DOCUMENT RESUME

ED 151 277 SO 010 727

AUTHOR Cort, H. Russell, Jr.; Peskowitz, Nancy
TITLE A Longitudinal Study of Man: A Course of Study. Volume II: Quantitative Results. Final Report.
INSTITUTION Antioch Coll., Washington, D.C.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE 30 Dec 77
GRANT SED-72-06289-A04
NOTE 264p.; For related documents, see SO 010 725-729 ; Product of the Social Science Research Project

EDRS PRICE MF-$0.83 HC-$14.05 Plus Postage.
DESCRIPTORS *Academic Achievement; Anthropology; Classroom Environment; Course Evaluation; Cultural Differences; *Curriculum Evaluation; Data Analysis; Educational Practice; Educational Research; Elementary Education; Followup Studies; Grade 5; Grade 6; Measurement Techniques; Methods; Post Testing; Pretests; Process Education; Program Evaluation; Research Methodology; *Social Studies; *Statistical Analysis; *Student Attitudes; Student Characteristics; Tables (Data); Teacher Characteristics
IDENTIFIERS *Man A Course of Study

ABSTRACT

This second volume of the summative evaluation of "Man: A Course of Study" (MACOS) presents results of quantitative analyses of what MACOS students seemed to learn, what they retained one year later, and how what they learned was different from what students in other social studies courses learned. The first part of the document compares MACOS and non-MACOS classes with regard to characteristics of students and teachers, and climate and processes in the classroom. Characteristics computed for classes included ability grouping; average age of students; and percentage of students who were white, female, and male. Characteristics computed for teachers included educational background, social studies objectives, and preferred teaching methods. No significant differences were found between the two groups of classes with regard to pretests of achievement, attitude, demographic characteristics, or class composition. The typical class in both groups was predominantly white, urban, and non-low income. The second part of the document presents pre- and posttest findings, and follow-up information. All classes were given attitude and achievement tests. Pretest instruments were also administered to teachers. Findings indicated that MACOS students scored significantly higher than non-MACOS students on the MACOS-specific posttest; scored similarly on more generalized tests of social studies skills; learned more about facts and terms than abstract concepts; and tended to like social studies more than non-MACOS students. (Author/DB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED151277

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Alphonse Buccino

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

FINAL REPORT

A LONGITUDINAL STUDY

OF

MAN: A COURSE OF STUDY

VOLUME II

QUANTITATIVE RESULTS

H. Russell Cort, Jr.
Nancy Peskowitz

Social Studies Research Project
Antioch College
1624 Crescent Place N.W.
Washington, D.C. 20009

December 30, 1977

This study was carried out under National Science Foundation Grant No. SED 72-06289 A04. Any opinions, findings, and conclusions or recommendations expressed in this report are those of the authors and do not necessarily reflect the views of the National Science Foundation.

SO 010 227

III. QUANTITATIVE RESULTS

A. Introduction

In this section, results of quantitative analyses are presented. The section is organized in two main parts: 1) analyses of the beginning characteristics of groups; and 2) findings. The first deals with the characteristics of MACOS and non-MACOS groups at the out . The second part presents the results of analyses of groups with respect to pretest, posttest, Follow-up 1, and Follow-up 2 outcomes.

In most analyses, variables have been treated as falling into one of three major groups:

- Inputs - measures of initial characteristics of students, teachers, and classes
- Processes and classroom climate[1] - measures of activities and attitudes toward the classroom made during the baseline year (after pretest, before posttest; sepcifically, Feb./Mar., 1975, midtest 2)
- Outcomes - posttest and follow-up achievement and attitude variables.

Furthermore, in most analyses the class is the unit of analysis. This was done to eliminate the correlation of scores among students within classes, and thus to provide independent units of analysis. Thus measures of student attitude or achievement are class means. Those means are based on results from students who took both pre and posttest. One consequence of this approach is that conclusions apply to classes, not to individuals. Analyses of the data collected in this study

1. Classroom climate (satisfaction with class, for example) could be regarded as an outcome. Some analyses were made treating climate variables as outcomes. Most analyses treated climate as a moderator or predictor of outcomes.

have not attempted to address questions such as: what kind of student seems to derive more (or less) from MACOS? Are there aptitude-context interactions (differential structural or group effects)? Are there aptitude-context-treatment interactions?[2]

Characteristics of classrooms were based on class percentages (for example, percentage of females; percentage of students for whom English was a second language; percentage of students not eligible for free lunch program; percentage of 5th grade students - to take non-graded classes into account - etc.).

Characteristics of teachers were treated as another set of attributes of classrooms. For cases in which a class had more than one teacher (e.g., team teaching situations), the attribute was the average for the teachers on variables for which it made sense to average (e.g., sex was not averaged, and not used as an attribute; Educational Scale VII scores were averaged and used). Class characteristics pertaining to students were based on all students for whom the project had data, not just pre-post students.

Analyses were also made of responses to some items or questions in pretest, posttest and Follow-up 1 and 2 in which the individual is the unit of analysis. In those cases, care has been taken to identify such analyses clearly and to delimit interpetations accordingly.

The aims of analyses have been twofold: 1) to determine whether there were differences in outcomes between MACOS and non-MACOS classes; and 2) to investigate relationships, between groups and within groups, of input and pro-

2. Analyses have been made of the possible effects of districts and schools.

cess variables to each other and to outcomes. The basic plan of analysis has been to: 1) create a reduced number of input and process/climate variables by a principal components analysis of sets of variables; 2) analyze differences in outcome between and within groups using analysis of variance and covariance, and fixed order multiple regression techniques; 3) analyze relationships among sets of variables for both groups and within each group, using canonical correlation analyses. At different stages, supplementary analyses have been made to test hypotheses related to the stage.

Several critical points need to be stated at the outset. First, the quantitative analyses and conclusions to be discussed in this section are not the total findings of this study. Much time and effort was spent obtaining and examining qualitative information from students and teachers. The information obtained bears importantly on questions of what was done in classes, what the context of classes was, what students made of their classes, what teachers saw as strengths and problems of different courses, how students remembered their social studies course the following year, and so on. The results of repeated interviews with students and teachers are given in Section V of this report. It is believed that they contribute substantively to a more complete understanding of MACOS and non-MACOS classes than can be obtained solely from the quantitative analyses. Other sources of information have also been examined and analyzed to obtain a fuller understanding of what teachers in both groups did and why. Second, a number of theoretical and methodological issues and problems are inherent in a study of this kind. One is the problem of inferring causality from correlations. The problem arises in part from the fact that the design of the

study is not experimental. It also arises from the fact that detailed, continuous observation of classes was not possible. Therefore, determination of what was actually done (the specification of processes) has to be inferred from secondary sources (ratings and reports by students and teachers, interviews with students and teachers). Also, there are issues concerning the fallibility of measures, and appropriate methods of analysis.

There is a large literature on the methodological problems and pitfalls involved in the analysis of non-experimental data and the interpretation of the effects of correlated and fallible variables. Analyses reported in this section have attempted to take cognizance of the methodological (and conceptual) problems in the field. The study has employed the strategy of cross-checking by use of several different analytic methods. But even that strategy does not assure decisive conclusions. Different methods or procedures apply different models and ask different questions of the data. Not all possible methods were employed. No one method can resolve unresolved theoretical issues, nor is it clear what combination of methods is optimum in the face of particular theoretical uncertainties. In the absence of replication and cross-validation, conclusions must be tempered.

The spirit of this inquiry has been to illuminate, not to pass judgment, but even illumination has its limitations in time and application. The most that this study says is that if one wants to use MACOS, here are some factors to consider based on how it was used by a number of teachers and their results. It is a study of classes in various settings in which no attempt was made to influence

what should be taught or how. On the contrary, explicit effort was made to avoid that. It is thus a study of a variety of MACOS and non-MACOS courses that were implemented in a variety of ways in a variety of settings.

B. Initial Characteristics of Groups

1. Classes and Districts

There were 108 classes distributed among 15 districts in 11 states.[3] The distribution of classes, by group and grade level, is shown in Table III-1.

Table III-1
Number of Classes by Group and Grade Level

Group	5	6	Non-Graded	Total
MACOS	20	20	17	57
Non-MACOS	24	20	7	51
Total	44	40	24	108

Table III-2 shows the distribution of classes by group and district.

Table III-2
Number of Classes by Group and District

District	MACOS	Non-MACOS	Total
1.	5	5	10
2.	4	4	8
3.	4	4	8
4.	3	3	6
5.	3	4	7
6.	2	2	4
7.	4	4	8
8.	7	4	11
9.	8	3	11
10.	4	4	8
11.	4	4	8
12.	1	3	4
13.	2	4	6
14.	4	4	8
15.	2	1	3
Total	57	51	108

3. California, Colorado, Florida, Illinois, Iowa, Nebreska, New Jersey Oregon, Pennsylvania, Virginia, Washington. Districts, teachers and students were promised anonymity in this study. Every effort has

The process by which districts, school and teachers entered the project were described in Section IIA. A further description of a number of characteristics of districts is given in Appendix B, Vol. III.

2. Assignment Characteristics of Classes

In order to ascertain whether there were systematic differences between MACOS and non-MACOS classes with respect to assignment of students, teachers were asked how classes were formed. In the majority of cases in both groups classes were formed on a random or heterogeneous basis, or on no special basis. In some cases teachers described classes as consisting of particular ability groups or formed on other bases such as interest, compatibility with the teacher, absence of reading problems, etc. The overall distribution of assignment characteristics for the two groups is shown in Table III-3.

Table III-3
Bases of Assignment to Classes

	No Special Basis Random, Heterogeneous Grouping	Some Form of Grouping or Selection	No Data	Total %*	N
MACOS	66.7%	24.6%	8.8%	100.1%	57
Non-MACOS	76.5	13.7	9.8	100	51

*May not add to 100% due to rounding

There was somewhat more variability of assignment characteristics in the MACOS classes. The two overall distributions were not, however, significantly different ($X^2 = 2.019$, $df = 2$, $p > .30$).

been made in this report to protect that anonymity. The computer data files contain no names, only code numbers. Coding has not been done according to alphabetical orders or other systematic bases.

3. Characteristics of Students: Pretests

Pretest characteristics of students, based on class averages, or half-class averages, are listed in Table III-4. Table III-4 shows the overall means, standard deviations (SD), standard error of the mean (SEM), the observed range of class means, and the number of classes (N) for the MACOS and non-MACOS groups. In the cases of the Questionnaire About Animals and People (AP) and the STEP test, the data are given for sub-tests as well as total scores.

It is important for the reader to bear in mind that Table III-4 gives: 1) statistics on classes (or half-classes), not on individual students; 2) statistics based only on students known to have been in a particular class all year; and 3) in the case of attitude scales, statistics based on average scale scores for individuals who passed predetermined exclusion rules with respect to number of scale items completed validly.[4] Achievement means are based on total raw scores. Thus, the figures pertain to classes and were derived from students who were associated with a class all year, and who, in the case of attitude measures, had what was deemed to be sufficiently valid data to be used in computing class means. With one exception noted below no comment on the significance of differences

4. For example, WWA consisted of 5 items. If a student omitted more than one on pretest, a mean was not computed for the student; similarly for posttest. Either case would exclude the student from the class means, pre and post. If the student omitted one item on pretest and one on post-test, the pre-post averages were still computed for the student and included in the class means. The exception was Study Choices (SS Ch). That was scored by counting the number of times social studies was chosen when paired with another subject. If a student omitted (or marked invalidly) one or more pairs containing social studies, the scores were not included in the class pre-post means. Generally losses were small by the exclusion criteria applied.

