualified in terms of the assumption that subjects were initially equated on the criterion variables seem to now be placed on a foundation of objective evidence.

TABLE 7

PRETEST MEANS AND TESTS OF SIGNIFICANCE FOR DIFFERENCES BETWEEN WITHIN PAIRS PRETEST SCORES ON CRITERION TESTS

Criterion Test	Pretest Mean Experimental	Pretest Mean Control	Number of Subject Pairs	T-Value
FAS I	28.14	28.10	20	113.0
FAS II	29.48	29.71	21	123.5
TOUS I	13.00	13.95	19	126.0
TOUS II	13.05	12.19	17	57.5
TOUS III	13.71	14.29	20	127.0
Stan. Ach.	47.29	47.00	18	97.5
Watson-Glaser	72.95	71.29	21	88.5
Sci.&Cult.Ach.	23.23	21.86	20	82.0
SOV	34.67	32.81	19	75.0

N	$T_{.01}$	$T_{.05}$
17	24	35
18	28	40
19	32	46
20	38	52
21	43	57

From this point it is a short step to performing the tests of significance appropriate to the seven null hypotheses set forth in the statement of the problem. It will be recalled that each of the first three of these hypotheses is to be evaluated with data collected through the use of two different evaluative instruments. Thus a total of ten tests of significance are to be performed in the main analysis. One of these ten must be performed on a slightly different basis, however, since pretest data on the variable involved, the "Science Opinion Survey", is not available. This test will be reported separately.

In the previous discussion and application of the signed-ranks test the null hypothesis involved was concerned with differences in raw scores. The main analysis, however, employs pretest - posttest growth scores as criterion variables. Thus it is differences in growth scores rather than differences in raw scores that is now the point of concern. But no particular change in procedure is required and the explanation of the signed-ranks test can be made applicable by simply substituting "growth scores" for "scores". Of course, a single-ended test is now appropriate in view of the fact that a prediction of the direction of departure from the null hypothesis is being made. According to the null hypothesis the differences between the growth scores of subjects in matched pairs are symmetrically distributed around a mean value of zero. Since the difference in growth scores was determined by subtracting control subjects' growth scores from those of the subjects exposed to the treatment, the alternative hypothesis is that the distribution of differences is not symmetrically distributed about a mean value of zero but is negatively skewed with a mean greater than zero.

The mean growth scores, T-values, sample sizes and tests of significance for the null hypotheses concerning growth on all criterion variables except the "Science Opinion Survey" score are presented in Table 8. An inspection of the results of these tests of significance will reveal that all of the hypotheses tested can be rejected at the .01 level except two. One of these two, that for the Watson-Glaser Critical Thinking Appraisal score, can be rejected at the .05 level. The results for the Stanford Advanced Science Achievement Test score clearly indicate that there is no significant difference in control group and experimental group growth scores on this variable. In fact the mean growth score for the control group is larger than that for the experimental group.

TABLE 8

T-VALUES AND TESTS OF SIGNIFICANCE FOR NULL HYPOTHESES CONCERNING GROWTH SCORE DIFFERENCES

Criterion Test	Mean Growth Score Experimental	Control	Number of Subject Pairs	T-Value
FAS I	1.81	-1.47	20	36.0
FAS II	1.48	-1.24	19	24.5
TOUS I	1.33	-1.23	21	31.0
TOUS II	1.24	0.81	20	42.5
TOUS III	3.00	0.71	19	15.0
Stan. Ach.	0.43	0.95	19	79.0
Watson-Glaser	8.19	0.62	21	50.0
Sci.&Cult.Ach.	6.48	0.00	19	9.5
SOV	4.24	0.19	21	32.0

N	$T_{.01}$	$T_{.05}$
19	38	54
20	43	60
21	48	66

Since the results on the test of significance for difference in growth in critical thinking ability border on rejection at the .01 level it was decided to repeat the test using a higher powered nonparametric method, the Randomization Test for Matched Pairs. This particular test represents the most powerful nonparametric technique available. However, when the number of pairs involved is at all large, the computation involved is beyond reason unless a high speed computer is used. This fact will become clear with a few words of description concerning the test. If indeed the treatment is ineffectual as far as fostering growth in critical thinking, then the difference between the growth scores of the control and experimental members of a matched pair depends simply upon which subject happens to be placed in which group; whether the difference is +D or -D is a matter of chance. If we proceed on this assumption, there are as many equally likely sets of growth score differences as there are ways to generate new ones by negating one or more of the obtained differences. It so happens that the number involved is 2^{21} or 2,097,152. The Randomization Test for Matched Pairs consists of generating all 2^{21} sets, calculating the sum of growth score differences for each one, and forming a critical region consisting of the $(\alpha)2^{21}$ most extreme (largest, if we subtract control subject growth score from experimental subject growth score) sums. The test statistic used with this critical region is the sum obtained for the actual set of growth score differences.

A program for the randomization test was written for the IBM 7044 computer and used with the data from the critical thinking test. The outcome was as follows. The sum of growth score differences in the actual data, 159, was equaled or exceeded in 14,100 cases out of the possible 2,097,152. Thus the actual sum of differences was within the largest 14,100 cases in the mathematically generated population of sums. Since this amounts to saying that it falls within the $(.0067)2^{21}$ most extreme sums, the null hypothesis can be rejected at the .01 level of significance.

In studying Table 8 one notes that in three cases the mean "growth" shown by the control group on the criterion variables is negative. This fact may raise a question as to whether or not rejection of the null hypotheses on those three occasions is a result of the "shrinkage" of control subjects rather than the growth of treatment subjects. One rather simple way to resolve this question is to test the hypothesis that the differences between the pretest and posttest scores of subjects in the treatment group are symmetrically distributed around a mean of zero. In effect, this approach uses the treatment subjects as their own controls. For this reason it has value apart from helping to resolve the question raised above. It provides an evaluation of the treatment in terms of the population of "seekers" themselves, the group whose response to the experimental course is the matter of specific concern. Of course, there are design weaknesses involved. But these are not so serious as to render the results valueless. The results of the analysis are set forth in Table 9. Included in the table are the results for the same analysis when performed with control group scores.

The contents of Table 9 indicate that with the data from the experimental group, the hypothesis of no difference between pretest and posttest scores can be rejected at the .01 level for every criterion test except the Stanford Achievement Test. On the other hand, rejection is not possible at the .05 level for any of the control group test scores.

Although a comparison of pretest - posttest growth scores is not possible where the "Science Opinion Survey" is concerned, the resulting loss of precision should not be overly serious. A number of pieces of evidence have been presented to indicate that the comparison groups were effectively matched on a subject to subject basis, both in terms of the matching variables themselves and in terms of the variables measured by nine of the criterion tests. It does not therefore appear unreasonable to believe that they were equated with regard to the variable measured

TABLE 9

T-VALUES AND TESTS OF SIGNIFICANCE FOR NULL HYPOTHESIS CONCERNING PRETEST-POSTTEST SCORE DIFFERENCES

Criterion Test	Experimental		Control	
	Number of Subject Pairs	T-Value	Number of Subject Pairs	T-Value
FAS I	16	23.0	17	112.5
FAS II	19	36.0	17	127.5
TOUS I	17	19.5	17	119.0
TOUS II	16	24.0	17	104.0
TOUS III	19	8.0	18	124.5
Stan. Ach.	19	59.0	18	90.5
Watson-Glaser	19	8.0	20	79.5
Sci.&Cult.Ach.	21	1.0	20	115.0
SOV	18	26.	19	105.0

N	$T_{.01}$	$T_{.05}$
16	24	36
17	28	41
18	33	47
19	38	54
20	43	60
21	48	66

by the "Science Opinion Survey", a variable which is similar to those measured by the other tests. The tests most similar in nature to the "Science Opinion Survey" are the "Iowa Science and Culture Achievement Test" and the processes of science subtest of the Test On Understanding Science. The correlation between posttest scores on each of these two tests and the "opinion survey" are, respectively, .58 and .57. In the light of these facts it seems legitimate to consider comparison of post-treatment test scores closely analogous to a comparison of pretest-post-test growth scores.

The Wilcoxon Matched-pairs Signed-ranks Test was used to test the hypothesis of no difference between control and experimental subjects' scores on the "Science Opinion Survey" with the following results. The data from a total of nineteen subject-pairs could be used in the calculations. The computed T-value was 18. Since the probability associated with this value is less than .005, the lowest value given in the table used, the hypothesis can clearly be rejected at the .01 level.

The difficulty which remains is that no estimate is available for the extent to which the experimental subjects gained in understanding of science, as measured by Kimball's test. However, it is possible to form a well founded opinion concerning the probable magnitude of growth by comparing the performance of the experimental subjects with that reported by Kimball (1965) for the well qualified science teacher, scientists, philosophy graduates and graduate science majors included in his study. This comparison is presented in Table 10.

Unless it is contended that the subjects in the experimental group possessed a remarkable understanding of science previous to the treatment, the relative level of performance indicated in Table 10 points to a sizeable growth in understanding during the treatment. In view of this fact, it appears that the results obtained with the "Science Opinion Survey" agree with those previously reported for subscore III of the _Test On Understanding Science_.

TABLE 10

COMPARISON OF EXPERIMENTAL GROUP'S PERFORMANCE ON THE "SCIENCE OPINION SURVEY" WITH THAT OF OTHER GROUPS

Comparison Group	Number of Items on which Experimental Group Showed a Higher Percentage of Correct Responses	Percentage of Total Number of Test Items* Represented by the Number in Column 2
Philosophy Graduates	20	69%
Science Majors	22	76%
Science Teachers	23	79%
Scientists	20	69%

*Total number of test items = 29

Subsidiary Analysis

Evaluation of Qualitative Data. In addition to the data collected through the use of the criterion instruments, two other devices, the "Descriptive Words Test" and the "My View of Science" essays, were used as sources of supplementary evaluation. However, the information provided by such "open ended" methods does not lend itself to statistical analysis and must be treated in a less formal way.

In view of such problems as interpretive bias, there is little which can be done in the way of objectively analyzing the "My View of Science" essays. However, this is not to say that they are of no value in apprais- ing the effect of the treatment. To the extent that the candid responses of the subjects are involved, they provide a kind of insight which could not be gained in any other way. Since the students in question tend to be open in their opinions and were provided no reason to be otherwise, the necessary candidness, and therefore validity, can be assumed. The entire collection of pretreatment and posttreatment responses are reproduced with- out change in Appendix D. It is suggested that they be read with a focus on the differences between pretreatment and posttreatment opinions.

Since the results obtained from the "Descriptive Words Test" repre- sent a collection of well over two thousand words, these responses were given a preliminary examination to determine their potential value as evaluative data. The results of this screening process indicated that there was insufficient justification for presenting the entire collection. Only in the case of the stimulus word "science" were there clear enough trends to support generalizations. Table 11 contains a summary of the initial and final responses given by the experimental group to this stim- ulus word. Table 12 presents a similar summary for the control group. An attempt has been made to present the data in a form which lends itself to ready interpretation. It is hoped that this presentation will encour- age the drawing of inferences beyond the limited number permitted by the investigator's fear of bias. It should be kept in mind that the initial and final responses presented were obtained at the nine week point and at the end of the school year, respectively.