Table III-4

Means, Standard Deviations (SD), Standard Errors of the Means (SEM), Range, and Number of Classes (N) for Pretests, by Group [1]
(Pretest)

Instrument	MACOS Mean	SD	SEM	Range	N	Non-MACOS Mean	SD	SEM	Range	N
1. Animals and People (AP)										
a. Total Score	19.01	4.26	.56	9.67-26.57	57	18.74	3.58	.50	9.00-25.54	51
b. Questions 1-4. Animals part	10.46	2.17	.29	3.50-15.20	57	10.06	1.69	.24	4.75-13.77	51
c. Questions 5-8 Netsilik part	8.55	2.66	.35	1.00-13.28	57	8.68	2.25	.32	3.89-13.00	51
2. STEP										
a. Total Score	28.79	5.73	.77	15.38-39.88	56	28.08	5.60	.79	16.84-40.60	50
b. Sub 1, Organizing Information	3.65	.80	.11	1.50-5.32	56	3.55	.77	.11	1.96-5.05	50
c. Sub 2, Interpreting Information	14.27	2.89	.39	8.00-19.68	56	13.93	2.78	.39	8.59-20.95	50
d. Sub 3, Assessing Adequacy of Data	2.80	.61	.08	1.50-4.25	56	2.71	.67	.09	1.26-4.05	50
e. Sub 4, Drawing Inferences, Generalizing	6.59	1.24	.17	3.33-8.92	56	6.50	1.21	.17	3.84-9.00	50
f. Sub 5, Reaching Conclusions	1.48	.38	.05	.86-3.00	56	1.38	.37	.05	.64-2.05	50

1. See Appendix E (Vol III) for pre, post and follow up statistics for these measures.

Table III-4 Continued

Instrument	MACOS Mean	SD	SEM	Range	N	Non-MACOS Mean	SD	SEM	Range	N
3. Interpretation of Data Test (IDT)	8.95	1.45	.19	5.80-13.11	57	8.78	1.54	.22	5.46-12.00	51
4. Social Studies Choices (SS Ch)	2.33	.75	.10	.86-4.44	57	1.99	.72	.10	.42-4.04	51
5. What Would You Think										
a. WWA	4.87	.26	.03	4.38-5.39	57	4.80	.31	.04	4.20-5.35	51
b. WWB	5.48	.27	.04	4.87-6.01	57	5.40	.36	.04	4.58-6.11	51
6. Children's Attitude Toward Problem Solving (CAPS, Factored)										
a. CAPS-1, Ability	2.84	.26	.03	1.95-3.36	57	2.88	.27	.04	2.30-3.40	51
b. CAPS-2, Interest	3.79	.20	.03	3.33-4.15	57	3.82	.19	.03	3.17-4.20	51
c. CAPS-3, Tolerance of Ambiguity	2.91	.26	.03	2.22-3.44	57	2.87	.25	.04	2.40-3.43	51
d. CAPS-4, Creativity	3.11	.20	.03	2.67-3.88	57	3.12	.19	.03	2.66-3.57	51

between groups will be made in this sub-section. Results of multivariate tests of beginning characteristics of the MACOS and non-MACOS groups will be given in Section III C-2.

4. Characteristics of Classrooms

A number of characteristics of classes were computed. These characteristics, such as average age, were based on data provided for each student by the teacher. Eleven such variables were computed for each class:

- average age of students, in months
- percentage of students who were female
- percentage of students who were white
- percentage of students not eligible for the free lunch program
- percentage of students for whom English was the primary language, not a second language (not ESL)
- percentage of students who had not previously had any MACOS
- percentage of students who were 5th graders (to take non-graded classes into account)
- average reading level, based on a 5-point scale on which the teacher classified each student on the basis of his/her latest reading achievement scores (1= more than 1 year above grade level; 5 = more than 1 year below grade level)
- average number of years students in class had been in present school
- size of class, based on the total number of students known to be in the class at time of posttesting, even if not at pretesting
- stability, the ratio of students for which there were pre-post data to the total number of students for whom there were any

pre, midtest or post data.[5]

Table III-5 gives the mean, SD, skewness (SK), kurtosis (K), observed range and number of classes for each group.

5. Characteristics of Teachers

Descriptive statistics of characteristics of the MACOS and non-MACOS teachers that are used in quantitative analyses of results and relationships are given in Table III-6. The table shows for each group the mean, standard deviation, standard error of the mean, observed range, and sample size for each variable.

The Objective Categories listed in Table III-6 came from the Social Studies Program Survey Form, Part II. Part II contained a list of 40 objectives. Groups of objectives were intended to be related to different major program emphases listed in Part I of the form. Teachers were asked to rate each objective with respect to its importance. Ratings of the group of objectives in each category were summed and averaged. The abbreviations listed in Table III-6 refer to the following category titles:

- Cat 1 (CT): citizenship transmission objectives
- Cat 2 (CD): cross-disciplinary/humanities/integrated concept objectives
- Cat 3 (IMP): inquiry modes and process objectives
- Cat 4 (SA): self-actualization objectives

5. This is not, obviously, a beginning characteristic of a class, nor a pure measure of stability. It is included here as an index of the class on which student means were based, relative to all students known to have been associated with the class at any time.

Table III-5
Mean, Standard Deviation (SD), Skewness (SK), Kurtosis (K), Range, and Number of Classes (N) for Classroom Characteristics, by Group

Characteristic	MACOS Mean	SD	SK	K	Range	N	Non-MACOS Mean	SD	SK	K	Range	N
1. Age in months	131.6	5.67	.42	-1.07	121.3-134.5	57	132.0	6.22	-.96	-1.29	120.6-142.7	50
2. Percent females	48.96	10.10	.16	1.35	20.00-77.78	57	51.09	9.99	-.11	-.20	31.03-76.92	51
3. Percent white students	87.91	16.66	-1.60	1.50	37.50-100	57	87.36	18.06	-1.73	2.28	29.06-100	51
4. Percent not eligible for free lunch	86.42	13.42	-1.20	.81	48.28-100	57	82.66	18.15	-1.61	2.95	16.67-100	50
5. Percent not ESL	98.38	4.12	-2.65	5.86	82.14-100	57	97.19	6.71	-3.66	15.66	60.87-100	51
6. Percent not previous MACOS students	93.11	19.58	-3.51	11.13	8.70-100	57	99.07	6.07	-6.94	45.45	57.14-100	50
7. Percent 5th graders	47.97	43.45	.10	-1.65	0-100	57	53.69	47.22	-.14	-1.84	0-100	51
8. Reading level (5-point scale)	2.81	.61	.64	1.11	1.28-4.56	57	2.74	.54	-.02	-.11	1.48-4.09	50
9. Years present school	3.91	1.09	-.26	.28	1.36-7.00	57	4.18	1.12	-.35	.30	1-6.7	51
10. Class size	28.98	10.87	1.88	8.96	9-82	57	28.20	6.00	-.37	.68	13-40	51
11. Stability	.73	.17	-.61	-.17	.3-1.0	57	.78	.13	-1.14	1.18	.4-.9	51

Table III-6

Means, Standard Deviations (SD), Standard Errors of the Mean (SEM), Ranges and Sample Sizes of Selected Teacher Variables, by Group

Variable	MACOS Mean	SD	SEM	Range	N	Non-MACOS Mean	SD	SEM	Range	N
1. Years teaching experience	9.26	7.37	1.01	0-33	53	9.80	7.67	1.16	1.00-40	44
2. Years taught in present district	6.04	5.72	.79	0-27	53	7.18	5.63	.85	1.00-28	44
3. Years taught present program (1st year =1)	2.39	1.06	.16	1-4	46	2.51	3.11	.51	1.00-16	37
4. Educational Scale VII										
a. Progressivism score	86.21	8.41	1.17	65-103	52	85.51	9.11	1.36	66-105	45
b. Traditionalism score	53.00	9.81	1.36	36-86	52	58.91	13.45	2.00	18-83	45
5. Teachers at Work (TAW)	14.42	3.37	.46	7-24	53	15.74	4.05	.60	8-26	46
6. Objectives										
a. Cat 1 (CT)	2.34	.62	.08	1-3.6	54	2.38	.55	.08	1.2-3.6	48
b. Cat 2 (CD)	1.67	.55	.07	1-3.0	54	1.91	.52	.08	1-3.7	48
c. Cat 3(IMP)	1.59	.51	.07	1-3.0	54	1.92	.50	.07	1-3.2	48
d. Cat 4 (SA)	1.94	.56	.08	1-3.6	54	2.35	.68	.10	1-4.0	48
e. Cat 5 (SS)	2.48	.61	.08	1-4.6	54	2.66	.56	.08	1-3.0	48
f. Cat 6 (V)	1.71	.49	.07	1-3.0	54	1.94	.53	.08	1-3.0	48
g. Polit	2.40	.72	.10	1-4.0	54	2.36	.60	.09	1-3.7	48
h. Gen	1.96	.51	.07	1-3.4	54	2.08	.41	.06	1.4-3.0	48

. Cat 5 (SS): the social sciences objectives

. Cat 6 (V): values clarification objectives

. Polit: political relevance objectives

. Gen: general social studies objectives

There were other background characteristics of the two groups of teachers that were not selected for use in major quantitative analyses. They are given below in a series of tables in order to provide a more complete descriptive summary of the beginning characteristics of the two groups. These tables will not be numbered separately, to save space. They are simply identified by variable. All tables are in percentages of teachers in each group for whom there were data. It should be noted that sample size may be greater in some cases than the number of MACOS or non-MACOS classes. That is because there were a few cases (in both groups) in which more than one teacher, (for example, in a team-teaching situation) completed the Teacher Master Record Form. Those responses are included here.

Sex

	Male	Female	N
MACOS	47%	53%	55
Non-MACOS	35	65	48

Race

	1	2	3	4	5	N
MACOS	2%		97%	2%		57
Non-MACOS	2		98			53

1. Black
2. American Indian
3. White
4. Oriental
5. Other

Age as of January 1974

Age	MACOS(N=58)	Non-MACOS(N=51)
20-25	22%	29%
26-30	29	14
31-35	19	21
36-40	9	12
41-45	10	12
46-50	5	2
51-55	2	4
56-60	2	2
61 or over	2	4
	100%	100%

Identify with an Ethnic Minority?

	Yes	No	N
MACOS	13%	87%	53
Non-MACOS	24	76	45

Education: Degrees or Certificates Earned (Cneck all Applicable)

	1	2	3	4	5	6	N
MACOS	20%	76%	9%	19%	3%	19%	59
Non-MACOS	23	67	12	18	10	18	49

1. AA
2. Bachelor's Degree
3. MAT
4. Other Master's Degree
5. Master's plus 30 hours
6. Other

Type of Teaching Situation at this time (Check all applicable)

	1	2	3	4	5	6	7	N
MACOS	43%	45	15	28	11	23	4	53
Non-MACOS	60%	38	23	13	2	21	11	47

1. Self contained
2. Team teaching
3. Departmentalized
4. Open-Space
5. Demonstration classroom
6. Non-graded
7. Other

What subject do you most enjoy teaching?

Subject Preferences	MACOS(N=57)	Non-MACOS(N=49)
Social Studies	21%	22%
Math or Science	23	20
Language Arts	12	10
Reading	12	6
Other Single Subjects	4	10
All Other (Combinations)	28	31
	100%	99%*

*Does not add to 100 due to rounding

Some teachers responded to this question by listing several subjects. The lists given by different teachers sometimes included social studies, sometimes not. Teachers giving multiple responses are classified above under 'All Other (Combinations)'. The category also includes two teachers who said 'All Subjects', and one who said 'No Preference'. The category 'Other Single Subjects' includes such subjects as music, art, health, etc.

Special Areas of Training Related to Social Studies: Pre-Service or In-Service

	1	2	3	4	5	6	7	8	9	10	11	12	N
MACOS	73%	62	75	55	50	60	57	63	53	37	45	32	60
Non-MACOS	62%	58	62	57	42	51	55	51	58	30	42	38	53

1. Social and emotional development
2. Developing cognitive skills
3. Teaching inquiry methods
4. Teaching how to analyze values and value conflict
5. Teaching interpersonal skills to students
6. Teaching social science methods and techniques
7. Developing self awareness in students
8. Use of questions as an educational method
9. Leading and/or evaluating classroom discussions
10. Teaching how to analyze social issues
11. Integrating social studies with other subjects
12. How to increase relevance of subject matter to students

Pre-Service Training in Social Studies
(All training before becoming a teacher of record)

. Did training include course(s) in social studies methods?

	Yes	No	N
MACOS	83%	17%	60
Non-MACOS	75	25	51

. If Yes, was it undergraduate, graduate, or both?

	Undergraduate	Graduate	Both	N
MACOS	66%	6%	28%	50
Non-MACOS	66	11	24	38

. Did pre-service teacher training include any other courses specifically concerned with teaching social studies?

	Yes	No	N
MACOS	21%	79%	58
Non-MACOS	21	79	47

In Service Training in Social Studies
(All training received since becoming a teacher of record)

. Have you received any in-service training pertaining to teaching social studies?

	Yes	No	N
MACOS	67%	33%	58
Non-MACOS	48	52	44

The Verbs for Objectives Form (VO), described in Section IIB 2 1, was not used in major quantitative analyses. This instrument was intended to provide one measure of teachers' tendencies to consider application-oriented objectives as important. It was scored by counting the number of times out of 6 a high application rated verb was selected from an alphabetized list of 30 verbs. The distributions of percentages of teachers choosing a given number of application verbs are shown below:

Number of Application Verbs Chosen Out of Six

	0	1	2	3	4	5	6	N
MACOS	18%	34%	23%	21%	4%	0%	0%	56
Non-MACOS	24	45	24	8	0	0	0	51

Based on a hypergeometric probability distribution, the exact probability of each possible number of selections is:

Number of Selections	P
0	.2267
1	.4295
2	.2684
3	.0682
4	.0070
5	.0002
6	.0000

From this it can be seen that the cumulative probability of picking 0 to 1 verb by chance is .6562, while the cumulative probability of picking 2 or more by chance is .3438. Each distribution (MACOS, non-MACOS) of selections was tested against a chance expectation by a Chi-square test of frequencies combined into 3 categories: 0, 1, 2 or more selections. For the MACOS groups, $X^2=3.690$, df=2, $p > .10$. For the non-MACOS group, $X^2=.206$, df = 2, $p > .90$. Thus while neither distribution exceeded a chance expectation, the MACOS group was somewhat more likely to pick a larger number of application type verbs than the non-MACOS group. When the distributions of the two groups were compared with each other, they were found not to differ significantly ($X^2=5.120$, df = 3, $p > .10$). Considering the whole distributions for each group, an index of predictive associ-

tion was computed.[6] If Group is taken as the dependent variable there is a 12% reduction in error of predicting group membership, knowing number of selections. There is, however, no reduction in error of prediction in the opposite direction. Since the overall distributions for the two groups were not significantly different, it was concluded that the strength of predictive association indices should not be regarded as of consequence.

C. Findings

1. Creation of Input and Process/Climate Principal Components (PC's)

To reduce the number of input and process variables, principal component analyses were made of sets of variables.[7] Sets were logical groups of variables. For example, one set was the three achievement tests: A Questionnaire about Animals and People (AP), the Interpretation of Data Test (IDT), and the STEP, Social Studies, Series II, Form 4a (STEP). Another set was the group of computed classroom characteristics such as average age of students, percentage of females, class size, etc. The consequence of this approach is that resultant principal components from different sets of variables were not necessarily uncorrelated. The reason for using logical groupings of variables was to retain interpretability of resultant composites. While it was possible to create a limited number of composites from an ensemble of meaningfully different variables which would have been uncorrelated, it was believed that, parsimony notwithstanding, the the ability to interpret relationships of composites to dependent variables was the

6. Hayes, William L., Statistics for Psychologists. New York: Holt, Rinehart and Winston, 1963, pgs. 606-610.

7. Nie, Norman H., et. al, SPSS: Statistical Package for the Social Sciences Second Edition. New York: McGraw Hill, 1975, PA 1 routine. Listwise deletion was used.

more important consideration.

Analysis was done using the total array of classes from both groups. For any set of variables, a maximum of two principal components (which are the first two with the largest variances) was retained. The decision rule employed was that if the first principal component (PC) accounted for 50% or more of the variance for the set, only it was retained. Otherwise, the second PC was retained in addition. The reason for this decision rule was parsimony. The aim was to reduce the number of covariates relative to sample sizes. There were eight sets of variables for which principal components were obtained and four of the sets had only one or two eigenvalues greater than 1.0. In three of the sets that had more than two eigenvalues greater than 1, examination revealed that the difference between the second and third eigenvalue was substantial and constituted a breaking point in a curve relating eigenvalues to components. There was one set (classroom process variables as rated by students) in which the breaking point came after the third eigenvalue, but inspection of component structures for this set showed that including the third component would have picked up only two more variables out of a total of 12. Therefore, it was decided to retain only the first two principal components for that set. In sum, to keep the number of covariates as low as possible relative to the number of classes to be analyzed, and to retain interpretability of composites (covariates), eight sets of logically defined or related variables were reduced to either 1 or 2 principal components each.

Table III-7 gives the structure matrix for each set of principal components (PC's). It also shows the eigenvalues of 1.0 or more for each set, and the cumulative percentage of variance accounted for by those eigenvalues. It can be seen in Table III-7 that there are a total of 13 PC's which will be referred to hereafter in the following abbreviated form:

Input Principal Components (PC's)

1. Student Pretest Achievement: Ach
2. Student Pretest Attitude: Att 1, Att 2
3. Teacher Experience: T Demo
4. Teacher Pre Attitudes: T Psy 1, T Psy 2
5. Classroom Composition and Characteristics: Class 1, Class 2

Process/Climate Principal Components (PC's)

1. Process Variables as Measured by Student Ratings: S Proc 1, S Proc 2
2. Process Variables from Teacher Ratings: T Proc 1, T Proc 2
3. Climate Variables as Measured by Student Ratings: Climate

Table III-8 shows the correlations of PC's for the total group, and for MACOS and non-MACOS groups. It can be seen in Table III-8 that there are non-trivial correlations among some PC's. For example, there is a substantial correlation between the pretest achievement PC (Ach) and the first component of the attitude pretests (Att 1). It can also be seen that in a few cases, there are substantial differences in the intercorrelations of PC's between MACOS and non-MACOS groups. For example, for the MACOS group, the correlation of the first student process component (S Proc 1) with the teacher experience component

Table III-7

Factor Correlations, Eigenvalues and Cumulative Percentage of Variance of Principal Components Analyses

PC	Variables	Polarity of Variables	First Factor (PC) Correlations	Second Factor (PC) Correlations	Eigenvalues[1]	Cumulative % of Variance
Ach	Animals and People (AP) Pre	High = +	.90	-	2.50	83
	Interpretation of Data Test (IDT) Pre	"	.90	-		
	STEP Pre	"	.94	-		
Att 1,2	Social Studies Choices (SS Ch) Pre	High = +	.42	.32	2.05	29
	What Would You Think Part A (WWA) Pre	"	.70	-.37	1.19	46
	What Would You Think Part B (WWB) Pre	"	.79	.12	1.06	62
	CAPS-1, Ability, Pre	"	.29	.08		
	CAPS-2, Interest, Pre	"	.32	.71		
	CAPS 3, Tolerance of Ambiguity, Pre	"	.71	.17		
	CAPS-4, Creativity, Pre	"	-.24	.63		
Class 1,2	Average Age		.69	-.55	2.58	22
	SD of Age		-.18	-.37	2.05	39
	Years in present school		.43	-.45	1.34	50
	SD years in present school		.48	-.08	1.05	60
	%-females		-.15	.03	1.01	69
	% not ESL (English as a second language)		.35	.53		77
	% not eligible for free lunch program		.64	.55		
	% Previous MACOS students		.07	-.02		
	% white students		.74	.55		
	% 5th grade students		-.70	.47		
	Class size		-.04	.52		
	Class stability		.25	-.16		

1. Only values ≥ 1.0 are reported.

Table III-7 Continued

Factor Correlations, Eigenvalues and Cumulative Percentage of Variance of Principal Components Analyses

PC	Variables	Polarity of Variables	First Factor (PC) Correlations	Second Factor (PC) Correlations	Eigenvalues	Cumulative % of Variance
T Demo	Years teaching experience		.96		1.87	62
	Years in present school system		.95			
	Years experience with present SS program		.20			
T Psy 1, 2	Educational Scale VII Progressivism		-.36	-.51	4.74	43
	Educational Scale VII Traditionalism		.06	.87	1.73	59
	Teachers at Work (TAW)		-.14	.57		
	Category 1 (Citizenship Transmission Objectives)*	The higher the score, the less the objectives in the category were rated as important	.63	-.42		
	Category 2 (Cross-disciplinary, humanities, integrated concept objectives)*	"	.75	.12		
	Category 3 (Inquiry modes and processes objectives)*	"	.79	.20		
	Category 4 (Self-actualization objectives)*	"	.77	.16		
	Category 5 (social sciences objectives)*	"	.73	-.25		
	Category 6 (Values analysis objectives)*	"	.79	.16		
	Political relevance objectives*	"	.70	-.15		

*From Program Survey Form

Table III-7 Continued

Factor Correlations, Eigenvalues and Cumulative Percentage of Variance of Principal Components Analyses

PC	Variables	Polarity of Variables	First Factor (PC) Correlations	Second Factor (PC) Correlations	Eigenvalues	Cumulative % of Variance
T Psy 1, 2 cont'd	General social studies objectives	The higher the score, the less the objectives in the category were rated as important	.87	-.12		
S Proc 1,2	Teacher Talk*	High means little talk	-.12	.07	3.14	19.6
	Speed	High means want to go faster	-.13	-.24	2.89	37.7
					2.21	51.5
	Listen	High means much time just listening to teacher	.07	.22	1.27	59.5
					1.19	66.9
	Discussion	High means not much discussion	-.13	.71		
	Stress	High means little stress on grades	.64	-.49		
	Compare	High means little emphasis on comparing	.12	.63		
	Joking	High means little joking, informality	-.56	.07		
	Memory	High means little emphasis on remembering facts	.70	-.22		
	Translation	High means little emphasis on restating or saying in own words	.58	.01		
	Interpretation	High means little emphasis on interpreting what things mean	.56	.52		
	Application	High means little use elsewhere of what's learned in school	.32	.38		

2 My Social Studies Class Form, Part I and II.

Table III-7 Continued

Factor Correlations, Eigenvalues and Cumulative Percentage of Variance of Principal Components Analyses

PC	Variables	Polarity of Variables	First Factor (PC) Correlations	Second Factor (PC) Correlations	Eigenvalues	Cumulative % of Variance
S Proc 1, 2, cont'd	Analysis	High means little emphasis on completeness, giving good reasons, making sense	.62	.26		
	Synthesis	High means little emphasis on making new things, creating from what was learned	.48	.34		
	Evaluation	High means little emphasis on deciding on right or wrong, good or bad	.17	.60		
	ODI	High means little opportunity for discussion or involvement	.03	.58		
	Test/Grade Stress	High is little emphasis on grades or right answers	.67	-.54		
T Proc 1, 2	Affect	High means emphasis of curriculum on affective content	.55	-.42	3.52	32.0
	Memory	High means more emphasis on remembering	.01	.58	2.32	53.0
	Comprehension	High means more emphasis on understanding	.05	.55	1.19	68.9
	Application	High means more emphasis on using knowledge, problem solving	.61	-.34		

Table III-7 Continued

Factor Correlations, Eigenvalues and Cumulative Percentage of Variance of Principal Components Analyses

PC	Variables	Polarity of Variables	First Factor (PC) Correlations	Second Factor (PC) Correlations	Eigenvalues	Cumulative % of Variance
T Proc 1, 2 Cont'd	Analysis	High means more emphasis	.65	-.10		
	Synthesis	High means more emphasis	.65	-.41		
	Eval'n	High means more emphasis	.68	-.09		
	Indiv	High means more individual activities	.19	.77		
	Group	High means more group discussion activities	.80	.23		
	PM	High means more perceptual motor activities	.38	.69		
	Total Group	High means more group activities and projects, including discussions	.83	.26		
Climate	Satisfaction	Low means more satisfaction	-.94		2.45	81.8
	Apathy	High means less apathy	.93			
	Difficulty	High means work is rated as less difficult	.84			

Table III-8

Correlations Between Principal Components, PC's for Total Group, T (N=81), MACOS, M (N=45) and Non-MACOS, NM (N=36)[1]