TABLE 11

SUMMARY OF EXPERIMENTAL GROUP RESPONSES ON THE "DESCRIPTIVE WORDS TEST"

Response Categories

Quality	Content and Product	Performance and Process	Unclassified
Repeated objective(3,2) curiosity(4,2) dynamic(2,1) imagination(3,1)	Repeated *theories(9,8) *nature([illegible],3) *facts(6,0) *knowledge(6,-1) questions(6,3) explanations(5,4) *laws(7,6) truth(2,-1) *ideas(1,0) hypotheses(2,1)	Repeated *experiment(15,5) observations(7,5) *research(4,3) thought(3,2) reasoning(2,0) *study(3,0) learning(1,0) *discovery(5,4)	Repeated culture(6,4) art(4,3) *scientists(1,0)
Omitted absolute abstract impersonal new infinite precise intellectuals repeatable modern mentality spirit of freedom	Omitted concrete material formulas inventions assumptions	Omitted answering try	Omitted field method technology mind lab test tube acid gray
Added value free(5) pure(4) rational(3) universal(2) empirical(3) creative controlled systemmatic aesthetic	Added relationships(5) understanding(3) generalizations(2) solutions(2) problems	Added prediction(7) why?(7) how?(4) *questioning(4) inquiry(2) *investigation induction(4) deduction(3) logic(2) comparing correlating classify validation defining	Added none

Legend:

(n) -- frequency of response
(x,y) -- x: frequency on posttest; y: change in frequency
* -- appears in control group posttest list

TABLE 12

SUMMARY OF CONTROL GROUP RESPONSES ON THE "DESCRIPTIVE WORDS TEST"

Response Categories			
Quality	**Content and Product**	**Performance and Process**	**Unclassified**
Repeated *required(1,0) *important(1,0)	Repeated theories(4,2) *chemicals(1,0) *space(1,0)	Repeated experiment(4,-7) research(3,-1) study(4,-2) discovery(2,-2)	Repeated scientists *technology(3,2) *lab(1,-1) *work(2,1)
Omitted imagination systemmatic variety	Omitted laws(2) advancement(2) truth matter principles evolution atmosphere solar system sound time power improvement	Omitted thought(1) learning(2) questioning(2) inquiry(1) classification(1) problem solving(1) explain(1)	Omitted culture job class
Added *ingenuity *mythical	Added *biology(3) *physics(2) facts(2) nature(2) knowledge(2) *electron *universe *ideas *plants *concepts *light *speed *rocks *animals *earth	Added questions(2) *answers investigation *search *explore	Added *test *professor *school *world *future *hope *destruction *test tube

Legend:

(n) -- frequency of response
(x,y) -- x: frequency on posttest; y: change in frequency
* -- does not appear in experimental group posttest list

The data in Tables 11 and 12 reveal the following general trends:

1. Initial Responses

   a. The experimental group tended to give a disproportionately greater number of responses reflecting the "quality" of science and scientists than the control group.

   b. The control group responses in the "content and product" category tended more toward concrete terms dealing with the topics of science than the experimental group responses did. The experimental subjects responded with a greater number of more abstract terms descriptive of the nature of scientific knowledge.

   c. Although there seems to be little difference in the responses given by the two groups in the "performance and process" category, the frequency with which words common to both lists appear in the control group list is slightly greater than in the experimental group list.

   d. The collection of "unclassified" responses in both tables contain a rather large number of "neutral" responses such as "lab", "job", "field", "work" and "gray". These responses seem to be of a free association character.

2. Initial Versus Final Responses

   a. In the case of the "quality" category the overall change in character which occurs in the control group list is minor. However, the response words in the experimental group list which project an error-free non-human character to science and scientists are omitted. The words retained, as well as those added, seem to represent a more realistic picture of the unique character of science and scientists.

   b. Although a number of the concrete words in the control group list under "content and product" are excluded from the final responses, they are replaced by words equally lacking in breadth of meaning. The experimental group's list of words shows a trend toward responses reflecting the logical, predictive, and explanatory nature of science. A number of key words such as "theory", "law", "explanation", etc., occur with increased frequency.

   c. The similarity of words in the "performance and process" category is destroyed by the addition of a considerable number of realistic new responses to the experimental group list and the omission of some equally realistic words from the control group list. The experimental subjects introduce a number of responses which reflect the curiosity drive in science. Again, a trend toward reflecting the logical and predictive nature of science is evident.

d. The "neutral" responses in the "unclassified" category are largely removed from the experimental group's list. On the other hand, the number of such responses increases among control subjects.

Evaluation of Evidence Regarding Nonadditivity of Treatment Effect. During the examination of the "goodness of fit" between the data and the normal curve model, evidence of heterogeneity of sample variances was discovered. At the time, the point of concern was the possibility of employing parametric statistics in the evaluation of the null hypotheses. Heterogeneity of variance is also significant in that it suggests the presence of a nonadditive treatment effect. In other words, there is an indication that the treatment does not result in a uniform change in criterion performance for all subjects. But the test of significance for the hypothesis of no difference in sample variances is also quite sensitive to lack of normality in the score distributions and therefore rejecting the hypothesis of homogeneous variance does not necessarily indicate a nonadditive effect. It does serve to alert the investigator to the presence of additional evidence for nonadditivity, however. As was stated earlier, a lack of between-groups homogeneity in the regression of posttest scores on pretest scores is one such type of evidence. If a treatment produces the same increment in response for all subjects, regardless of their initial level of response, the relationship between pretest scores and posttest scores should be the same for a group of subjects exposed to the treatment and a group not so exposed. When, as in the present study, this is not true, the indication is that the treatment effect is differential along the pretest score dimension; it is nonadditive and the nonadditivity is correlated with pretest score. A logical extension of this argument yields the conclusion that if the correlation between a given variable and pretest scores is not the same as with posttest scores, the treatment involved is differential in terms of that variable. Of course, when a treatment effect is differential in terms of several variables this may be a mere result of the fact that those variables are correlated. Obviously, such is not the case if the direction of the differential effect and the direction of the correlation are not the same. When several correlated variables all undergo a similar pretest-posttest change in their individual correlations with a criterion variable it is unfortunately not possible to determine which is (are) causally related to the differential treatment effect.

Another approach to the question of nonadditivity would be to examine the relationship between a given variable and growth scores. A relatively high correlation would indicate an effect which is differential in terms of the variable involved. The direction of the correlation would indicate the general nature of the departure from additivity; positive correlations suggesting that the treatment favors those who rank high in the variable and negative correlations suggesting the reverse. Thus rather than examining the homogeneity of regression of posttest scores on pretest scores, the correlation between pretest and growth scores could be studied. Preferably, both sources of evidence would be employed since an analysis based upon correlations in small samples is open to the question of reliability. Although parametric statistics provide approaches which are perhaps more reliable, such methods are yet to be developed in nonparametric form. Regardless of the approach employed, one fact should be kept in mind; nonadditivity of treatment effect may be apparent rather than real. Often times, for instance, the presence of a "ceiling effect" in a test instrument being employed is mistaken for a nonadditive treatment. Since measures of attitude, aptitude, and achievement are relative and little is known about the equality of units or the zero point on the measurement scale used, such errors are difficult to avoid.

In view of what has been said above, it is evident that it will not be possible to draw infallible conclusions regarding the nonadditivity of the effect produced by the "Science and Culture treatment". However, to the extent that somewhat limited ex post facto evidence permits, reasoned speculation is possible. A summary of the available evidence is presented in Table 13. In order to make a meaningful comparison of product moment correlation coefficients possible, coefficients have been transformed to z_r equivalents through the use of Fisher's transformation (Tate, 1962, p. 423).

An examination of the regression coefficients presented in Table 13 indicates that a suggestion of nonadditive effect is present in the coefficients for several test scores; the first subscore in <u>Facts About Science</u>, the second and third subscores in the <u>Test On Understanding Science</u> and the "Science and Culture Achievement Test" score. In the first of these four cases there is a rather strong suggestion of a more pronounced treatment effect among subjects who tend to obtain low pretest scores. The same sort of nonadditivity is indicated in the other three cases but to a much lesser extent. The hint provided by the regression between the scores on the <u>Facts About Science</u> subtest is supported by a high negative z_r for the correlation between pretest score and growth score. Although this correlation may merely reflect the presence of a ceiling effect on the test, it seems rather large to be accounted for in this way. This

TABLE 13

SUMMARY OF EVIDENCE REGARDING NONADDITIVITY OF TREATMENT EFFECT

Criterion Tests	Regression Coefficients for Regression of Posttest Scores on Pretest Scores: Control	Combined	Experimental	Correlation between Pretest Score and Growth Score	z_r Equivalents for Multiple Correlation of I.Q. and GPA With Growth Score	z_r Equivalents for Multiple Correlation of I.Q. and GPA With Pretest and Posttest Scores: Pretest	Posttest
FAS I	.71	.60	.42	-.60	-.26	.33	.10
FAS II	.84	.81	.78	-.25	.59	.30	.65
TOUS I	.94	.74	.70	-.33	.33	.44	.74
TOUS II	.82	.86	.68	-.29	.48	.29	.71
TOUS III	.66	.70	.90	-.08	.45	.50	.69
Watson-Glaser	.57	.59	.58	-.58	-.38	.60	.35
Sci.&Cult.Ach.	.89	.86	.65	-.34	.18	.41	.48
SOV	.87	.88	.76	-.34	.71	.21	.24

opinion is justified further by the fact that the mean posttest score obtained by the class is eight raw score points below the highest possible score. The picture obtained thus far is complicated somewhat by the fact that the correlations between academic aptitude (intelligence and grade point average) and performance also suggest a greater treatment effect for the less able students. The considerable shift in correlation from pretest to posttest scores denotes the presence of a differential effect. The growth score correlation fixes the direction of that differential as being toward favoring the lower ability students. Thus it is possible that either or both academic aptitude and pretreatment response level determine the degree to which a student is affected by the course, if indeed the presence of nonadditivity has been established.

Not only is the initial indication of nonadditivity less definite in the case of the two Test On Understanding Science subscores and the "Science and Culture Achievement Test" score, but there is a lack of consistency with the additional evidence available. Although the correlations between pretest scores and growth scores are negative in each case, they are too low to provide confirmatory evidence. In fact, no real correlation exists between growth and pretest scores where understanding the process of science is concerned. In only one case, that of subscore II on the Test On Understanding Science, is evidence for the presence of a selective effect to be found in the correlation of academic aptitude with pretest and posttest scores. The shift in correlation involved here is rather distinct and is supported by a transformed correlation coefficient of .48 between academic aptitude and growth score. This, of course, indicates a greater treatment effect for the more able students. Thus there is no confounding involved and the evidence for a selective effect based on pretest score and that for one based on academic aptitude can be considered on their own merits. It is well within reason that the negative correlations between pretest and growth scores were produced by the ceiling effect. Sixty-five percent or more of the experimental group subjects came to within four points of the maximum possible score on both subtests of the Test On Understanding Science. Although this was not true for such a large percentage on the "Science and Culture Achievement Test", a large number of the test items were not valid by virtue of the fact that they dealt with material not treated in the course. These items are included in the test in anticipation of an expanded version of "Science and Culture" during 1967-68. The evidence indicating that higher ability students grew more in their understanding of science than their classmates cannot be discounted so easily. To the extent that trust can be placed in the correlations involved, this was the case to some degree at least.

A survey of Table 13 will reveal one additional hint of nonadditivity not suggested by the comparison of regression coefficients. The evidence referred to is the high negative correlation between pretest score and growth score for the Watson-Glaser Critical Thinking Appraisal. Additional evidence for the suggested nonadditivity can be found in the correlations of academic aptitude with pretest, posttest and growth scores. Once again, however, it is not possible to determine whether the apparent difference in treatment effect is determined by the student's critical thinking ability previous to the treatment, his academic aptitude or both. Of course, the possibility of a ceiling effect cannot be ignored. However, as was the case with the Facts About Science subscore, it is improbable that such an effect could completely account for the rather substantial correlational evidence.

On the whole, the evidence is not strong enough to support a claim that the "Science and Culture" course is generally selective in its effect. As would be expected, not all students profited uniformly. In isolated cases there is some indication of a systemmatic difference in response. But, while sufficient to bar the use of an analysis of covariance in those cases, it is not sufficient to suggest that the "Science and Culture" course is inappropriate for some particular segment of the population studied.