PC	Group	Ach	Att 1	Att 2	Class 1	Class 2	T Demo	T Psy 1	T Psy 2	S Proc 1	S Proc 2	T Proc 1	T Proc 2
Att 1	T	56											
	M	45											
	NM	68											
Att 2	T	-01	05										
	M	-02	11										
	NM	02	01										
Class 1	T	73	48	-14									
	M	73	28	-13									
	NM	72	70	-15									
Class 2	T	03	07	15	-15								
	M	-08	-02	35	-32								
	NM	12	15	-03	02								
T Demo	T	-14	01	06	-12	01							
	M	-23	-12	05	-16	-20							
	NM	-02	17	06	-07	23							
T Psy 1	T	04	-09	11	05	-15	-20						
	M	08	-15	03	04	-05	-02						
	NM	01	-00	15	08	-25	-47						
T Psy 2	T	-03	-07	03	-10	-04	21	-02					
	M	02	01	-23	-08	00	14	-46					
	NM	-05	-10	22	-12	-05	30	26					
S Proc 1	T	41	26	-03	43	-08	-31	01	-26				
	M	44	27	08	48	-04	-51	01	-22				
	NM	37	25	-15	37	-12	-04	04	-29				
S Proc 2	T	-05	-01	-02	-08	-10	04	07	29	08			
	M	-11	05	-22	-13	07	02	-09	38	-12			
	NM	06	-04	18	-01	-27	06	13	08	23			
T Proc 1	T	-00	18	08	13	-01	10	-07	-21	-03	-43		
	M	-08	03	18	16	-16	07	08	-32	-01	-49		
	NM	10	38	-01	10	17	19	-20	02	-08	-25		
T Proc 2	T	02	02	09	01	08	17	-14	45	-21	18	06	
	M	23	11	09	09	27	-00	-30	36	-04	01	15	
	NM	-20	-02	00	-08	-06	42	-18	41	-42	15	19	
Climate	T	16	17	21	11	11	-06	-02	-39	04	-40	22	-33
	M	06	03	13	10	-34	-06	13	-34	29	-43	31	-20
	NM	25	26	41	12	45	-05	-02	-32	-17	-25	-04	-26

1. Decimal points are omitted. See Appendix E for similar table for largest sample, based only on student PC's.

(T Demo) is -.51; for the non-MACOS group the correlation is -.04. The same correlation for the total set of classes is -.31.

The following, based on Table III-7, are interpretations of each PC.

Input Principal Components (PC's)

Pretest Achievement of Students

Ach: generalized pre achievement level. The three achievement tests, STEP, IDT and AP, all load highly on this factor.

Pretest Attitudes of Students:

Att 1: tolerant of ambiguity (CAPS-3); rates unusual beliefs and behaviors more positively (WWA, WWB); tends to choose social studies over other courses (SS Ch).

Att 2: interest in complex problem solving (CAPS-2); thinks of self as having creative ideas (CAPS-4).

Classroom Characteristics

Class 1: older, white, non-poor classes (age, percent white, percent not eligible for free lunch).

Class 2: larger, younger, homogeneous classes (class size, percent not ESL, percent white, percent not eligible for free lunch, average age)

Teacher Experience

T Demo: more experienced teachers, although not necessarily with present program.

Teacher Attitudes

T Psy 1: generalists; teachers who tend to see many objectives as desirable, not certain categories as critical.

T Psy 2: conservatives; teachers who score higher on ES VII traditionalism and TAW, and lower on ES VII progressivism.

Process/Climate Principal Components (PC's)

Process as Rated by Students

S Proc 1: generalized informality; no apparent emphasis on particular types of activities or on grades, from students' point of view.

S Proc 2: generalized traditional class; stress on grades, relatively little discussion or comparing; more emphasis on remembering facts, in students' perceptions.

Process as Rated by Teachers

T Proc 1: emphasis on group activities and higher order cognitive processes.

T Proc 2: emphasis on knowing, remembering, and individual work.

Classroom Climate as Rated by Students

Climate: students like the teacher and class. Correlates positively, to a limited extent, with Att 2, T Proc 1 and 2; negatively with T Psy 2.

Plots of the distributions of each PC were made for the total group of classes for which there were complete data for all variables and separately for MACOS and non-MACOS groups. Histograms of each PC for the total group are shown in Appendix D. Most PC's show symmetric, normal-appearing distributions for total sample and each group. For the total group, and for MACOS and non-MACOS groups, there is some negative skewness for T Demo. In effect, in both groups, years of teaching experience tend to lump toward the lower end of the distributions. Class 1 shows some bi-modality in both groups, while Class 2 had three non-MACOS classes that were off to the low end of the distribution, producing negative skewness for that PC for non-MACOS and total group. There was some negative skewness in S Proc 2, and positive skewness for T Proc 2 for the non-MACOS distributions, with only slight effect on the overall distributions, as can be seen from the histogram in Appendix D. Normal plots and cumulative ogives, not shown, indicated essentially what can be seen in the histogram; viz., most of the PC's have reasonably symmetrical distributions that appear normal, and do not show outliers. The distributions of PC's were considered

reassuring for use in subsequent analyses, although it may be noted that normality is not a necessary assumption for tests of differences between groups of approximately equal size and equal variances. Table III-9 gives descriptive statistics for the PC's for MACOS, non-MACOS, and the total group.

2. Initial Comparability of Groups

A question of interest is whether the two groups were comparable initially. A multivariate analysis of variance was made to compare the MACOS and non-MACOS classes with respect to the input PC's.[8]

Two separate analyses were made. The first used only student-based PC's (derived from class means, not individual students), with resulting samples of 55 MACOS classes and 47 non-MACOS classes.[9] The PC's were pretest achievement and attitude, and classroom characteristics (Ach, Att 1, Att 2, Class 1, and Class 2). The second analysis included, in addition, 3 teacher PC's pertaining to experience and attitudes (T Demo, T Psy 1, and T Psy 2). This analysis had sample sizes of 47 MACOS and 41 non-MACOS classes.

Neither analysis produced a significant difference between the groups. For the comparison using only the student PC's, $F_{5,96} = .631$, $p \leq .68$, while for the comparison using both student and teacher PC's, $F_{8,79} = .287$, $p \leq .98$. It appeared that the groups could be considered as not different at the outset with

8. The program used was the Multivariate Analysis of Variance Program of the Biometric Laboratory, University of Miami.

9. The analysis described here did not include 4 classes from a district that did not participate in the follow-ups, plus an additional 2 classes for which there was missing STEP data, either pre or post. The classes in the district that were not included here were included in other analyses to be described later.

Table III-9
Means, Standard Deviations (SD), Skewness (Sk)
Kurtosis (K), and Range for PC's, by Group 1/

PC	MACOS Mean	SD	Sk	K	Range	Non-MACOS Mean	SD	Sk	K	Range
1. Ach 1	.12	1.51	.23	-.57	-3.4-3.0	.01	1.52	.23	-.88	-2.9-3.0
2. Att 1	.12	1.28	-.08	-.96	-2.3-2.7	-.06	1.49	-.20	-.42	-3.0-2.6
3. Att 2	-.04	1.07	.24	-1.26	-1.7-2.0	.20	1.03	.06	.55	-2.1-3.1
4. Class 1	.05	1.59	-.59	-.56	-3.7-2.4	-.00	1.58	-.40	-.22	-4.0-2.8
5. Class 2	.26	1.14	.35	.26	-2.3-3.5	.20	1.51	2.72	10.52	-6.9-1.8
6. T Demo	-.02	1.31	1.53	2.79	-1.5-4.7	.04	1.20	.40	1.02	-1.5-2.6
7. T Psy 1	-.58	2.03	.37	.17	-5.3-4.4	.31	2.02	-.21	-.88	-3.6-3.7
8. T Psy 2	-3.1	1.12	.34	-.67	-2.3-2.4	.45	1.55	-.47	-.13	-3.6-3.5
9. S Proc 1	.14	1.75	.03	-.66	-3.8-3.5	-.01	1.84	-.12	.18	-4.2-4.3
10. S Proc 2	-.41	1.66	-.84	1.27	-6.1-2.2	.59	1.58	.76	-.43	-1.2-4.5
11. T Proc 1	.49	1.88	-.16	-1.18	-3.2-3.5	-.28	1.45	-.09	-.74	-2.2-2.2
12. T Proc 2	-.60	1.38	-.10	-.46	-3.8-2.2	.58	1.36	-.04	-.71	-2.6-2.8
13. Climate	.44	1.38	.39	-.36	-1.9-3.9	-.64	1.69	.07	-.96	-4.0-2.8

TOTAL GROUP

PC	Mean	SD	Sk	K	Range
1. Ach 1	.00	1.52	.00	-.69	-3.38-3.02
2. Att 1	-.02	1.37	-.10	-.73	-2.98-2.72
3. Att 2	.06	1.06	.18	-.54	-2.06-3.09
4. Class 1	-.04	1.57	-.43	-.47	-4.00-2.84
5. Class 2	.18	1.31	-1.50	8.16	-6.87-3.46
6. T Demo	-.01	1.24	1.14	1.58	-1.54-4.72
7. T Psy 1	-.07	2.14	.21	-.28	-5.25-5.51
8. T Psy 2	.04	1.34	.03	-.29	-3.56-3.45
9. S Proc 1	-.01	1.79	.02	-.42	-4.23-4.28
10. S Proc 2	.03	1.67	-.20	1.41	-6.09-4.45
11. T Proc 1	.04	1.78	.03	-.90	-3.71-3.51
12. T Proc 2	-.04	1.49	-.05	-.44	-3.77-2.82
13. Climate	-.02	1.58	-.04	-.29	-4.03-3.86

1. Sample sizes are: MACOS, 45; non-MACOS, 36. These are samples for which there were complete teacher and student data, including posttest and follow-up measures.

with respect to the array of these composite variables (PC's) of student pre-test, classroom and teacher characteristics.

Normally one would not make a further examination of the significance of individual variables if the overall multivariate test is not significant. That practice, to protect the Type I error rate, has typically been followed in this study. What follows is a departure from that principle, undertaken because of continuing interest in the nature of the two non-experimentally formed groups of classes. Table III-10 gives the means and standard deviations of the student and classroom PC's for each group, the difference between means, and the F-statistic and p-value of the univariate test of difference for each PC.[10] It may be seen in Table III-10 that none of the student pretest and classroom characteristics PC's reaches significance on a univariate basis.

Table III-11 gives the same statistics for the univariate analyses of differences that include the teachers PC's as well as the student and class ones. As before, none of the student or class PC's reaches significance. However, the second teacher attitude principal component (T Psy 2) was significant at the .05 level ($p \leq .013$), and the first teacher attitude PC was nearly so ($\leq .058$). The original variables that correlated strongly with the T Psy 2 component were the Educational Scale VII Traditionalism score and the Teachers at Work (TAW)

10. It will be recalled that the pretest means and standard deviations for the individual variables that were combined in the principal component composites were presented earlier in Tables III-4, 5, and 6. Tests of significance of differences of the individual variables were not given for the reason that it was believed to be methodologically more sound to make an over-all multivariate test with as much power as possible before attempting to interpret a test of each of a large number of variables separately.