CHAPTER V

SUMMARY AND CONCLUSIONS

From the data which has been presented, it is evident that the group of twenty-one eleventh and twelfth grade students exposed to the "Science and Culture Treatment" showed substantial mean gains in score on a variety of evaluative instruments. The data also shows that a group of effectively matched subjects who were not exposed to the treatment failed to achieve such mean gains. The results of the main analysis have indicated that this observed difference in growth could have been obtained by chance in only one of one hundred replications of the experimental comparison. Thus, since logical and empirical evidence was presented to exclude other systemmatic differences as explanations, the greater growth exhibited by the experimental subjects can, with considerable confidence, be attributed to the effect of the treatment. Since neither of the comparison groups showed a significant growth in one of the dependent variables studied, the treatment must be considered ineffectual for that variable.

Translated into specific terms and placed in relationship to the seven hypotheses presented in the statement of the problem, these outcomes indicate that subjects who were students in the experimental course, "Science and Culture", evidenced a significantly greater increase in:

1. understanding of the scientific process, as measured by the Test On Understanding Science and the "Science Opinion Survey", than subjects who were not.
2. understanding of scientists as an occupational group, as measured by the Facts About Science Test and the Test On Understanding

Science, than subjects who were not.

3. understanding of science as an institution and its relationship to other institutions in our society, as measured by the Facts About Science Test and the Test On Understanding Science, than subjects who were not.

4. critical thinking ability, as measured by the Watson-Glaser Critical Thinking Appraisal, than subjects who were not.

5. the importance which they place upon theoretical values, as measured by the Study of Values, than subjects who were not.

6. understanding of the character of scientific and nonscientific segments within cultures and knowledge of the evidence for interaction between them, as measured by the "Iowa Science and Culture Achievement Test", than subjects who were not.

But they also indicate that subjects who were students in the experimental course evidenced no greater increase in substantive scientific knowledge, as measured by the Stanford Advanced Science Achievement Test, than subjects who were not.

The above positive results are corroborated at several points by evidence obtained from the "open-ended" techniques used in the study, the "Descriptive Words Test" and the "My View of Science" essays. In isolation such data would be of questionable significance since the symbols of learning cannot be equated to learning itself. However, when placed on the background provided by the criterion test data, it is indicative of more than the fact that students have learned to produce the accepted verbal symbols. There is good reason for contending that they understand the meaning of these symbols as they relate to science. The "My View of Science" essays are especially valuable in that they serve the additional function of reflecting student attitudes toward science. In a number of cases they provide evidence of a fact which was clear to the course instructors from the outset: some of the experimental subjects selected the

"Science and Culture" course because of a desire to fulfill a science requirement while avoiding a formal science course. This fact has obvious significance for interpreting the outcome of the study. It should be kept clearly in mind when the question of bias due to "self-selection" is considered. Moreover, an examination of the pretreatment-posttreatment essays will reveal that the treatment was able to overcome a considerable anti-science attitude in several cases. The fact that no significant difference in growth was observed on the Stanford Science Achievement Test is not surprising for two reasons. First, the "Science and Culture" course was not designed to increase students' knowledge of science facts. Secondly, the intrasessional history of control group subjects included a mean number of semesters of science courses which was almost four times that for experimental subjects.

In conclusion, the findings of this study rather clearly demonstrate that the experimental course, "Science and Culture", is effective in producing a broad increase in the scientific literacy of secondary school students such as those which comprised the experimental group. The substantial significance of the results obtained with an analysis employing nonparametric statistics justifies placing considerable confidence in this assertion. Since the need for fostering scientific literacy and attitudes is a matter of major educational concern, the potential value of such a course is evident. Although additional research is necessary before any conclusions can be drawn concerning the effectiveness of these course materials with other learners and other teachers, there seems to be reasonable cause for optimism.

APPENDICES

APPENDIX A

THE EXPERIMENTAL COURSE--DESCRIPTIVE MATERIALS

A Description of the Experimental Course

"Science and Culture" is a course which treats science as a continuing and cumulative human activity whose past history is essential to understanding its present complexion. It is concerned especially with matters of intellectual history and deals not with dry facts and chronology, but with pivotal moments wherein the kinds of scientific questions asked and the answers given produced changes in the mentality of Western man and hence represent essential chapters in the unfolding of modern science and the Western intellectual tradition. This design is dictated, not only by its promise as a means to a desired end, but also by virtue of the fact that it provides a more complete and accurate picture of science than could be achieved with more traditional approaches. Within this framework, the authors are committed to the following propositions: 1) science cannot be regarded as a thing apart from all other approaches to experience, 2) science is a vital part of culture and its content and course of development are powerfully influenced by the complexion of contemporary culture, 3) cultures are influenced and conditioned by the nature of scientific thought just as they are influenced by art and man's ideas about himself and the nature of life. Thus science and culture are viewed as inseparable; they are considered to overlap and influence each other to an extent which is startling to minds accustomed to the continued separation of science from what is fondly called the "humanities".

In its present form the course consists of three units. The first unit involves an orientation to the study of ideas, practice in the use of historical perspective, and an introduction to the science versus culture issue. In the second unit the focus is on the development of science during its interaction with philosophy, religion and politics in classical Greece. Particular emphasis is placed upon the nature of rationalism, the development of a faith in an order in Nature, and the character of "humanistic science". The third unit concentrates on the scientific "revolution" of the sixteenth and seventeenth centuries. This revolution is traced from its beginning in the fusion of practical and philosophical temperaments in the culture of the Middle Ages to its culmination with the Newtonian synthesis. Emphasis is placed upon the role of naturalistic art and technology in promoting "true science", the metaphysical presupposition of the first men of physics, and the role of the "new science" in creating the pattern of values in contemporary society.

The discussion method is used exclusively in teaching the course. A continual effort is made to bring students to develop their own ideas. The classroom atmosphere is structured so as to encourage independent thinking based upon all available evidence. Little or no emphasis is given to the mere giving of answers. Although the instructors present ideas, they do not dictate their blind acceptance.

Specimen of Materials From Unit III

The Scientific Revolution

Introduction

In this unit, as in the previous one, we will be concerned with discovering and studying the reciprocal relationships between science and other facets of culture within a period of the past. We will, as before, not be concerned with merely reciting a sequence of discoveries and ideas, but with viewing a process of intellectual change in which scientific problems are our central focus, but not our entire field of vision. We will be particularly concerned with those cases in which problems were not only solved, but forced a change in mentality in the process. Individuals and singular events will continue to be viewed as signposts along the road of mental evolution, important not in isolation, but as reflections of a way of thinking that is part of the track of intellectual development leading up to the present. At no time should it be let slip from our attention that it is only a stage in a long process that is under examination. As a fragment, this stage has its fullest possible meaning only when viewed in the context of as much of the entire process as we know. Hence, we must continually view this new material in the framework provided by our study of ancient science and culture.

In the brief consideration of the Alexandrian era which completed Unit II we saw evidence of the positive force produced by Aristotle's having returned sense experience to the place denied it by pure rationalists. During the several centuries immediately following Aristotle's death (322 B. C.), Alexandrian science continued to advance by virtue of maintaining a willingness to "interrogate nature". In its new environment it soon left behind many purely speculative and intuitive notions held by its Athenian ancestor. The atomic theory of Democritus, which had previously received almost no acceptance, began to receive attention. To a limited degree, ingenuity within the Museum even extended to designing a variety of improved instruments and pieces of apparatus. It must be pointed out, however, that the military engines, water pumps and similar devices which appeared during the 3rd century B. C., cannot be credited to the scientists of the Museum with certainty. They may well have had a more traditional origin as the handwork of craftsmen. In the case of Archimedes, however, the legends concerning his war machines are too numerous to be rejected as totally without foundation. It is not difficult to imagine how a well payed "professor" in the Museum might feel obliged to sacrifice "good taste" long enough to briefly apply his knowledge to practical problems. In Archimedes' case it is clear that practical applications were not undertaken out of desire. As Plutarch remarks, (Life of Marcellus, Chapter XVII) "He looked upon the work of an engineer and everything that ministers to the needs of life as ignoble and vulgar."

The 3rd century B. C. might well lay claim to being the golden age of science in ancient times. Certainly it is the period which contributed most to modern science. This period of glory was short lived, however. Although the decline of Alexandrian science was gradual, by 100 B. C. the creative outburst that had made the Alexandrians exceptions to the intellectual decay around them, had died out. The activity which remained consisted of merely pursuing in greater detail the already existing lines of thought. Certainly, potential for further progress at the time of decline was present. The essential beginning had been made and at least several of the seeds of modern science were evident. The importance of observation had been recognized and the pursuance of simple research had reached surprising efficiency. The capacity for organizing and systematizing information and expounding a subject from fundamentals to conclusions was established. The fund of information which had been wrestled from nature was sizable. Despite the fact that science stood on the threshold of the modern era, it seemed unable to cross it. Furthermore, this strange paralysis continued for approximately a millenium and a half. When modern science began in the sixteenth century it was not, in large, the result of a discovery of new facts, but the result of looking at the old facts in a new way. It is this story which will be the focus of Unit III. But before beginning that story two other subjects commend themselves to our attention. The first is an attempt to remove the paradox which we now see raised at the conclusion of Unit II. The second is an overview of the significant features of the period between the close of the Alexandrian age and the earliest stages of the "break through" which gave rise to modern science, the Scientific Revolution.

Decline of Alexandrian Science - Internal Causes

Although the failure of science to progress beyond the stage reached in the Alexandrian era was not due simply to its own inadequacy, this was part of the reason. Aristotle's reincorporation of sensory evidence into the scientist's proper field of attention was the huge asset with which the workers in the Museum began. However, during the period following Aristotle and his immediate successors, the use of systemmatic observation as a basis for science failed to mature beyond the place Aristotle gave it in the beginning. In fact, it actually underwent a relapse and the uneasy struggle between experimental methods and the use of reason alone began again. Gradually the phenomena of nature, rather than being the starting point in the search for knowledge, fell toward the position of being appealed to "after the fact", when they could serve to support already formed ideas. Furthermore, whether sensory evidence was being used as a starting point or in an attempt to prove a conclusion reached by reason alone, the patience to go beyond a cursory examination of phenomena was too seldom present. Only the astronomers provide any significant exception to this generalization. It seems that the excessive admiration for pure rationalism and the undisciplined eagerness for generality present in the old Greek tradition was still making itself felt. Even a man as well known for his devotion to experiment as Archimedes exhibited a serious weakness for emphasizing proof over discovery and logical deduction from "self-evident principles" over induction based on observation and experiment. The Greek mind seems to have harbored such a strong preference for abstraction and generality over facts, that its genius was not apt for laboring over individual cases in order to reach principles by inductive generalization. Even Aristotle himself made little systemmatic use of the fruits of his observational skill beyond developing schemes of classification. This stubborn urge within the Greek intellect was never really brought under rein and as a consequence the development of the kind of attention to fact essential to complete scientific mentality was never attained.

Beyond their failure to follow and improve upon the lead of Aristotle where it would have been wise to do so, the Alexandrians made the error of being content to accept his authority where they should not have - in the areas of physics and astronomy. As we have learned previously, these were the most intuitive and fanciful portions of the Aristotilian corpus. The biological works, wherein Aristotle excelled, were left largely ignored. Thus in physics and astronomy the Alexandrians interpreted whatever new information they gained with minds set in the Aristotilian mold. That is to say, they would consciously or unconsciously regard things in ways which made them fit with their Aristotilian prejudice. The resulting distortion not only helped to preserve Aristotle's faulty ideas, but by adding apparent supporting evidence, made them increasingly harder to excape. New knowledge, rather than serving as a stimulus to and a means for, achieving an improved theoretical structure, was serving to intrench the traditional one. The likelihood of progress under such circumstances is obviously slight. As we remarked earlier, when the Aristotilian views were finally overthrown (well over a millenium later) to make way for modern science, it was not a consequence of some sort of crucial new facts, but of the ability to look at old facts from a new point of view. Although it is a premature question, one cannot help but wonder why these later men were not afflicted with "paralysis" - were they so constituted psychologically that they could question authority more readily than the Alexandrians?

Decline of Alexandrian Science - External Causes

In addition to internal factors, there were several significant influences acting from outside Alexandrian science which help to explain its decline. One of the most powerful of these was a change in the position of science in society. The Museum represented the first attempt in human history to organize and subsidize science. Evidently the study of nature had come to be recognized, in some quarters at least, as an important ingredient in society. This fact remains the same regardless of whether it was the prestige value or the promise of practical returns from science which motivated the Ptolemies. Although the latter is less likely, each was probably a consideration to some extent. The important point is that science had risen from the position which led Socrates to reject it as mere folly. Most of us would concur with his judgement that little of value could come from much of the science of his time - science founded on imaginative speculation. Socrates would have leveled an equally pointed criticism against the science of the Museum. This criticism would have been aroused by the fact that the interest of the Alexandrians was in natural science alone and not in wisdom in general - philosophy in the old sense. Having seen to what extent the science of earlier Greek thinkers was adjusted to and so, distorted by, the religious, social, and political preconceptions which they held, we are hardly of a mind to share Socrates' position. The very nature of philosophy requires that it include the realm of values, while scientists deal with an objective world which they strive to keep free from values. This does not necessarily mean that science and philosophy cannot, or should not, exist together in a single mind in a mutually beneficial relationship. Ideally, this would be the case where both philosophers and scientists are concerned. However, it is essential that this blending does not result in destructive interference of one with the other. Such a relationship is difficult to maintain and the early Greek thinkers, in their role as pioneers, failed to even see the distinction between science and philosophy and so were unaware of the possibility of such "dangerous" interaction. Furthermore, they were, as we have said, not well armed with facts and therefore their science was prone to distortion from their philosophy. This being the case, it would seem better that they forgo the possible benefits derived from the combination than suffer what we have seen to be the possible detriment.

There is one serious omission in the above argument which makes the conclusion an error. We have treated science and philosophy as though they were entities apart from a particular society and culture. Lifted in this way from the context of their "life support system" and viewed in largely theoretical terms, it does indeed seem that science would have profited from being separated, as it was in Alexandria, from philosophy. However, this conclusion is premature until we have examined how this separation affected the position of science in the society of the time. The Greeks of the Alexandrian age, as well as their ancestors, expected men of learning to be interested in adding to man's entire fund of wisdom. Essentially, they saw only one subject of interest and therefore it would hardly be considered a virtue to be concerned only with the phenomena of nature at the expense of ignoring questions of morals, ethics, beauty, immortality, and other matters bearing on man, his purpose, and the significance of his doings. In fact, the latter topics were considered to be the most fundamental and hence of first importance. In cutting out a small section of reality and dealing with it as if it were independent of all other reality, the study of nature had developed a kind of specialization which was out of character with the mentality of the age, a mentality which found appeal in generality, unity, universality, etc. As a consequence, the dissatisfaction which Socrates would have felt with regard to science if he had been living, was felt by Alexandrian society. With this dissatisfaction came disinterest as well. As it retreated from philosophy, science drifted out of the main current of interest and came to rest in the "backwater". It became too scientific for the temper of the times. To make matters worse, even at this stage of development it had almost as little application to the problems of everyday life as it had in the past. Unlike the science of today, Greek science failed to reach a place where it could make a claim to support from society by virtue of direct or indirect benefits which it returned to it. Thus, in all ways, science had ceased to be, or had failed to become, a significant element in the culture of the time. This was a significant cause for its decline. It is the height of irony that a development which was among the most responsible for the movement ot science to the threshold of the modern era - the virtual attainment of objective empericism through removing man from the subject matter, was in no small part, responsible for its coming to rest before crossing.

Perhaps the observations which we have been attempting to convey will be more clear if we look at matters in quite a different way. No toiler can continue to exert the necessary effort to carry on his task without feeling that some <u>purpose</u> appropriate to his character is being served. It may be that he feels that not only he, but his fellow man or, if you prefer, his society, will profit from his efforts. Similarly, an area of <u>endeavor</u> must serve some real purpose - fulfill some function appropriate to culture of the time if it is to flourish for long in that culture. Lest this point be misunderstood, it should be realized that the function involved need not be a practical one to be appropriate. Certainly the arts can make little or no claim to practical value and yet they have survived when more practical activities have proven inviable. It must be kept in mind that the needs of a given culture and therefore the merits of a particular activity in that culture are difficult to judge from an "external viewpoint". The necessity of a historical point of view when judgements are being made in this realm is hence obvious. Furthermore, it is not necessary that the practitioners of a "craft" be explicitly aware of all the ends which their efforts serve for themselves or for their culture. In fact, the personal ends which are achieved for different individuals engaged in the same activity may differ greatly. The basic ends served for their culture, on the other hand, are common to all. Although the relationship between an individual's practice of his "craft" and his personal needs is an interesting subject, it is a concern of psychology and hence we shall leave it here. Our concern is with the relationship between the nature of a culture and the <u>common ends</u> which the activity of the workers in a "craft" serve for that culture, i.e., the relationship between the nature of a culture and the function which a "craft", as an activity, serves within that culture.

Now the ends which science can serve at almost any stage of its development are many. Accordingly, the ends which it does serve in a given culture depend upon the functions which the character of that culture make appropriate. The important point to add to this observation is that the nature of science is molded to a significant degree by the function which it is serving. Thus the nature of a culture acts, through the role which it grants to science, as a crucial factor in determining the nature of its methods and, to some extent, its contents. At this point it is essential to distinguish between the intensity of scientific activity and the meaning of that activity for the progress of science itself. The more closely science conforms to the function or role which the culture makes appropriate, the greater its level of activity will _tend_ to be. When science fails to conform to the "culture-fostered role", the probability is great that it will suffer a loss of "support" and hence undergo a decline in its level of activity. However, level of activity and progress are not synonymous. The fact that science is flourishing as a result of filling an "appropriate" role does not _necessarily_ mean that it is also developing in profitable ways, i.e., progressing in the sense of becoming "better science". The role which it is taking may be causing it to develop in ways which are detrimental to its progress toward better methods, more valid laws, and more powerful theories.

Although it will not be appropriate to pursue the point here, it must be added that we are talking about a "two-way street" -- as science plays its role, it tends to bring about changes in the culture and, therefore in what is appropriate to the culture. These changes may further reinforce the role and add to the force pressing science toward the particular character being fostered. On the other hand, they may alter the "role demand" of the culture and result in a new role and hence a change in the direction in which the character of science moves. Further, the disposition of "traffic" on this two-way street, the extent to which culture is influencing science, differs greatly over time. We are _studying_ a period wherein the formative influence of culture on science outweighs the reverse side of the relationship. We are _living_ in a time when the character of science is such that it holds the upper hand.

In essence, we have outlined a theory which focuses on particular _critical elements_ of science and culture and proposes a specific pathway between them along which the reciprocal influence we have been concerned with might act. Although we have found it possible to speculate intelligently about this relationship without such a theory, we may well be able to become more systematic in our inquiry and gain a greater understanding of our subject with its help. To review, the essential elements in this theory are: 1) cultures foster certain functions for the institutions within them, science included. These functions are determined by the ends which are appropriate to the particular culture--the"good" which represents the basic and pervasive concern of the culture and which is supposed to be attained by the way of life practiced in the society, 2) the level of activity in science is related to the extent to which it pursues the "culture fostered goals", 3) the character of the method and content of science is molded to a significant degree by the function it is serving, i.e., the goals it is being used to accomplish, 4) thus, when science responds to the concerns which are fundamental to the culture in which it is found and functions in sympathy with those concerning the nature of its methods and content (concepts, laws, and theories) come to reflect their character.

APPENDIX B

SPECIAL DIAGNOSTIC INSTRUMENTS

Selected Items from Study of Values

Part I

Directions: A number of controversial statements or questions with two alternatives are given below. Indicate your personal preferences by writing appropriate figures in the boxes corresponding to each question. Some of the alternatives may appear equally attractive or unattractive to you. Nevertheless, please attempt to choose the alternative that is relatively more acceptable to you. For each question you have three points that you may distribute in any of the following combinations.

1. If you agree with alternative (a) and disagree with (b), write 3 in the first box and 0 in the second box on your answer sheet.

2. If you agree with (b); disagree with (a), write 0 in the first box and 3 in the second box.

3. If you have a slight preference for (a) over (b), write 2 in the first box and 1 in the second box.

4. If you have a slight preference for (b) over (a), write 1 in the first box and 2 in the second box.

Do not write any combination of numbers except one of these four. There is no time limit, but do not linger over any one question or statement, and do not leave out any of the questions unless you find it really impossible to make a decision.

1. The main object of scientific research should be the discovery of truth rather than its practical applications. (a) Yes (b) No

2. Taking the Bible as a whole, one should regard it from the point of view of its beautiful mythology and literary style rather than as a spiritual revelation. (a) Yes (b) No

3. Which of the following men do you think should be judged as contributing more to the progress of mankind? (a) Aristotle (b) Lincoln

6. Which of the following branches of study do you expect ultimately will prove more important for mankind? (a) mathematics (b) theology

10. If you were a university professor and had the necessary ability, would you prefer to teach (a) poetry (b) chemistry and physics?

12. If you should see the following news items with headlines of equal size in your morning paper, which would you read more attentively? (a) SUPREME COURT RENDERS DECISION (b) NEW SCIENTIFIC THEORY ANNOUNCED

15. At an exposition, do you chiefly like to go to the buildings where you can see (a) new manufactured products (b) scientific apparatus?

18. If you had some time to spend in a waiting room and there were only two magazines to choose from, would you prefer (a) SCIENTIFIC AMERICAN (b) ARTS AND DECORATIONS?

21. Are you more interested in reading accounts of the lives and works of such men as (a) Alexander, Julius Caesar, and Charlemagne (b) Aristotle, Socrates, and Kant?

22. Are our modern industrial and scientific developments signs of a greater degree of civilization than those attained by any previous society, the Greeks, for example? (a) Yes (b) No

28. All the evidence that has been impartially accumulated goes to show that the universe has evolved to its present state in accordance with natural principles, so that there is no necessity to assume a first cause, cosmic purpose or God behind it. (a) I agree with this statement (b) I disagree

Part II

Directions: Each of the following situations or questions is followed by four possible attitudes or answers. Arrange these answers in the order of your personal preference by writing, in the appropriate set of boxes on the answer sheet, scores of 4, 3, 2, and 1. In each case the first box in the set corresponds to statement (a), the second box to statement (b), and so on. To the statement you prefer most give 4, to the statement that is second most attractive 3, and so on. Thus when you finish a given question you will have placed numbers in the first four boxes of the set corresponding to that question. You may think of answers which would be preferable from your point of view to any of those listed. It is necessary, however, that you make your selection from the alternatives presented, and arrange all four in order of their desirability, guessing when your preferences are not distinct. If you find it really impossible to state your preference, you may omit the question. Be sure not to assign more than one 4, one 3, etc., for each question.

33. If you could influence the educational policies of the public schools of some city, would you undertake --
    a. to promote the study and participation in music and fine arts
    b. to stimulate the study of social problems
    c. to provide additional laboratory facilities
    d. to increase the practical value of the courses

35. If you lived in a small town and had more than enough income for your needs, would you prefer to --
    a. apply it productively to assist commercial and industrial development
    b. help to advance the activities of local religious groups
    c. give it for the development of scientific research in your locality
    d. give it to the Family Welfare Society

36. When you go to the theater, do you, as a rule, enjoy most --
    a. plays that treat the lives of great men
    b. ballet or similar imaginative performances
    c. plays that have a theme of human suffering and love
    d. problem plays that argue consistently for some point of view

39. At an evening discussion with intimate friends of your own sex, are you more interested when you talk about --
    a. the meaning of life
    b. developments in science
    c. literature
    d. socialism and social amelioration

40. Which of the following would you prefer to do during part of your next summer vacation (if your ability and other conditions would permit) --

    a. write and publish an original biological essay or article

    b. stay in some secluded part of the country where you can appreciate fine scenery

    c. enter a local tennis or other athletic tournament

    d. get experience in some new line of business

41. Do great exploits and adventures of discovery such as Columbus's Magellan's, Byrd's and Amundsen's seem to you significant because --

    a. they represent conquests by man over the difficult forces of nature

    b. they add to our knowledge of geography, meteorology, oceanography, etc.

    c. they weld human interests and international feelings throughout the world

    d. they contribute each in a small way to an ultimate understanding of the universe

43. To what extent do the following famous persons interest you --

    a. Florence Nightingale

    b. Napoleon

    c. Henry Ford

    d. Galileo

45. Viewing Leonardo da Vinci's picture "The Last Supper", would you tend to think of it

    a. as expressing the highest spiritual aspirations and emotions

    b. as one of the most priceless and irreplaceable pictures ever painted

    c. in relation to Leonardo's versatility and its place in history

    d. the quintessence of harmony and design

"Science Opinion Survey"

Directions: Below are twenty-nine statements concerning the nature of science. Read each one carefully and decide whether you agree or disagree with what it says. When you have decided, place "A" in the blank to the right of the statement if you agree and "D" if you disagree. In some cases you may feel that you do not have a clear opinion but please make a response anyway by deciding which you tend more toward, agreement or disagreement.