Table III-10
Means, Standard Deviations, N's Differences Between Means, and P-Values of Univariate F-tests (df 1, 100) Between MACOS and Non-MACOS Classes Adjusted for 5 Pretest and Classroom PC's

Principal Component(PC)	MACOS Mean	MACOS SD	MACOS N	Non-MACOS Mean	Non-MACOS SD	Non-MACOS N	Difference	P
Pre-Achievement								
Ach	.07	1.60	55	-.11	1.60	47	.18	.589
Pre-Attitude								
Att 1	.11	1.26	55	-.17	1.63	47	.28	.324
Att 2	-.01	1.06	55	.10	1.13	47	-.11	.632
Classroom Characteristics								
Class 1	-.01	1.59	55	-.00	1.60	47	-.01	.976
Class 2	.17	1.21	55	-.17	1.68	47	.34	.237

Table III-11

Means, Standard Deviations, N's Differences Between Means, and P-Values of Univariate F-tests (df 1,86) of Differences Between Means of MACOS and Non-MACOS for 5 Student Based and 3 Teacher-Based Background PC's

	MACOS			Non-MACOS			Difference	P-
	Mean	SD	N	Mean	SD	N		
Pre-Achievement								
Ach	.10	1.48	47	-.11	1.53	41	.21	.503
Pre-Attitude								
Att 1	.09	1.26	46	-.23	1.49	41	.34	.284
Att 2	-.04	1.05	46	.14	1.06	41	-.18	.439
Classroom Characteristics								
Class 1	-.01	1.58	47	-.04	1.53	41	.03	.928
Class 2	.26	1.15	47	.06	1.49	41	.20	.465
Teacher Characteristics								
T Demo	-.08	1.31	47	.15	1.42	41	-.23	.440
T Psy 1	-.47	2.18	47	.39	2.02	41	-.86	.056
T Psy 2	-.32	1.11	47	.39	1.49	41	-.71	.013

score. On both measures, non-MACOS teachers on the average scored significantly higher (at the .05 level) than MACOS teachers; i. e., in the direction of traditionalism. The variables that correlated particularly well with the first teacher attitude component (T Psy 1) were summated ratings of the importance of different categories of objectives. There were three categories for which there were significant differences between MACOS and non-MACOS teachers, on the average. Those were: Category 2 (objectives intended to be related to a cross disciplinary, humanities, integrated concept); Category 3 (objectives intended to be related to inquiry modes and processes); and Category 4 (objectives intended to be related to self-actualization as a goal of instruction). MACOS teachers, on the average, rated those groups of objectives more positively (as being important) than non-MACOS teachers. Thus there is indication that the two groups of teachers were different with respect to certain attitude variables. There is no indication that the classes of students in the two groups were significantly different with respect to certain preachievement and pre-attitude variables. There is no indication that the classes of students in the two groups were significantly different with respect to classroom demographic characteristics, using PC's as measures.

Subsequent analyses of covariance and regression analyses will statistically remove initial differences, although they may remove some treatment effect as well. In light of possible initial differences between the groups of teachers, homogeneity of regression tests of outcomes on inputs become particularly important; such tests and their results are described

below.[11] Furthermore, caution is needed in interpretations of results and in formulating hypotheses about possible causes of effects. With respect to the latter, it is restated here that material obtained from interviews with students and teachers and presented in Section V is considered to be an important source of information bearing on possible explanatory hypotheses.

3. Major Outcomes

a. Dependent Variables

Dependent variables are organized according to whether they were first year outcomes (posttest), Follow-up 1 outcomes (FU-1) or Follow-up 2 outcomes (FU-2). Posttest measures were:[12]

AP: a MACOS course content specific test

IDT: the Interpretation of Data Test (interpretation and use of ethnographic data)

STEP: STEP Social Studies Test (social studies skills and knowledge)

SS Ch: social studies choices; the number of times social studies was picked in preference to 5 other subjects

WWA: What Would You Think, Part A, attitudes toward unusual behavior, customs or beliefs

WWB: What Would You Think, Part B, attitudes towards people who might do such things or have such customs or beliefs

CAPS 1: ratings of ability of self as problem solver

CAPS 2: ratings of interest in solving problems

CAPS 3: ratings of tolerance of ambiguity in problems

11. To anticipate, the results were that the hypothesis of homogeneity of regression of outcome variables on teacher composite variables were accepted at the .05 level.

12. All outcome variables are based on class means.

CAPS 4: ratings of creativity of self as thinker

Follow-up 1 measures were:

Skills: the average of ratings of 8 scales which for each item, range from 1-learned how to do this last year and found it a great advantage this year, to 5-didn't learn how to do it and wish I had because it would help in social studies this year.[13]

Know: the average of ratings of 7 scales pertaining to knowledge of topics that may have been studied last year; same scale values as for skills (i. e., values ranged from 1-learned this last year and found it a great advantage this year, to 5-didn't learn this and wish I had because it would help in social studies this year)[14]

Interest: responses on a 5 point scale to the question of how interesting you find social studies this year compared to last year; 1 is a lot more, 5 is a lot less.[15]

Follow-up 2 measures were:

SS Ch F: 6 pair comparisons of social studies, arithmetic, science English, scored for number of times social studies was picked (0 to 3).[16]

---

13. My Social Studies Class This Year and Last (MSSCTYL), Part III, items 1-8. The items were: 1) how to make or use maps; 2) how to make or use graphs; 3) how to find information in the library; 4) how to write reports; 5) how to tell the difference between facts and opinions; 6) how to support your ideas or opinions with evidence or facts; 7) how to work with other students in small groups; 8) how to look at all sides of a question before deciding what you think. See Appendix. A.

14. Ibid., items 9-15. The items were: 1) how people and their environment affect each other; 2) the history and/or customs of our country; 3) the history and/or customs of other countries; 4) different beliefs people have; 5) how different animals behave and why they behave the way they do; 6) similarities and differences in ways animals and people behave; 7) learning and understanding more about myself and other people. See Appendix A.

15. Ibid., Part II, item 3i.

16. This was a modification of SS Ch to take into account the fact that 7th grade programs do not normally include reading and spelling.

WWAF: same as WWA

WWBF: same as WWB

WWAPF: two WWA-type items concerning attitudes toward unusual behavior of a hypothetical peer

WWBPF: attitudes toward the person that might behave in this unusual manner

AP1-4F: the man and other animals part of the MACOS specific questionnaire (AP)

SS: an absolute rating on a 4 point scale of how much social studies is liked this year; 1 - dislike very much; 4 - like very much;

b. Multivariate Comparisons and Analyses

With 20 outcome or dependent variables, one might expect by chance alone to find significant differences at the .05 level in one or two cases if each variable was tested individually, assuming independence of variables. As a means of guarding against that possibility, a multivariate analysis of covariance was made.[17] The independent variable was Group (MACOS, non-MACOS); the covariates were the input, process and climate PC's. All 20 outcome variables were used simultaneously as dependent variables.

Two analyses were made. The first used only the 8 student based PC's as covariates (pretest achievement, attitude, classroom characteristics, student based ratings of processes and classroom climate). This provided a sample size with complete data of 97 (54 MACOS and 43 non-MACOS classes).[18]

17. The program used was again the Multivariate Analysis of Variance Program of the Biometric Laboratory, University of Miami.

18. By including follow-up measures in the analyses of classes, 4 classes from 1 district that did not participate in the follow-ups were lost. The other 7 losses from the 108 classes were also mostly from follow-ups except for 2 cases for which there was missing STEP data, either pre or ppst.

The second included, in addition, the 5 teacher based PC's (teacher experience, attitudes, and ratings of processes). The sample sizes for the second analysis were 45 MACOS and 36 non-MACOS classes, or a total of 81 classes. The rationale underlying these two analyses was as follows. The analysis using only the student-based PC's gives the maximum sample sizes having complete data sets and thus involves the samples of most general interest with respect to outcomes. It also minimizes the number of covariates. It adjusts outcome means statistically for differences between groups that might have been associated with pretest characteristics of classes, average perceptions of class activities or processes, and average perceptions of classroom climate. Any or all of these may be related to curriculum, and to the effects of curriculum. Thus, the analysis essentially addresses the question of whether, by equating statistically for differences in potential operational and mediating characteristics of courses as perceived by the classes of students, as well as in initial characteristics, there is still reason to believe there are differences between the two groups with respect to the total set of outcomes under consideration. That is, the analysis tests the null hypothesis that the MACOS and non-MACOS groups of classes were not different with respect to an array of outcome variables, when those outcomes were statistically adjusted for input and intervening characteristics. The reasons for doing the second analysis with the teacher PC's included were to examine the possiblity that differences in teacher characteristics may have influenced observed multivariate differences, if any, between groups, and to establish characteristics of the samples for which subsequent investi-

gations of relationships of variables, including teacher variables, were of interest. It should be noted that in both cases, the analyses address the question of differences in outcomes associated with groups, not of possible relationships among variables within or between groups that may help explain differences. The latter will be addressed in subsequent analyses.

The results for both analyses showed a significant difference between groups for the multivariate F-test, ($p \leq .001$ in both cases).[19] Table III-12 gives the p-values of the univariate tests of each outcome variable adjusted for the 8 student PC's and, in addition, for the 5 teacher PC's. Adjusted and unadjusted means, standard deviations and differences between the MACOS and non-MACOS groups will be presented following a statement about homogeneity of regression tests made of each variable. At this point, the results support the hypothesis that the two groups differ when all outcome variables taken together are adjusted statistically for student (and student and teacher) background, classroom process or activities, and classroom climate variables. The univariate tests of adjusted outcome variables point to at least six outcome variables in posttest and in each follow-up in which there is indication of interpretable differences between groups.

1.) Homogeneity of Regressions Tests

Homogeneity of regression tests of each outcome variable were made for each case (the large group and the reduced group). The available program

19. For the analysis using the 8 student-based covariates, $F_{20,68}=3.681$, $p \leq .001$. For the analysis using all 13 covariates, $F_{20,47}= 3.215$, $p \leq .001$.

Table III-12
P-Values of Univariate F-tests of Differences Between MACOS and Non-MACOS Classes on 20 Posttest and Follow-up Outcome Variables Adjusted for 8 Student PC's, and for 13 Student and Teacher PC's

Period	Outcome Variable	P-Values for F-tests (df = 1, 87) Using 8 Student PC's as Covariates (MACOS N = 54; Non-MACOS N = 43)	P-Values for F-tests (df = 1, 66) Using 13 Student and Teacher PC's as Covariates (MACOS N = 45; Non-MACOS N = 36)
Posttest	AP	.001***	.001***
	STEP	.277	.528
	IDT	.346	.605
	SS Ch	.497	.916
	WWA	.079	.018*
	WWB	.595	.412
	CAPS-1	.064	.063
	CAPS-2	.795	.411
	CAPS-3	.619	.895
	CAPS-4	.103	.446
Follow-up 1	Skills	.046*	.115
	Know	.011*	.009**
	Interest	.013*	.056
Follow-up 2	AP1-4F	.002**	.029*
	SS Ch F	.293	.600
	SS	.459	.581
	WWAF	.296	.978
	WWBF	.030*	.318
	WWAPF	.423	.235
	WWBPF	.504	.940

*p ≤ .05
**p ≤ .01
***p ≤ .001

Note: these designations are made as a visual aid to the reader, since subsequent reference will be made especially to outcome variables with p-values of .05 or less.

could not handle 13 covariates simultaneously.[20] Therefore, the homogeneity tests were made separately for the 8 student-based PC's and the 5 teacher-based PC's. These analyses in both cases involved again listwise deletion, but for the specific variables involved in a given analysis. That is, each class was included for a variable if there were data for the class on the particular outcome and particular set of PC variables rather than dropping the whole class if there were any data missing on any other PC or outcome variable. Thus, the sample sizes ranged from 97 to 102 for different analyses. For the reduced group, the homogeneity tests were made only for the 81 cases for which there were complete data for teachers and students. The reason for that was further analyses including teacher variables were to be made of that set of classes, and thus an analysis of homogeneity of regression specifically for that set was wanted.

For the total group, using the 8 student-based PC's,[21] there were two outcome variables for which the F-test for homogeneity of regression was significat at the .05 level: WWA (posttest) and WWAPF (Follow-up 2).[22] In order to assess which covariate or covariates were associated with non-homogeneity of regression, the homogeneity tests were redone for the two outcome variables for each covariate. When WWA was regressed on each PC individually, all regression slopes were homogeneous except for Climate. For WWAPF, there were interactions for ACH, Class 1, S Proc 1, S Proc 2, Climate

20. One Dimensional Analysis of Covariance with Homogeneity of Regression Test, prepared by Dr. C. Mitchell Dayton, College of Education, University of Maryland.

21. Ach, Att 1, Att 2, Class 1, Class 2, S Proc 1, S Proc 2, Climate.

22. The F-test for among-slope differences for WWA was $F_{8,84} = 2.121$ $p \leq .042$ and for WWAPF, $F_{8,80} = 5.371$, $p \leq .001$. Group sizes for posttest variables were MACOS = 55; non-MACOS = 44.

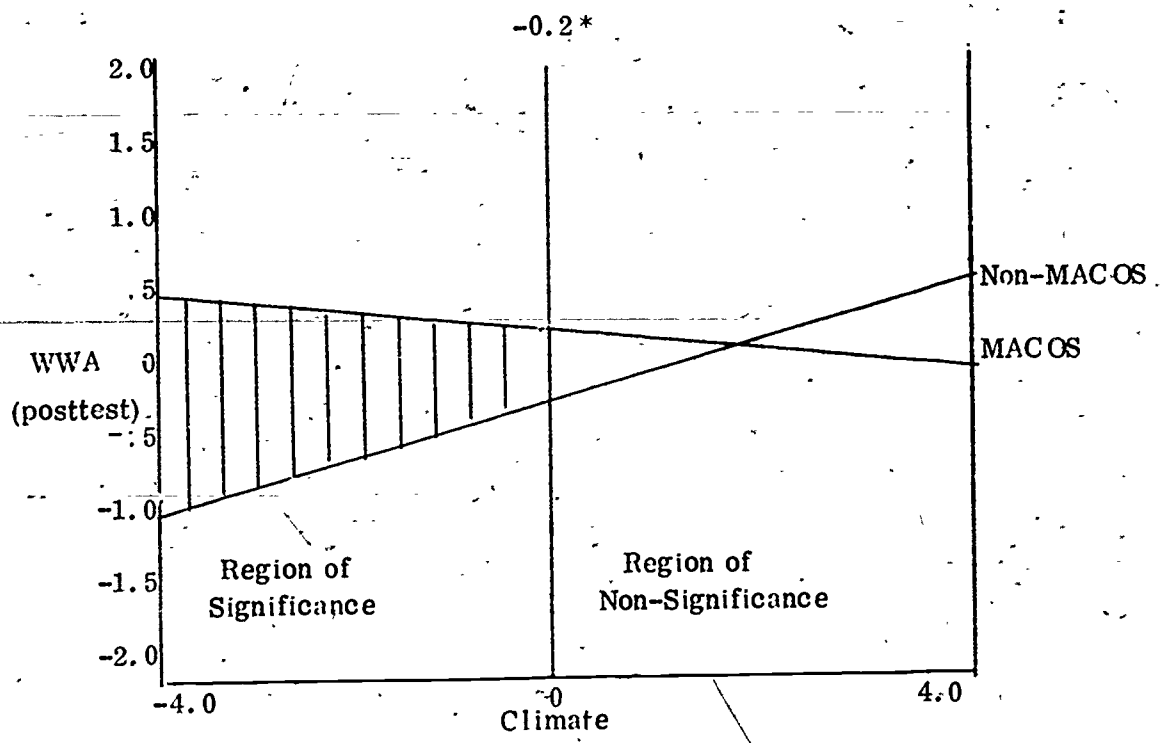
Results of homogeneity of regression tests for the 8 student-based PC's for the reduced group of classes were that there was non-homogeneity of regression for WWA, WWB, WWAPF and WWBPF.[23] Again, analyses for individual PC's indicated that for all except WWAPF, Climate was the only variable for which there was heterogeneity. When the 5 teacher PC's (T Demo, T Psy 1, T Psy 2, T Proc 1, T Proc 2) were used to test homogeneity of regression for each outcome variable, all regressions were homogeneous. These tests, of course, could only be done with the reduced group since that was the group with complete teacher data.

The regression lines of WWA and WWB regressed on Climate for the two groups are given in Figures III-1, 2, 3 and 4. These plots show the regions of significance and non-significance of difference between the slopes as determined by the Johnson and Neyman technique, based on 95% confidence limits around the difference between regression lines.[24] Each figure also includes for each group, the unadjusted and adjusted means and standard deviations of the outcome variables, and the regression intercept, slope, and standard error of the slope (SE_B). Figure III-1 shows the regression lines of WWA for MACOS and non-MACOS classes for the total sample of classes (N = 108). Figure III-2 shows the lines for the same variable for the reduced groups of classes (N = 81). Figures III-3 and 4 show the regression lines for WWB (posttest) and WWBF (Follow-up 2) for the reduced group of MACOS and non-MACOS classes, since only in those samples were the

23. F-test p-values with 8 and 63 degrees of freedom were: WWA, $p \leq .025$; WWB, $p \leq .024$; WWBF, $p < .048$; WWAPF, $p = .000$; WWBPF, $\leq .043$.

24. Johnson, T. O., and Jackson, R. W. B., Modern Statistical Methods. Chicago: Rand McNally, 1959.

Figure III-1
Neyman-Johnson Regions of Significance Between MACOS and Non-MACOS Regression Lines for WWA Regressed on the Climate PC (total sample)

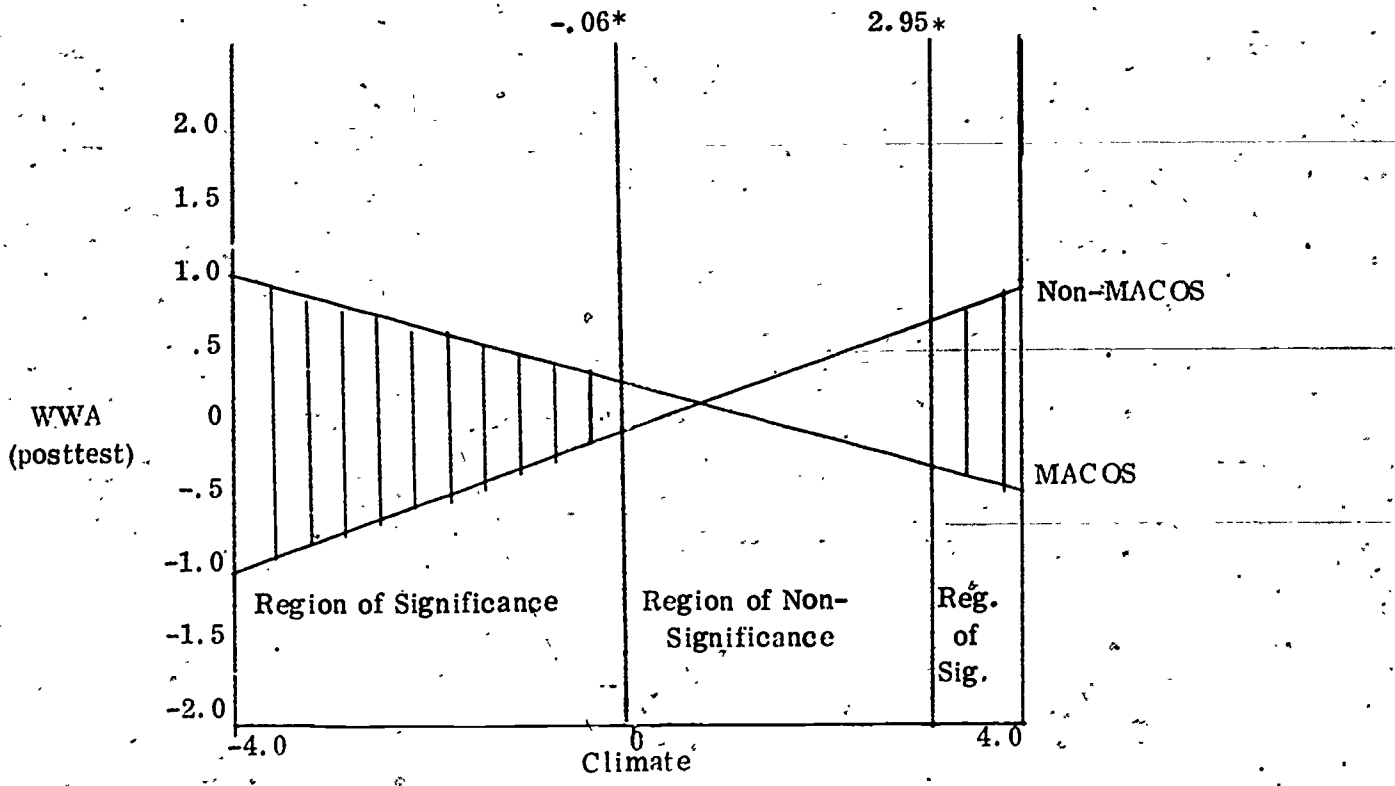


	Unadj Mean	SD	Adj Mean	Adj SD	N	Intercept	Slope	SE_B
MACOS	.21	.93	.16	.92	57	.24	-.06	.09
Non-MACOS	-.26	1.03	-.20	1.02	51	-.14	.24	.08

*Boundary of region of significance.

Figure III-2

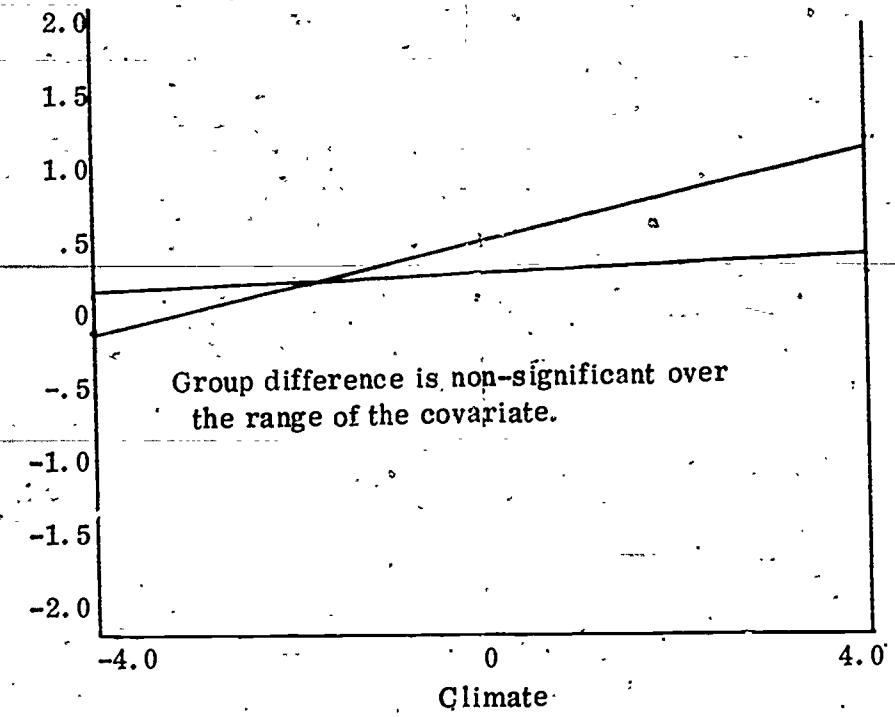
Neyman-Johnson Regions of Significance Between MACOS and Non-MACOS Regression Lines for WWA Regressed on the Climate PC (reduced sample)



	Unadj Mean	SD	Adj Mean	Adj SD	N	Intercept	Slope	SE_B
MACOS	.19	.89	.16	.89	45	.27	-.19	.09
Non-MACOS	-.29	1.04	-.26	1.04	36	-.13	.25	.10