1. The most important scientific ideas have been the result of a systematic process of logical thought. ____

2. Classification schemes are imposed upon nature by the scientists -- they are not inherent in the materials classified. ____

3. Thanks to the discovery of the scientific method, new discoveries in science have begun to come faster. ____

4. The primary objective of the working scientist is to improve human welfare. ____

5. While a scientific hypothesis may have to be altered on the basis of newly discovered data, a physical law is permanent. ____

6. The scientific investigation of human behavior is useless because it is subject to unconscious bias of the investigator. ____

7. Science is constantly working toward more detailed and complex knowledge. _____

8. A fundamental principle of science is that discoveries and research should have some practical applications. _____

9. While biologists use the deductive approach to a problem, physicists always work inductively. _____

10. The ultimate goal of all science is to reduce observations and phenomena to a collection of mathematical relationships. _____

11. The best definition of science would be "an organized body of knowledge." _____

12. Science tries mainly to develop new machines and processes for the betterment of mankind. _____

13. Any scientific research broader than a single specialty can only be carried out through the use of a team of researchers from various relevant fields. _____

14. Investigation of the possibilities of creating life in the laboratory is an invasion of science into areas where it does not belong. _____

15. Team research is more productive than individual research. _____

16. Many scientific models are manmade and do not pretend to represent reality. ____

17. Scientific investigations follow definite approved procedures. ____

18. Most scientists are reluctant to share their findings with foreigners, being mindful of the problem of national security. ____

19. The essential test of a scientific theory is its ability to correctly predict future events. ____

20. When a large number of observations have shown results consistent with a general rule, this generalization is considered to be a universal law of nature. ____

21. The scientific method follows the five regular steps of defining the problem, gathering data, forming a hypothesis, testing it, and drawing conclusions from it. ____

22. One of the distinguishing traits of science is that it recognizes its own limitations. ____

23. The steam engine was one of the earliest and most important developments of modern science. ____

24. Scientific research should be given credit for producing such things as modern refrigerators, television, and home air conditioning. _____

25. If at some future date it is found that electricity does not consist of electrons, today's practices in designing electrical apparatus will have to be discarded. _____

26. By application of the scientific method, step by step, man can solve almost any problem or answer almost any question in the realm of nature. _____

27. Scientific method is a myth which is usually read into the story after it has been completed. _____

28. Scientific work requires a dedication that excludes many aspects of the lives of people in other fields of work. _____

29. An important characteristic of the scientific enterprise is its emphasis on the practical. _____

"Iowa Science and Culture Study Achievement Test"

Directions: Mark the appropriate blank on the answer sheet to indicate the best answer for each question.

1. The creative periods in science and technology in history tend to be associated with periods of
   a. economic and political stability
   b. economic stability and political change
   c. political stability and economic change
   d. economic and political change

2. The existence of large scale slavery during the time that Greek culture reached its peak in Periclean Athens was generally
   a. an asset to the growth of science because it provided men with time for study and thought
   b. an asset to the growth of science because it tended to further distinguish scholars from technicians
   c. detrimental to the growth of science because it made the implements and processes for the control of nature the province of slaves
   d. detrimental to the growth of science because it fostered a shift in focus from theoretical to practical knowledge

3. Why is science described as a cyclic activity:
   a. because scientific achievements appear periodically in history
   b. because all facets of a fundamental problem must be investigated before significant results can be expected
   c. because one finding usually leads directly to another
   d. because the fundamental ideas are often completely altered in time

4. Why is science described as a human enterprise?

   a. because science is often thought to be an impersonal study of text book materials

   b. because science is whatever scientists do to discover more about the universe in which we live

   c. because people involved with science are extremely dedicated

   d. because the activities with which most people are involved illustrate little of humanistic qualities

5. A reflection of the degree to which ancient Greek science was related to human life and institutions is that

   a. Pythagoras had a semi-mystical theory of a universe somehow made entirely out of numbers

   b. inscribed above the entrance to Plato's Academy were the words "Let no one enter here who would be both politican and scientist."

   c. many Greeks thought of the universe as analogous to a city - state

   d. men of science are treated with almost religious reverence in Homeric mythology

6. How does the fact that science is self-correcting separate it from the field of the humanities?
   a. in the humanities the result of human effort stands as it was created while in science the products of man's mind are constantly changing
   b. in science most people do not appreciate what others are doing except those in very closely related fields
   c. no one in the humanities ever creates the same product whereas this is common in the sciences
   d. whenever the product of efforts in the humanities can be found, it is known as a single expression of a man unrelated to anything else in the field

7. Why are scientists so interested in developing theories?
   a. the development of theories enable scientists to view at one time a total group of observations
   b. this provides an expression of all that is known by the given scientist
   c. theories enable explanations which in turn suggest possible experimentation
   d. the observations are only valuable when they are applied to an idea or when they are fully classified

8. The nature of Shakespeare's achievement in writing "Othello" is comparable to
   a. Bell's invention of the telephone
   b. Galileo's discovery of the law of inertia
   c. Copernicus' conception of a helio-centric solar system (planets revolving around the sun)
   d. none of the above, since Shakespeare's work involved more than the rational intellect

9. Which of the following best describes the fundamental objective of a scientist?
   a. learning more of the facts of nature so that they can be passed on to future generations
   b. searching for answers to questions about the universe in which he lives
   c. trying to record all the information about mankind and his understandings so that it can be studied as an area of knowledge
   d. attempting to produce more and better products in order that we can have a better life

10. The contrast between classical and modern art is similar to that between classical and modern science because in both cases
   a. the modern discipline represents an attempt to improve upon nature while the classical discipline sought only to represent nature
   b. the same concepts are involved but the modern discipline involves seeing them in new kinds of relationships
   c. the modern discipline involves a greater separation from everyday reality (the sensible world?)
   d. none of the above

11. What does a scientist do when the results of a given set of experiments cannot be explained by prevalent theories of the time?
    a. calls in a team of experts active in the field to arrive at another theory
    b. ignores the experiments since they are not in keeping with prevalent concepts
    c. discards the theory since it is obviously wrong
    d. expands the theory to include the new facts

12. Which of the following would characterize the fundamental goal of a scientist?
    a. identify problems which need answers
    b. record observations about the universe
    c. use data collected from experiments in support of theories
    d. find a cure for a disease like cancer

13. Which of the following is the most common driving force of a scientist?
    a. gaining recognition among his associates
    b. securing greater financial rewards
    c. finding ways of improving the health and happiness of others
    d. adding to the general understanding of the world in which we live

14. Science, as an intellectual activity, can be considered an art in the sense that

a. some scientists paint or indulge in other "artistic activities" in their spare time

b. scientists are required to express themselves in the literary form when they prepare articles for publication

c. when a scientist designs and builds special equipment he must be "creative"

d. scientists who formulate theories employ their imagination in a creative way

15. In the course of history the sciences and the arts have tended to

a. flourish at different times and in different places

b. flourish at same time but in different places

c. flourish at different times but in the same places

d. flourish at the same times and in the same places

16. If you were to site evidence to support the notion that cultures are influenced by man's opinions of what he is and what life is, you would be least likely to select the culture first exposed to

a. Darwin's theory of evolution

b. Copernicus's model of the solar system

c. Freud's theory of personality

d. Mendel's theory of heredity

17. Today the largest share of the financial support given to scientific research comes from

    a. universities

    b. industrial corporations

    c. government agencies

    d. wealthy individuals

18. Many people condemned the scientists who developed the atomic bomb alleging they created a monster which may end civilization as we know it. Which of the following is the best defense for these scientists?

    a. scientists are amoral, they cannot be responsible for the uses to which their discoveries are put by other men

    b. Germany was also seeking to develop atomic weapons which they would have used to enclose the world

    c. the material benefits possible from atomic energy is greater than its potential for destruction

    d. the expansion of human knowledge about man and his universe must always go forward

19. Which of the following is an example of Post hoc ergo propter hoc thinking?

    a. "The weather has gotten worse since the atomic bomb tests."

    b. "I know it because I saw it with my own eyes."

    c. "Man is mortal. Socrates is a man. Socrates is mortal."

    d. "He can't see the forest for the trees."

20. What is meant by "cultural lag"?
    a. scientific developments cannot keep up with man's needs
    b. several decades or more often pass before achievements come into wide-spread use by man
    c. man's culture often causes him to be neglectful of looking to future needs
    d. several generations are required for man to adapt to change

21. Why are science and technology often confused?
    a. scientific knowledge is used for technological advances
    b. many of the same people are involved in both kinds of endeavors
    c. technology is the immediate forerunner of science activities
    d. they are synonymous terms and hence should be confused

22. If you wanted a better transistor radio, who would be the most likely person to contact? Why?
    a. an engineer since he would have conducted the pure research necessary to do the job
    b. an engineer since he would be involved in applying basic principles to make such advances
    c. a scientist since he knows more about basic electronics than anyone else
    d. a scientist since he must be the person to make basic discoveries necessary for changes in design

23. Why are observations alone of little value in science?

    a. they do not explain anything of the universe by themselves

    b. they illustrate only the accumulation of basic understanding

    c. they can not be used to support or to correct a fundamental idea

    d. they are useful only when recorded logically and reported to other scientists

24. Which of the following is a fundamental characteristic of a good hypothesis?

    a. it explains some of the observations

    b. it can be eventually proven as a law

    c. it is testable

    d. it will lead to specific conclusions

25. Which of the following activities of a scientist illustrates the creative aspects of science?

    a. defining the problems to be approached

    b. formulating theories to serve as explanations

    c. designing specific experimentation

    d. identifying justifiable conclusions

26. In which of the following ways are scientists and historians most similar?

    a. in their reverence for past achievements and developments in their respective fields

    b. in their rigidly defined rules for the treatments of evidence and the conduct of research

    c. in their desire to understand their present culture

    d. in their ability to remain free of the culture in which they operate

27. In which of the following is the path of science least likely to be affected by the culture of the times

a. selection of problems for study

b. discovery of laws

c. creation of theories

d. formulation of concepts

28. Scientists are most apt to be able to work in the area of pure rather than applied science when their "patron" is

a. a wealthy individual

b. a university

c. an industrial corporation

d. a government agency

29. Why is the potential usefulness of a given piece of research of little concern to a scientist?

a. he thinks of only one problem at a time

b. he can not consider use until the answer to the first problem is known

c. he will explain his observations, develop a model, and draw conclusions and not spend more time on the same problem

d. he is primarily concerned with an explanation of natural phenomena and not how the information can be used "practically."

30. How does a trial and error investigation differ from a more desirable type usually employed by scientists?

    a. in trial and error there is no model enabling some prediction

    b. few discoveries are made utilizing a trial and error procedure

    c. a trial and error experiment is by nature less complicated and involved

    d. no one starts out with a trial and error approach until all other possible procedures have been investigated

31. Freud's psycho-analytic theory received strong opposition when it first appeared. After World War I, however, it rose to a place of considerable influence. This can be at least partially understood in the light of

    a. the general scientific respectibility granted to psychology after its successful application in warfare

    b. the relaxed attitude of the scientific community after its release from war time pressures

    c. the fact that Freudean psychology provided an explanation for man's inability to solve social problems by corporate action

    d. the discovery of physiological correlates for Freud's ego, id, and super ego

32. The Greeks felt that science was good because

    a. it was intellectually appropriate

    b. it promised advantages to mankind

    c. it was a form of worship of God's majesty

    d. all of the above

33. The basic changes which took place in the orientation of thought during the period of the Renaissance and Reformation included all but which of the following

    a. religious to secular

    b. qualitative to quantitative

    c. continuity to atomicity

    d. limits to infinite extention

34. Which of the following fails to describe a feature of the process of science?

    a. an inquiry approach to problems

    b. an activity of the minds of man

    c. testing theories proposed to explain natural phenomena

    d. collection of information concerning the universe to be learned as a body of knowledge

35. Which of the following best characterizes science as an activity?

    a. observations and descriptions lead to experimentation and temporary explanations of natural phenomena

    b. it is a human discipline that is unchanging although new information is added to the old

    c. science is concerned with trying to find a means of making a better life for people

    d. collections of observations of the universe results in adding to our books of scientific knowledge

36. How has science developed over the years?

a. unifying concepts and theories have developed from seemingly unrelated facts and ideas

b. each generation has produced its new science which is often unrelated to the science of the past

c. it has been centered around relatively few great minds which have dominated certain periods

d. new ideas are produced each day which replace old ideas

37. In the ideological realm, Darwin's theory of evolution made a damaging break in

a. Aristotle's idea of final causes

b. Plato's doctrine of ideal forms

c. the theological idea of divine creation

d. all of the above

38. The split between the sciences and the humanities, which is so much a feature of our times, first became obvious

a. during the Scientific Revolution of the 16th and 17th centuries

b. during the Industrial Revolution of the 19th century

c. during Rome's domination of the Greek speaking world

d. during the Middle Ages

39. Basic to the development of the split referred to in item 38 above was

    a. the overthrow of the intellectual assumptions inherited from the Greeks

    b. the identification of science with industrialism

    c. the separation of science and philosophy

    d. the movement of scientific and artistic productivity to separate geographical centers

40. In the economic realm, Darwin's theory of evolution seemed to give scientific blessing to

    a. government control over industry

    b. unfettered competition

    c. socialism

    d. machine production

41. At the close of the nineteenth century

    a. the major problems of both science and society seemed to have been finally solved

    b. while the general framework of scientific theory seemed secure, the future of society was clouded

    c. while the present state and immediate prospects of society were secure, science faced an overwhelming number of basic dilemmas

    d. both science and society were in a state of turmoil

42. During the Renaissance the visual arts

    a. concentrated on ideal forms and hence contributed to a climate which favored the reform of the science of the Middle Ages

    b. concentrated on ideal forms and hence contributed to a climate which retarded the reform of the science of the Middle Ages

    c. sought to achieve realism and hence contributed to climate which favored the reform of the science of the Middle Ages

    d. sought to achieve realism and hence contributed to a climate which retarded the reform of the science of the Middle Ages

43. Francis Bacon, a philosopher of the early Renaissance, set out to preach the doctrine that "The true and lawful end of science is that human life be enriched by new discoveries and powers." His position is a reflection of the fact that

    a. throughout the Middle Ages the natural philosophers were more remote from the craftsmen than at any other time

    b. he was aware of the justification which the then emerging society would demand for its support of science

    c. natural philosophy had been relatively sterile when compared with the inventiveness of the crafts

    d. all of the above

44. In what sense can science be called a modern superstition?

    a. it contains a number of almost mythical ideas

    b. it includes such things as astronomy and astrology

    c. some scientists hold deep religious convictions

    d. none of the above

45. How does Aristotelian physics give evidence of being affected by Greek cultural thought?

    a. it makes use of Homeric mythology

    b. it is strongly colored by ideas of harmony, symmetry, purpose, etc.

    c. it makes a clear distinction between technology, science, and arts

    d. it is based on ideas derived from the crafts of metallurgy, music, and ceramics

46. Which of the following best describes how science and culture interact?

    a. science affects the culture of man while man's culture affects his science

    b. science affects man's culture and has provided the framework for its advance

    c. science has provided the necessary knowledge needed for cultural advances

    d. science and culture act independently of each other; however, the results of both are inseparable in the daily lives of people

47. The conflicts between science and religion historically have generally resulted in which of the following?

a. the triumph of science by widespread acceptance of the irrefutable evidence supporting scientific claims

b. compromises by religion either through capitulation or through resolving scientific claims with Biblical and religious interpretation

c. the continued existence side by side of religious and scientific opinions

d. the suppression of the points of conflict by scientific accomodation of religious views

48. Which of the following best illustrates how the Greeks used rational inquiry to decide significant moral and political questions?

a. the Aristotelian idea of the nature of the make-up of the universe

b. the Socratic-Platonic concept that knowledge is virtue

c. the Galenic explanation of how the human body works

d. the Aristotelian idea of the make-up of social structure

49. How does the concept of science exhibited during the sixteenth and seventeenth centuries depart from Greek science?

   a. Greek science is based strictly upon observed phenomena while conceptual schemes which are not readily observable are the base for later science

   b. Greek science is illustrated by rational inquiry while reductionism characterizes the science of the sixteenth and seventeenth centuries

   c. the science which follows Greek times began anew in terms of departure and outlook

   d. although the nature of science did not change, the use of it did

50. Why do most scientists attempt to explain their results instead of simply reporting them in the literature?

   a. this prevents misinterpretation on the part of the others who might read the report

   b. this allows the investigator to relate the results to past efforts and point to some future experiments

   c. this insures that they vocalize and communicate directly with other scientists

   d. this provides a means of directing the minds of people in the matters of trying the same experiments in other research laboratories

APPENDIX C

COMPARATIVE STATISTICS FOR SAMPLE POPULATIONS

Experimental Group

Sex and Grade:

	11	12	Totals
M	5	2	7
F	8	6	14
	13	8	21

Cumulative Grade Point Average:

Mean = 3.03
Standard Deviation = 0.55
Range = 1.96 - 4.00

Science Grade Point Average:

Mean = 2.66
Standard Deviation = 0.72
Range = 1.36 - 4.00

Intelligence Quotient:

Mean = 121.0
Standard Deviation = 12.06
Range = 99 - 146

Control Group

Sex and Grade:

	11	12	Totals
M	5	2	7
F	3	11	14
	8	13	21

Cumulative Grade Point Average:

Mean = 3.01
Standard Deviation = 0.60
Range = 1.97 - 4.00

Science Grade Point Average:

Mean = 2.83
Standard Deviation = 0.70
Range = 2.00 - 4.00

Intelligence Quotient:

Mean = 116.7
Standard Deviation = 13.10
Range = 95 - 146

APPENDIX D

ESSAYS: "MY VIEW OF SCIENCE"

Subject 1

Pretreatment

Science has always been of interest to me, but I can not always understand it because of technical properties which I can not grasp completely. When I compare Science to Social Studies or English, I usually put Science in the middle. Whenever I hear the word science, my first image is a man in a white lab coat with a dissected frog beside him. Upon deeper inspection, I realize that Science is the base of "learning life" and why, how and when. I always feel that some accepted institutions of science are rather futile, however. An instance might be the approximate age of the earth. What good is it to know? Does it help me to know something else?

Although science interests me (especially actual nature), I can not see why a person can be so involved with matters which are senseless themselves, but are of vital importance when incorporated with something else. I don't know how they do one without the other.

Posttreatment

My personal view of Science is rather confused at this time. I have seen thru this course that Science and technology are separate realms, but I cannot view one without the other. Wherever I see Science, I think of technology and vice versa.

Generally, science is the collection of all facts about our universe and us, that deals directly with our lives. Not like history, etc., it can be applied anywhere and at any time. (eg. In 1 A. D. the grass grew because it got chlorophyll manufactured from sunlight, just as it does today.)

Seeing it in the way I wish, science deals not only with facts, but with objects obtained by facts (technology).

Subject 2

Pretreatment

The pretreatment view of subject 2 was not obtained.

Posttreatment

Science being an inquirey into nature, its ultimate that of understanding nature and the universe. Its means? a) question formulation b) observing -- collecting data d) formulation of theories e) looking for more problems --

Concerned more with the essense of things than how they can be used to help man --

Interested in understanding -- not changing --

Subject 3

Pretreatment

The scientist is a man who works theorising and/or gathering data. He hopes someday he will make a discovery to better the world in which he lives. He is very friendly with his scientist friends. Sometimes they talk about a new physics theory, an advancement in biology or sometimes they even talk about their families. Scientists (especially those at Cape Kennedy) have difficulty communicating with their colleagues so they make up a whole new language full of terms like "A-OK" and other humerous names for the tools of their trade. Scientists are allright as long as they stay in their laboratories and get us to the moon first. I should add that scientists have an odd sense of humour and little appreciation of the "finer" things in life.

Posttreatment

Science is a search for knowledge about nature, or search for the truth. But this search is somewhat meaningless to anyone who is not a scientist himself. Yet the laws of science, the methods and rules of science have meaning for me as they can be applied in everyday life. And of course science is important to everyone because of its great influence on technology which effects everyone.

Yet I feel it is important for this uncovering of knowledge to continue. It is important for man to understand his environment, so he may live better. Whether this is important or not to the scientist, it is important to me.

Subject 4

Pretreatment

The field of science is a very large one. Science does not start and end in the labratory of a scientist who is only interested in finding answers, but extends to the factories, industries, and even to the home. There are many branches of science going from technology to pure science. Science is a specialized field, and I feel as the ladder progresses from technology to pure science the degree of specialization increases. Also, just as a ladder extends and the top seems to get farther and farther out of reach, science as it gets closer and closer to pure science seems to be beyond the relm of general understanding by other people and thus it tends to form it's own world, and the public begin to loose interest and not care about science.

Posttreatment

Science is an ever changing thing -- new theories, facts, and discoveries are made all the time. However, in its state of constant change the methods, etc., of finding answers and discoveries are relatively constant. To me, science has largely been supported on the creativity and discoveries of a few "talented" men. Science is supposed to answer questions of the universe, etc., but in my mind science is not on my list as one of the most important necessities to live or survive -- just as religion. I admit -- technology does make life "easier" -- and cures for diseases do make the chances of living better (survival of the fittest) -- and that science is responsible -- indirectly for these things. But, I still do not believe in the "religion of science".

Subject 5

Pretreatment

To me science could just as well not exist. I do not understand it. In fact I hardly believe in it, because no matter how hard I may study or try to interpret it, it still remains non-sensical to me. I can observe its practical purposes true, however beyond that it holds no magnetism or desire for me. I exclude it from my life as much as possible because it confuses me and I resent it because I can not understand it. I much prefer the humanities which interest me very much and which I enjoy pursuing. Science to me is nothingness. I don't even <u>want</u> to understand it. It holds nothing for me, no appeal at all. I am probably distrustful of it too, for we are often mistrustful of things which we are unable to comprehend. The very fact that I don't care about it alienates me from the sciences yet I feel in close companionship with my fellow humanists. I am content to let science go its own way, follow its own callings, with no help or communication from me. In short, I lead a happy life without understanding or communicating with the sciences.

Posttreatment

My opinion of science has changed quite radically in the past year. Although previously I was one of those who considered science something of a "modern superstition" I now feel that I understand it much better.

Science to me is now not so far-off or incomprehensible as before. I can understand how science can so greatly effect our culture and our society. Likewise I can understand how our society and its values can influence our science.

On the other hand, in one way my view of science has not changed.

I still believe that there is a certain aptitude for science or for scientific thinking. I have resigned myself to the fact that there are many scientific principles and ideas which I will never be able to understand.

Nevertheless I can see that the great degree to which science and culture interact, influence and affect one another, neccesitates a greater understanding of the values, ideas, and attitudes within each.