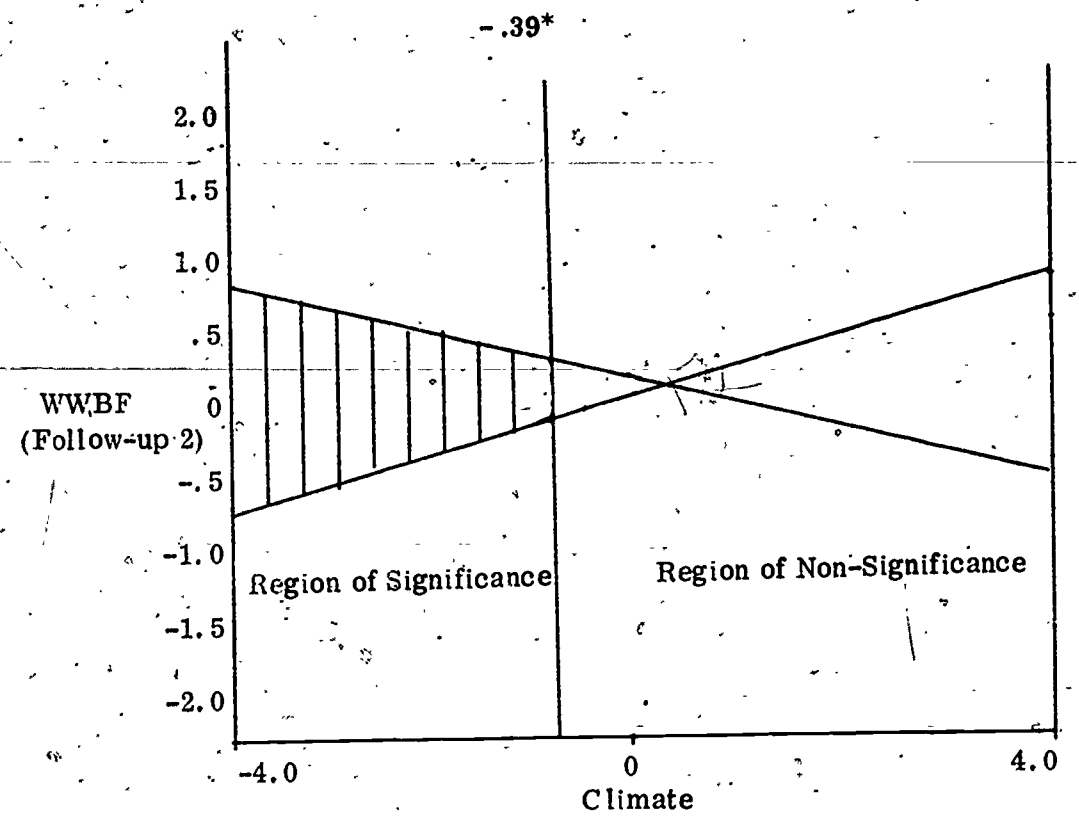
*Boundary of regions of significance.

Figure III-3
Neyman-Johnson Regions of Significance Between
MACOS and Non-MACOS Regression Lines for WWB
Regressed on the Climate PC
(reduced sample)



	Unadj Mean	SD	Adj. Mean	Adj SD	N	Intercept	Slope	SE_B
MACOS	.22	.77	.15	.74	45	.21	.03	.08
Non-MACOS	-.07	1.06	.03	1.02	36	.11	.27	.10

Figure III-4
Neyman-Johnson Regions of Significance Between
MACOS and Non-MACOS Regression Lines for WWBF
Regressed on the Climate PC
(reduced sample)



	Unadj Mean	SD	Adj Mean	Adj SD	N	Intercept	Slope	SE_B
MACOS	.20	.85	.19	.85	45	.26	-.13	.08
Non-MACOS	-.15	.94	-.14	.94	36	-.05	.15	.09

*Boundary of region of significance.

tests of heterogeneity significant with respect to Climate.[25]

Figures III-1 and 2 show for WWA that both for the total sample of classes (Figure III-1) and the reduced sample (Figure III-2), there is an interaction with respect to classes that rated classroom climate as poor (the negative end of the Climate PC scale). MACOS classes in that range rated unusual customs or beliefs more favorably on posttest, even if they had rated classroom climate lower in midtest 2 than non-MACOS classes. There is indication of the opposite being the case at the upper end of the Climate range for WWA in the reduced sample (Figure III-2). That may be the result of a change in sample characteristics owing to loss of classes in the reduced sample, although except for the MACOS regression slopes, the other parameters of WWA for the two groups are essentially the same in the reduced sample as in the total sample. Generally, the slope for the MACOS classes is flat or negative. The slope is not significant for the full sample (the 95% confidence limits include zero); it is just significant in the reduced sample. The slope for the non-MACOS classes is positive in all cases, and significant except for WWB for the reduced group. The implication is that Climate bore little or no relationship to how MACOS classes rated WWA or B items, while there did seem to be some positive relationship between ratings of classroom climate variables and ratings of customs and of people who hold them in the non-MACOS classes.

---

25. The apparent slopes of the regression lines shown in the figures are exaggerated by the scaling of the Y-axis relative to the X-axis. Also, boundaries of regions of significance are not drawn to exact values on the x-axis.

Figure III-3 presents an interesting situation. Climate was the one variable in the set of variables that contributed most to the significance of heterogeneity for the group slopes. However, that variable per se was not significant at the .05 level (the p-value of the F statistic was .075). The plot of regions of significance indicates that there was not a significant difference of regression slopes of WWB on Climate over the range of the climate variable. Therefore, one may conclude that while collectively the set of PC's produced a heterogeneity of the regression hyperplanes, and while Climate was the likely candidate for accounting for the non-homogeneity, it per se contributed more than any of the other covariates, but not significantly, to the overall heterogeneity.

The overall conclusion is that there is reason to consider interactions of Climate and WWA (posttest) responses between groups. For WWB posttest and Follow-up 2 in the reduced sample, classroom Climate may have different effects in the two main groups (MACOS, non-MACOS). Figure III-4, for WWBF, again suggests a different effect of Climate between the two groups on ratings of people who might have unusual beliefs or customs. However, Climate in the non-MACOS groups is associated with more negative ratings. Both for WWB and WWBF, the MACOS regression slopes are not significant while the non-MACOS ones are. With respect to WWAPF, there appears to be a complex interaction based on several factors. Johnson and Neyman confidence intervals and regions of significance were computed for each PC for which there was heterogeneity of regression, although they are not presented here. The position taken is that interpretations of differences between groups for this variable is complex at best, and that

little emphasis or weight should be put on those two measures. They are included in other analyses, but not interpreted. The consequences of heterogeneity of regression for WWA and WWB are more important within the overall context of the study.

The overall importance of the analyses of homogeneity of regressions of dependent variables on the various covariates is that, with the exception of the attitudes towards customs and people measures, there do not appear to be interactions between groups. Thus, analysis of covariance is supported as an analytic strategy, as is the use of fixed-order, stagewise regression.

As a final step in examining the nature of distributions of adjusted outcomes, residuals (observed minus expected values) of each outcome were computed using the 8 student-based PC's as predictors. Scatter-plots of standardized residuals plotted against standardized predicted dependent variables for the total groups were examined. They are shown in Appendix C. Examination of the scatter plots suggest that there were no unusual patterns in the swarms of the different variables, nor were there more extreme values than would be expected by chance. The same plots for MACOS and non-MACOS classes were made, with similar results. It was believed that distributions of residuals would not seriously bias further analyses.

Several conclusions were drawn from all these analyses. There were differences in outcomes between the two groups of classes (MACOS, non-MACOS) when they were adjusted for student, or student and teacher, input, process and climate variables, and when all outcome variables were analyzed simultaneously.

The assumption that regressions of outcome variables on the PC's for the two groups were homogeneous was supported for the majority of variables in the large and reduced samples. Where the assumption of homogeneity of regression could not be sustained for WWA (and also for WWA, WWBF, and WWBPF in the reduced sample), it was found that the one covariate associated with non-homogeneity between groups was climate. The variable WWAPF showed interactions between groups on a number of predictors.

2) Analyses of Differences

On what outcome variables were there differences between MACOS and non-MACOS classes? Table III-13 and III-14 summarize the unadjusted and adjusted means, standard deviations, differences (and p-value of the F-statistic) for each outcome variable, for the largest possible group of classes with all variables for a given outcome. These tables were obtained from the analyses of covariance described above. Table III-15 and III-16 give the same information for the reduced group. In this case, the analyses were held just to the 81 classes with complete sets of data in order to provide a comparable summary specifically based on the sub-sample including all teacher variables on which subsequent analyses were planned.

Table III-13 shows that when no adjustment in standardized outcome means for the largest group was made, there were outcome differences in:

- AP, the MACOS course content specific test, with MACOS classes outscoring non-MACOS classes on posttest, on the average, and also AP1-4F, the Man and Animals part of AP, in Follow-up 2, a year later;

Table III-13

Unadjusted Outcome Means, Standard Deviations, Differences and P-Values of Outcome Variables (z-score form)[1]

Period	Outcome Variable	MACOS Mean	MACOS SD	MACOS N	Non-MACOS Mean	Non-MACOS SD	Non-MACOS N	Diff[2]	P
1st Year (Posttest)	AP	.38	.96	55	-.46	.82	47	.84***	.000
	STEP	.01	1.03	55	-.05	.94	47	.06	.758
	IDT	.10	.99	55	-.15	.98	47	.25	.197
	SS Ch	.12	.89	55	-.20	.96	47	.32	.079
	WWA	.18	.92	55	-.24	1.06	47	.42*	.036
	WWB	.17	.90	55	-.18	1.05	47	.35	.067
	CAPS-1	-.15	1.04	55	.21	.96	47	-.36	.074
	CAPS-2	.12	1.13	55	-.06	.82	47	.18	.376
	CAPS-3	.07	1.06	55	-.16	.95	47	.24	.242
	CAPS-4	.04	.91	55	-.05	1.10	47	.09	.643
Follow-up 1	Skills	.15	1.08	54	-.15	.85	44	.30	.138
	Know	-.29	.90	54	.30	1.07	44	-.59**	.004
	Interest	.29	.90	54	-.40	.92	44	.69***	.000
Follow-up 2	AP1-4 F	.22	.91	54	-.28	1.02	44	.50*	.012
	SS Ch F	-.07	.99	54	.09	1.03	43	-.16	.456
	SS	-.04	.91	54	-.03	1.15	44	-.01	.958
	WWAF	.08	.95	54	-.11	1.08	44	.19	.351
	WWBF	.14	.85	54	-.23	1.13	44	.37	.071
	WWAPF	.06	.94	54	-.09	1.14	44	.15	.480
	WWBPF	.03	.93	54	-.11	1.14	44	.14	.499

*p ≤ .05
**p ≤ .01
***p ≤ .001

Note: Even though all p-values are given, differences with p-values of .05 or less are designated with asterisks as a visual aid to readers in identifying variables of particular interest.

1. PC's used were student based: Ach, Att 1, Att 2, Class 1, Class 2, S Proc 1, S Proc 2, and Climate. Outcome variable z scores are based on total group means for which there were outcome variables (N's = 102-108)

2. The test of the difference is the F-test.

Table III-14
Means, Standard Deviations, Differences and P Values of Outcome Variables (z score form) Adjusted for 8 Input, Process/Climate PC's[1]

Period	Outcome Variable	MACOS Mean	MACOS SD	MACOS N	Non-MACOS Mean	Non-MACOS SD	Non-MACOS N	Diff[2]	P
1st Year (Posttest)	AP	.36	.67	55	-.43	.57	47	.79***	.000
	STEP	-.06	.48	55	.04	.44	47	-.10	.310
	IDT	.06	.62	55	-.11	.61	47	.17	.244
	SS Ch	.05	.78	55	-.11	.84	47	.16	.367
	WWA	.11	.75	55	-.15	.86	47	.26	.161
	WWB	.07	.78	55	-.06	.91	47	.13	.509
	CAPS-1	-.19	.94	55	.25	.87	47	-.44*	.035
	CAPS-2	.07	.97	55	.00	.71	47	.07	.720
	CAPS-3	-.05	.77	55	-.01	.69	47	-.04	.839
	CAPS-4	.09	.77	55	-.11	.93	47	.20	.295
Follow-up 1	Skills	.21	1.02	54	-.23	.80	44	.43*	.046
	Know	-.26	.81	54	.27	.96	44	-.53*	.011
	Interest	.22	.85	54	-.31	.87	44	.53**	.010
Follow-up 2	AP1-4F	.21	.64	54	-.27	.71	44	.48**	.003
	SS Ch F	-.10	.94	54	.13	.97	43	-.23	.293
	SS	-.11	.79	54	.07	1.00	44	-.18	.399
	WWAF	.07	.83	54	-.10	.94	44	.17	.434
	WWBF	.16	.72	54	-.25	.96	44	.41*	.041
	WWAPF	.06	.91	54	-.10	1.10	44	.16	.518
	WWBPF	.02	.84	54	-.09	1.04	44	.11	.602

*p ≤ .05
**p ≤ .01
***p ≤ .001

Note: Even though all p-values are given, differences with p-values of .05 or less are designated with asterisks as a visual aid to readers in identifying variables of particular interest.