Subject 6

Pretreatment

Def. I

Science involves experimentation and analyses, the results of which are carefully recorded. Before this is begun however a specific problem must be kept in mind, so that the experimentation will have direction and purpose. Any activity engaged in by man which is gone about using this "scientific method" becomes a science.

Def. II

Science is a class in school that I am either (a) good at or (b) really hurtin' at depending upon whether it is (a) biological science or (b) physics. I suspect my lack of ability where the latter is concerned is due to the fact that I have never been in a good physics class (with a good teacher). Since a good teacher is necessary to stimulate enthusiasm. and enthusiasm is needed to learn a subject well, I attribute my lack of knowledge regarding physics to the absence of a good teacher in my background.

Posttreatment

Although this course seems to have been centered around the value of science for knowledge and not for practical application, I, not being a scientist, view science from a sociological-technological point of view. In other words, I tend to think of science in terms of how its practical applications will affect me and my culture.

Further, I am somewhat wary of scientists' smug amorality because of the dangerous potential of their discoveries -- especially atomic energy -- However, I realize that for science to advance, (assuming that this advance is a good thing) scientists must not be hampered by humanistic concerns.

And it is for this reason, the amorality and unconcern scientists have for the dire affects of their work on society that I argue that science (no matter how creative) can never join the other humanties. As long as it wishes to remain separate, it will be separte.

Subject 7

Pretreatment

Science is a means of exploration for gaining knowledge of the world we live in. It is a process, whereby you use experimentation and thought to make new discoveries which may later bennefit or hinder man. Science gives you laws and theories to use in your experimentation, to help you gain knowledge of what occurs in nature. Science seems worthwhile only if you are curious about why things happen, or want to go on in science and become famous by introducing to man some new discovery (such as a new drug for medical purposes). In all science is a personal experience which inables you yourself to gain knowledge -- or to have the science teacher tell you the knowledge other scientists have gained through their experimentation.

Posttreatment

Science is sort of an art, of creating and experimenting to learn about nature. It creates views to explain our present knowledge and to predict what will happen.

A scientist must be objective in what he does, but he may be somewhat subjective in his theories.

Subject 8

Pretreatment

I feel that science is not as creative as, for example, art or music. It is limited mainly to the scientific method and any departure from the same would be heresy. The scientific method is held divine.

The only real purpose of science is experimentation. Even this is frequently limited to "recipe" repetition of experiments. The only time I have really enjoyed myself in science is when I did something from beginning to end, from conception to completion. This occured, I believe, twice or perhaps thrice in my entire high school science career.

The only reason I have really prefered (or still do) English and Social Studies to Science is that I have recieved more of a chance to express my originality, to be creative, in these courses than in the sciences.

It is unfortunate that I was not given more opportunity to be creative in the sciences. Science has a definite edge over Social Studies and English in this respect. One can't take a person, starve him, and see how he will respond to social laws. Likewise, one can't "test" a story on the public. Both involve the human element, the uncertain, variable element. Science does not. On can experiment in science. Perhaps this has been science's nemesis and also its gift, because experimentation, while encouraging creativity, also emphasizes methodology, and with this comes more "recipe" repetition.

Posttreatment

This course has changed my attitude towards science. Admittedly, at the beginning of the year I did not distinguish between science and technology. I do now.

I also now identify several important attitudes and outlooks. For example, I realize now that scientists' goal may be to answer questions about the universe. This includes all the methodology of science. The identification of the problem, the searching for facts, the theory and its proof, prediction -- all these fall under this general goal. However, several paths of this methodology may be changed or omitted. The Greeks, for example, did search for facts, but through intellectual probing, not experimentation. Alexandrian science, as another example, does not place emphasis on prediction.

Subject 9

Pretreatment

Science, to me, is a field of many topics, ideas and concepts. It is not just one, main, definite idea and field, but a mixture of many concepts which all combine thus forming the broad and mysterious field of Science.

Although there are many branches of Science -- such as Physics, Chemistry, Biology, Geology and so on -- there are indeed present in the Sciences some common factors among them which unites them into one broad topic: science. Such common uniting factors among the sciences which I feel to be quite obvious to all -- are as listed in the following sentences: 1) The curiosity in which all scientists display in performing in their field of science. 2) The experimentation done by all scientists. Out to find the real truth in nature. 3) The hypothesis, questions theories which are made to help seek this real truth.

Posttreatment

My view of science is much clearer than it was at the beginning of this course. Science, to me, is the study of nature -- collecting observations, data, experimenting, making hypothesis's, laws, theories, facts, and being objective. Before this course I often confused science with technology, but now am quite aware of the differences between the two. Science is not studied for it's use to society or it's practicality. It is studied just to find more about the world in which we live.

Science is studied in an objective manner -- value-free, no association with personal opinions or desires. It is merely collecting facts, observing, experimenting, (etc). It is aesthetic to the scientist actually engaged in his work -- but non-aesthetic to the rest of society. The rest of society does not understand science -- they get no meaning from it. Science is not for the purpose of bettering society -- that is just an indirect result. Technology deals with the practical application of science. Technology is the branch of science with the purpose in mind of bettering society.

Subject 10

Pretreatment

When someone first mentions the word science I think Albert Einstien in a typical laboratory filled with testtubes, chemicals, white coats and the like. On the blackboard is the formula $E = m C^2$. However should I think of science as more than just a typical everyday thing then there seems to be some meaningfulness in experiments and observations. I think of the atom bomb and hula hoop etc. The difference between science and technology lies in that of mechanical ability and motives. The scientist has created an idea which he believes will benefit himself and society. He experiments and builds upon the idea and constructs the finished product. The technologist takes over from there by actually making the product fit for an assembly line but does not put the origanal creation together. Everything in science is rational, logical and systematic and anything that is not "right" is not rational.

History is the link between science and the humanities as all fields have a beginning, a progression and eventually some type of end.

Posttreatment

Science, although boring for me, serves the useful purpose of seeking knowledge for knowledge's sake, which can later be practically utilized by technologists for the betterment of today's societies.

I might add that scientific knowledge which cannot be useful eventually, seems ridiculous even though I understand the purpose of pure science is to discover and learn, disregarding any practical value.

Subject 11

Pretreatment

The way I view science is, as a determining facter of the world. Science does not really affect me at all. I think it as an occupation, but wrong again. I said, well the scientists are doing a good job, I'll just sit back and watch things fly by. So there is an average american response, much different then what you'll find in the other essays.

Posttreatment

Science is a fastly progressing unit, without it the world, would be clumsy and we would all stumbal about. My view of science is that we can not do without it, and I think the more mass support there is of science, the better off the world will be. Science like Religion, both are competaters in very different fields, but both try to explain points about the universe. My view of science, as a very humanistic field, and it is only when techonoly enters it to it can it be destructive.

Subject 12

Pretreatment

Science is an essential part of our or any society. It is by the use of science that many of our most important advancements for the betterment of the world are made. Science offers challenge of the unknown, and the knowledge of nature.

A scientist must be a person who has the ability to search for facts, and interpret them correctly. He must be a person who is capable to use creative ability to find problems and then to find ways in which to solve them.

Science today seems to be in a much larger scale than ever before. More attention is paid to it and it seems important to our country to be advanced over other countries in science. Science seems almost commercialized. For at present it takes expensive equipment and an alert and intelligent scientist to light upon a discovery.

Posttreatment

Science to me is the pursuit of knowledge and understanding of nature and the universe. Science plays an important, rather essential part in the culture of today. Science's influence is seen in almost every aspect of life. I do believe that science goes along with technology in the sense of which I am speaking of it. With the application of the scientific discoveries by technology I believe the influence is greatest. I believe that a scientist could follow his own patterns of research and completely isolate himself and the meaning of his science would be nothing less.

Science has an influence on the religion and beliefs of a society.

Subject 13

Pretreatment

Science is the study of nature to find new ideas. These ideas are used to better the life of man and make it easier for him. His motivation is to find something new and useful such as a cure for cancer.

Science is extremely hard to follow because it is so technical and mathmatical. The ordinary person often has no interest in what science is achieving because he can't understand why it works. Science is hard to study but if you finally pass that stage it is very interesting and thought provoking.

Science classes are rarily boring but it isn't as much fun to be given the answer or idea. In the humanities you have to figure out things as they apply to you. Science can't really be applied to a person.

Posttreatment

Science is an attempt to gain knowledge of the universe. It is concerned only with discovering the natural laws behind phenomena and broadening these laws to generalizations and theories. Science is not concerned with the application of its knowledge. One important aspect of science is its ability to predict results of the future.

Although science is not particularly interesting to the common man it does have aesthetic values for the scientist. Science does not worry about this because its goal is to gain knowledge not to ease life for the common man.

Subject 14

Pretreatment

I am definitely a "humanist", in that I believe that people and society are more important than facts and hypothesis. I want nothing to do with what I consider the narrowness of science. Then, before I took this course I read all the works of a French aviator, philosopher, writer, mathematician, and science: Antoine de Saint-Exupery. You have probably never heard of him but to me he was the greatest humanist who ever lived. This brought me to think: How can he be devoted to humanity and to science at the same time? I thought one had to be all one thing or the other. This disturbed me greatly and lead me to think deeply about the subject. Through this thinking and through the writings of people such as Paul Tillich I have concluded that there is no friction between science and religion or humanity. They live in different realms in one's mind. Science deals with reality, religion deals with eternity. This may be oversimplified; you may not agree -- but I see no fractured culture.

Posttreatment

It disturbs me that over and over again in class we had expounded to us the theory that scientists are amoral, and therefore cannot be held responsible for the uses to which their discoveries are put by other men. Indeed, it is none of their business or their concern.

First, this ivory tower let-the-world-hang-itself is the worst possible attitude in which to approach understanding between sciences and humanities. It simply perpetuates and widens any existing rift. Why then are we led to believe this about the scientific profession?

This leads to my last and most important point, simply, this attitude is not the _only_ nor the _prevelant_ one among scientists, at least as far as I can tell. A famous example: Openheimer stated after the completion of the bomb that all physicists now knew hell, and now bore the full stamp of the human condition. Scientists have never been able to escape their own humanity, and today, now more than ever, cannot afford to escape the humanity of the world.

If science stands concerned, if it wishes to bridge the gap it now professes to see, it _must_ come at least half-way in the effort. If it remains in its ivory tower -- real or imagined -- it will become socially irrelevant.

N. B. Thank-you, both Mr. Cossman and Dr. Fitch, for challenging me to think, and allowing me to think even when my results weren't always what you would have wished. We were always at least honest. You've got something important -- keep it going.

Subject 15

Pretreatment

Since my father is in the field of science education, my views are perhaps a little more favorable towards science. To me science is a tremendously large field of curiosities. There are so many things in science that a person could get interested in and become really involved in. I'm constantly exposed to ideas of science, especially in the earth science field, and it's obvious how easy it is to take the step into the sea of science. It seems like that's the only thing my father ever thinks about.

I view the scientist as a friendly guy from the same sort of environment that I am from ("just plain folks"). He isn't exceptionally brilliant (there are some exceptions, but I usually think of a scientist as a science student), but he is enthusiastic about his occupation and works hard at it. Not any harder than a person in any other field, but probably puts in more hours than most other occupations.

Science is an honorable profession in which great heights can be reached with hard work and dedication.

Posttreatment

Science is an important part of society. It is a means of striving for answers to why things are the way they are and what man's relationship is to the overall picture of the universe. It also helps to solve problems of unemployment.

Subject 16

Pretreatment

I think of science as a bunch of absent-minded proffessors in white lab coats making-up experiments in a place full of complicated apparatus and guinea pigs. Science, to me, is learning about how and why things happen the way they do, and doing alot of fairly unrelated experiments by following the directions out of a textbook. I don't look at science as a whole but as separate parts like Biology, Chemistry, and Physics. Science is also the cyclotron, the A-bomb, and the computer. In short, science is a mass of unrelated ideas and objects.

Posttreatment

Today I think of science basically as a process and the end products of that process, i.e., theories and concepts and laws. I no longer view it as an area of knowledge that employs complicated machines and complex models, but as a group of basic principles upon which we base new ideas and concepts. I think of science as a pure, more or less mental process, and technology as the application of this pure science.

Subject 17

Pretreatment

Science is discovering by hypothesis and experimentation and involves all forms of nature and its componants. It also includes the physical sciences and has as its purpose: to explain the world in which we live. After explaining "it" then puts the theories to work for mankind. Our society has therefore become oriented and bows down to the methodic ways of the god Science. After realizing the horrors involved with the A-Bomb and the fear of the unknown in space, I have come to fear science. It seems forever in quest for understanding the unknown, and although it may have helped in history and in conviences for today -- it seems to be taking the world down a very scary and dark path. It may be so that when other revolutions occurred in the past that people were unsure and that this is what is happening now. I seem to be unable to flourish in our scientific revolution without some hesitation.

Posttreatment

Science in my view is the "pure" science, that which questions and explores all realms of nature solely for the purpose of gaining the knowledge therein. The main purpose is to explain every natural phenonama by means of observation and experimentation incorporated into theories; which may be altered as new results affect them, or laws; which in turn will explain future theories. Science is ever changing, and often discards theories and basic ideas in the light of new, disproving evidence. This evidence must be impartial, as the final end should be understood by any one of the world's scientists, and all laws must have universal agreement. Nature is impartial -- all men should be able to see, eventually, the truth of its phenonoma -- there is no room for bais, prejudice or narrow mindedness.

Therefore science is a direct and open minded inquiry and explanation of all realms of natural phenomana. It seeks these ends for the ends alone.

Subject 18

Pretreatment

I view science as various fields of study by scientists. Some scientists are biologists, chemists, physicists, and others. All of these scientists have varied backgrounds, which is necessary for their field of work. The scientist tries to find out various things in his field, and may try to teach these things he has found to students in classes. His instruction is his way of telling someone else what his field is about.

The material in science is to be learned and kept as much as possible in the mind. The greater amount of material you have, the better you can possible do in science, in understanding it and using it when you can.

Most of the material in science is specific information in the various fields of science. Formulas of various types can be found, because of the varying fields.

I could try to go on and take other sorts of material, but I don't think, to myself, anyway, that this is necessary to give a general view of what I think science is.

Posttreatment

Science is a field of study, which is unrelated and related to society. Some areas are used to find out various things which could be used to benefit society. Others are used to learn more about the universe as a whole.

Thus, science is something which is somewhat difficult to understand as one simple subject. Science is more than that, being two-fold.

Subject 19

Pretreatment

I wish I were better in science. It seems to me to be one of the most vital areas of work today. Science is growing constantly, making discovery after discovery. It has made a massive effect on our lives. A person going into science has all sorts of doors open to him. The scientist is wanted and needed by society, and society is willing to pay the price to get him. The scholarships and grants and high paying jobs are open to him. I associate science with action, with progress, yet, I know, that a world of only science would be empty of the values and thoughts and ideals that make our lives fuller and deeper.

Posttreatment

Science is the means to a better material world for most people. For a few, it is an intellectual disinterested activity, as most intellectual activities seem to be. Many people seem to place a futile trust in science to replace a diminishing trust in God or a spiritual world. It seems to me, though, that science is too confined in its scope and arid in its spiritual value to act as a God replacement. Science today doesn't have the power in itself to give final answers, and I think most scientists are aware of that. It's unfortunate that they haven't been willing or perhaps able to communicate the meaning of science to the rest of the world. This is a science-oriented world in many ways, and yet no one seems to really know what science is all about. I can't help feeling that even the scientists themselves are uncertain of their place in society.

Subject 20

Subject 20 dropped the course.

Subject 21

Pretreatment

My view of science is more of a technological "help American society" concept. I realize that there are theorists and true scientists lurking behind the scenes developing new ideas, but the technologists are the ones who keep giving us "mirical flouriments" for our tooth paste and sugar free sugar, so they should come to mind first. I feel that science is the most important factor in our lives today, because science is striving to create new and improved aids for life. The technologists are doing a much greater part, than, for instance, the writers of novels and literature, who centre around only one facet of our lives, that of our leisure time and how we choose to spend it.

Posttreatment

Science is discoveries that makes life easier. Science developes things which allow us to live better than any of our ancestors.

It is also an endevor to improve things that exist.

It is not only old men in white coats in a nice clean laboratory, but also the technologists who create things with the scientists' discoveries.

Subject 22

Pretreatment

I view science as being a cummulation of knowledge. This knowledge when applied is for the purpose of making this world a better place in which to live. In other words, science's purpose is to aid man in a "material sense".

Science uses various procedures to acquire this knowledge. Such things as assumptions, experimentation, data and conclusions are prety much standard procedure in science. The results or findings of science unlike many fields cannot be thought of as always being true. For example, ideas in science in previous times have been discarded through the discovery of new facts. These new facts likewise may become obsolete in time with new developments in equipment etc. and the laws which explain these facts may also be found to be false in time.

Posttreatment

Science as I see it is concerned with increasing man's body of knowledge. This knowledge's aim is not to aid in practical products to society, although scientific knowledge often results in this. Sciences view is merely knowledge for knowledges sake alone.

Science has no set rules for investigation. Scientists may use almost any creative or imaginative means to yield facts. It is in this sense that it is aesthetic.

Science asks questions of why. Why is something the way it is? Not just on the practical level as: How did it become this way? Science searches for the ultimate nature of things. In doing so it may step on the feet of other things such as religion, but its purpose is not to refute these other bodies but merely to understand the universe.

LIST OF REFERENCES

LIST OF REFERENCES

Books and Pamphlets

Allen, G. E., 1966. General Education in the Sciences: The Role of History and Philosophy, Commission on Undergraduate Education in the Biological Sciences, III.

Allport, G. W., Vernon, P. E., and Lindzey, G., 1951. A Study of Values: A Scale for Measuring the Dominant Interests in Personality, Houghton Mifflin, Chicago.

Blommers, P. and Lindquist, E. F., 1960. Elementary Statistical Methods, Houghton Mifflin Company, Chicago.

Conant, J. B., 1951. On Understanding Science, Mentor Books, New York.

Cooley, W. W. and Klopfer, L. E., 1961. Manual for the Test On Understanding Science, Educational Testing Service, Princeton.

Dressel, P. L., and Mayhew, L. B., 1954. General Education: Explorations in Evaluation, American Council on Education.

Educational Policies Commission, 1944. Education for All American Youth, National Education Association.

Guilford, T. P., et al, 1951. "A Factor-Analytic Study of Creative Thinking", Reports From the Psychological Laboratory, University of Southern California.

Klopfer, L. E., 1960. History of Science Cases for High Schools, Department of School Services and Publications, Wesleyan University.

Lindquist, E. F., 1953. Design and Analysis of Experiments in Psychology and Education, Houghton Mifflin Co., Chicago.

McGeoch, J. A. and Irion, A. L., 1952. The Psychology of Human Learning, Longmans, Green, New York.

McLindon, J. C., 1960. Teaching the Social Studies, American Educational Research Association, National Education Association.

Meredith, W. M., 1967. Basic Mathematical and Statistical Tables for Psychology and Education, McGraw Hill Book Co., New York.

Ray, W. S., 1960. An Introduction to Experimental Design, Macmillan Co., New York.

Siegel, S., 1956. Nonparametric Statistics for the Behavioral Sciences, McGraw Hill Book Co., New York.

Snow, C. P., 1964. The Two Cultures, Mentor Books, New York.

Spranger, E., 1928. Types of Men, Stechert Publishing Co., London.

Stice, G., 1958. Facts About Science, Manual for the Facts About Science Test, Educational Testing Service, Princeton.

Tate, M. W., 1955. Statistics in Education, Macmillan Company, New York.

Walker, H. M. and Lev, J., 1953. Statistical Inference, Holt and Co., Inc, New York.

Watson, G. and Glaser, E. M., 1964. Watson Glaser Critical Thinking Appraisal, Manual for the Critical Thinking Appraisal, Harcourt, Brace, and World, Inc., New York.

Articles and Periodicals

Bloom, B. S., 1953. Thought Processes in Lectures and Discussions. Journal of General Education, 7.

Caldwell, O. W., 1931. Students Attitudes Regarding Unfounded Beliefs. Science Education, XV.

Carpenter, F., 1956. Educational Significance of Studies on the Relation Between Rigidity and Problem Solving. Science Education, 40.

Campbell, D. T., and Stanley, J. C., 1963. Experimental Designs for Research on Teaching. Handbook of Research on Teaching, Handbook of Research on Teaching, N. L. Gage, ed., Rand McNally and Co., New York.

Cooley, W. W. and Klopfer, L. E., 1963. The Evaluation of Specific Educational Innovations. Journal of Research in Science Teaching, 1.

Crumb, G. H., 1965. A Study of Understanding Science Developed in High School Physics. Journal of Research in Science Teaching, 3.

Cummings, H. H., 1957. The Social Studies in a Scientific Age. Science and the Social Studies, 27th Yearbook, National Council for the Social Studies, NEA Publication.

Dewitt, P. P., 1959. Attitudes Related to the Study of College Science. School Science and Mathematics, XXXIX.

Dressel, P. L., 1960. How the Individual Learns Science. Rethinking Science Education, 59th Yearbook, National Society for the Study of Education, University of Chicago Press.

Ennis, R. H., 1962. A Concept of Critical Thinking. Harvard Educational Review, 32.

Furst, E. J., 1950. Relationship Between Test of Intelligence and Tests of Critical Thinking and Knowledge. Journal of Educational Psychology, 43.

Guilford, J. P., 1950. Creativity. American Psychologist, 5.

Henderson, K. B., 1958. The Teaching of Critical Thinking. Phi Delta Kappan, 39.

Holton, G. 1965. "Introduction", Science and Culture--A Study of Cohesive and Disjunctive Forces, G. Holton, ed., Houghton Mifflin Co., Chicago.

Houle, C. O., 1943. Evaluation in the Eight Year Study. Curriculum Journal, 14.

Kleinman, G., 1965. Teachers' Questions and Student Understanding of Science. Journal of Research in Science Teaching, 3.

Klopfer, L. E. and Cooley, W. W., 1963. The History of Science Cases for High Schools in the Development of Understanding of Science and Scientists; A Report on the HOSC Instruction Project. Journal of Research in Science Teaching, 1.

Pella, M. O., O'Hearn, G. T. and Gale, C. W., 1966. Referents to Scientific Literacy. Journal of Research in Science Teaching, IV.

Rogers, E., 1960. The Research Scientist Looks at the Purposes of Science Teaching. Rethinking Science Education, 59th Yearbook, National Society for the Study of Education, University of Chicago Press.

Strauss, S., 1931. Some Results for the Test of Scientific Thinking. Science Education, XVI.

Trent, J. H., 1965. The Attainment of the Concept Understanding Science Using Contrasting Physics Courses. Journal of Research in Science Teaching, 3.

Washton, N. S., 1966. Teaching Science for Creativity--Needed Research. Journal for Research in Science Teaching, IV.

Wessell, G., 1941. Measuring the Contribution of Ninth-Grade General Science Courses to the Development of Scientific Attitudes. Science Education, XXV.

Zapf, R. M., 1938. Suggestions of Junior High School Pupils, Part II, Effects of Instruction on Superstitious Beliefs. Journal of Educational Research, XXXI.

Unpublished Material

Kimball, M. E., 1965. Opinions of Scientists and Science Teachers About Science. Unpublished Ed. D. Dissertation, Stanford University.

Snider, R. M., 1966. A Project to Study the Nature of Physics Teaching Using the Flanders Method of Interaction Analysis. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Chicago.

Sorensen, L. L., 1966. Change in Critical Thinking Between Students in Laboratory-Centered and Lecture-Demonstration Centered Patterns of Instruction in High School Biology. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Chicago.