1. PC's used were the student-based input, process and climate ones: Ach, Att 1, Att 2, Class 1, Class 2, S Proc 1, S Proc 2 and Climate.

2. The test of differences is the F-test

Table III-15

Unadjusted Means, Standard Deviations, Differences and P Values of Outcome Variables (in z-score form) for the Reduced Group[1]

Period	Outcome Variable	MACOS Mean	MACOS SD	MACOS N	Non-MACOS Mean	Non-MACOS SD	Non-MACOS N	Diff[2]	P
1st Year (Posttest)	AP	.40	.89	45	-.46	.82	36	.86***	.000
	STEP	.09	.89	45	.03	.89	36	.06	.767
	IDT	.14	.94	45	-.01	.95	36	.15	.491
	SS Ch	.05	.84	45	-.18	.99	36	.23	.251
	WWA	.19	.89	45	-.29	1.04	36	.48*	.029
	WWB	.22	.77	45	-.07	1.06	36	.29	.155
	CAPS-1	-.20	.98	45	.09	.82	36	-.29	.167
	CAPS-2	.17	1.20	45	-.09	.86	36	.26	.283
	CAPS-3	.13	1.14	45	-.09	.79	36	.22	.332
	CAPS-4	-.01	.85	45	-.08	.87	36	-.07	.705
Follow-up 1	Skills	.10	1.11	45	-.12	.86	36	.22	.339
	Know	-.31	.90	45	.34	1.13	36	-.65**	.005
	Interest	.34	.88	45	-.39	.93	36	.73***	.001
Follow-up 2	AP1-4F	.21	.95	45	-.25	1.01	36	.46*	.038
	SS Ch F	-.08	1.00	45	.04	.97	36	-.12	.592
	SS	-.04	.85	45	-.12	1.03	36	-.08	.688
	WWAF	.12	.94	45	-.10	.96	36	.22	.294
	WWBF	.20	.85	45	-.15	.94	36	.35	.083
	WWAPF	.17	.89	45	-.12	.86	36	.29	.141
	WWBPF	.05	.87	45	-.08	1.07	36	.13	.546

*p < .05
**p < .01
***p < .001

Note: Even though all p-values are given, differences with p-values of .05 or less are designated with asterisks as a visual aid to readers in identifying variables of particular interest.

1. PC's used were: Ach, Att 1, Att 2, Class, 1, Class 2, S Proc 1, S Proc 2 and Climate
2. The test of differences is the F-test.

Table III-16

Means, Standard Deviations, Differences and P Values of Outcome Variables (in z-score form) Adjusted for 8 Input, Process and Climate PC's for the Reduced Group[1]

Period	Outcome Variable	MACOS Mean	MACOS SD	MACOS N	Non-MACOS Mean	Non-MACOS SD	Non-MACOS N	Diff[2]	P
1st Year (Posttest)	AP	.41	.63	45	-.46	.58	36	.87***	.000
	STEP	.09	.42	45	.03	.41	36	.06	.548
	IDT	.11	.63	45	.03	.63	36	.08	.611
	So Ch	-.05	.69	45	-.05	.81	36	.00	.981
	WWA	.13	.71	45	-.22	.83	36	.35	.090
	WWB	.13	.66	45	.05	.90	36	.08	.666
	CAPS-1	-.25	.87	45	.15	.73	36	-.40	.065
	CAPS-2	.11	.97	45	-.03	.70	36	.14	.538
	CAPS-3	.04	.78	45	.01	.55	36	.03	.889
	CAPS-4	.05	.74	45	-.15	.75	36	.20	.300
Follow-up 1	Skills	.19	1.01	45	-.23	.79	36	.42	.085
	Know	-.22	.81	45	.23	1.02	36	-.45	.063
	Interest	.28	.84	45	-.31	.88	36	.59*	.011
Follow-up 2	AP1-4F	.21	.67	45	-.25	.71	36	.46*	.013
	SS Ch F	-.14	.95	45	.11	.92	36	-.25	.309
	SS	-.14	.77	45	.00	.93	36	-.14	.524
	WWAF	.06	.82	45	-.02	.84	36	.08	.701
	WWBF	.19	.72	45	-.13	.64	36	.32	.110
	WWAPF	.19	.85	45	-.15	.83	36	.34	.125
	WWBPF	.02	.79	45	-.04	.97	36	.06	.795

*p .05
**p .01
***p .001

Note: Even though all p-values are given, differences with p-values of .05 or less are designated with asterisks as a visual aid to readers in identifying variables of particular interest.

1. PC's used were: Ach, Att 1, Att 2, Class 1, Class 2, S Proc 1, S Proc 2 and Climate.

2. The test of differences is the F-test.

. WWA, the scale presumably measuring attitudes toward unusual customs or beliefs, with MACOS classes more positive on the average at posttest;

. Know, a scale in Follow-up 1, October 1975 (2 months into the following year) indicating advantage of having learned certain types of content, with non-MACOS classes responding more in the direction of wishing they had learned such things than MACOS classes;[26]

. Interest, a single scale item comparing interest in social studies this year to last year, with MACOS classes on the average finding this year less interesting.

Table III-14 shows that when standardized outcome means were adjusted for all 8 student-based input and classroom variables (PC's)and for student-based process and climate measures, the significant differences between groups were largely the same. The changes from the unadjusted means with respect to the .05 level of significance, were:[27]

. CAPS 1 was significant and favored the non-MACOS classes; this was a measure of perception of ability of self as a problem solver.

. Skills, a scale in follow-up 1 indicating perceived advantage in social studies this year of having learned certain skills last year, with MACOS classes responding in the direction of wishing they had learned such skills;

. WWA did not show a significant posttest (1st year) difference between MACOS and non-MACOS groups.

---

26. It should be remembered that Skills, Know and Interest were scaled such that the higher the score, the more negative the meaning and vice versa.

27. The reader should note that all follow-up measures were adjusted for performance on pre-tests (Ach, Att 1, Att 2) given in Sept/Oct, 1974 as well as on other measures. They were not adjusted for performance on posttest, 1975 (1st year outcomes).

.WWBF, a measure presumably of attitudes towards peoples or groups who have unusual customs or beliefs, with MACOS classes on the average scoring more positively a year after their course than non-MACOS classes;

Table III-15, giving the unadjusted means and differences for the reduced sample (N=81), shows essentially the same pattern of results as the comparable table for the larger sample of classes (Table III-11). In Table III-16 it can be seen that when the reduced sample of class means have been adjusted for the student based input, process and climate PC's the results are similar to those for the larger sample, but not identical, by holding to the .05 level as the criterion of significant difference between MACOS and non-MACOS classes. In this case, differences for CAPS-1, Know, and WWBF were not significant by that decision criterion, although it can be seen in the column of p-values that CAPS-1 and Know fall just short of that criterion, and WWBF is far short of it ($p \leq .110$).

In which outcome variables should one have most confidence, based on the covariance analyses made to this point? A reasonable approach is to identify those variables on which the total samples differed significantly and for which regressions were homogeneous. By that criterion, the two groups of classes (MACOS, non-MACOS) differed significantly with respect to the following:

Posttest

AP, the course specific questionnaire

CAPS-1, perceived ability of self as problem solver

Follow-up 1

Skills, average ratings of subsequent advantage of having learned how to do certain things the preceding year;

Know, average ratings of subsequent advantage of having learned certain topics the preceding year

Interest, class average ratings of interest in social studies this year, compared to last

Follow-up 2

AP1-4F, the Man and Animals part of AP.

If one adds as a further criterion of confidence that the results were also consistently significant for both analyses (total groups, reduced groups), one would eliminate CAPS-1, Skills and Know, although CAPS-1 and Know continued to be very close to being significant at the .05 level, more so than Skills. For all three variables, if the assumption is made that the reduced sample is similar to the total sample (an assumption is supported by the generally similar unadjusted means and variances of two groups in each sample), one could view the changes as the result of reduced power, due to the small sample size. It is worth continuing interest in them.

What Would You Think A and B (WWA, WWB) show indications of differentiating between groups at this point. The indications are not consistent, however, and there also appear to be interactions between groups for these variables with Climate in some cases. Thus, they cannot be designated at this stage as variables in which one can have high confidence with respect to differences in outcomes.

3) Multiple Regression Analyses

a) Analyses with Group Entered Last

The covariance analyses indicated on which outcome variables the two groups differed. They have not indicated the relative contributions of the different input, process and climate variables if those are viewed as predictors of outcomes. To examine the contributions of different sets of predictors to outcomes, fixed-ordered stagewise regression analyses were employed. In such analyses, the primary question of interest is: which sets of PC's add a significant increment of variance accounted for, relative to the total proportion of variance (R) accounted for by all predictors? The analyses were made for the MACOS and non-MACOS groups separately in order to evaluate differences in the possible contributions of different sets of PC's. They were also made for the combined sample, using Group (MACOS, non-MACOS) as a dummy variable. Two sets of such analyses were made in which the order of predictors was varied. The first entered Group last after all other variables. That procedure is analogous to the covariance analyses made in that it tests the null hypothesis that, when the variance accounted for by all other variables has been partialled out, group membership does not account for a significant increment in the remaining variance. The analysis was also made entering Group after input and background variables, and before process and climate variables. That analysis is analogous to asking: if all that has been removed is variance from pretest and background characteristics, does Group classification (MACOS, non-MACOS) contribute a significant increment to the total variance that could be accounted for?

Except for Group as noted above, sets of PC's were always entered into the analyses in a predetermined order. The order was based on assumptions of

a temporal priority of variables with respect to effects or outcomes that is implicit in an input-process-outcome model. Pretest variables for students were entered first (achievement-Ach; attitude - Att 1, Att 2); Classroom characteristics (Class 1, Class 2) were entered next, then process variables based on students ratings (S Proc 1, S Proc 2); and finally Climate. For that order, Group was entered either after the class set (Class 1, Class 2) or last, after the climate PC. For analyses that included teacher variables, teacher background and attitude variables (T Demo; T Psy 1 and T Psy 2) were entered after class characteristics, and teacher process variables were entered before student process variables. Group in this case was entered after the teacher background variables, and also last after the climate PC in the second run.

One reason for maintaining a fixed order of entry of sets of variables was to minimize effects of multicollinearity from one analysis to another. As has been noted, there are non-trivial correlations among some of the PC's. It was believed that by imposing an order based on a conceptual model of the nature of the variables, the effects of multicollinearity that can lead a regular stepwise regression analysis to capitalize on unstable correlations were held constant. It is true that different orders of entry can change the increment of variance associated with a particular set of variables. Normally variables entering the analysis later are less likely to show significant increments. Thus, one may anticipate different increments for Group when it is entered after input than when it is entered last. Similarly, entering Climate always last or next to last reduces the likelihood of detecting a significant increment with it. It was taken as a

working hypothesis, however, that if variables were to be ordered in a temporal sequence, climate could be viewed as a consequence of students and teachers interacting in a social studies course. There are, of course, other hypotheses, but that was the model adopted for purposes of these analyses.

In presenting results, increments from sets of PC's are combined into groups designated input, process, climate and Group (MACOS, non-MACOS). As noted above, input contains all student pretest PC's, classroom characteristics PC's, and when teacher variables were included, teacher background and pre-attitude PC's. The test of significance of an increment in proportion of variance associated with each of the four sets of variables (input, process, climate, group) is the F-test of increments.[28] The denominator for all tests is $(1-R^2y.ABCD)/(n-k-1)$, where $R^2y.ABCD$ is the total multiple R^2 for the analysis and k is the total number of variables in all sets. That model does involve a small sacrifice in power, especially for sets containing 5 or 8 variables. For the large sample analyses using student PC's for tests of the Group increment or climate, the F-test has, at the .05 level of significance, a power between 50-60% to detect an increment of 5%; it has a power between 85-90% to an increment of 10%. These values are based on no assumptions about the size of the total R^2.[29] For the same sample there are 5 input PC's. The F-test has a power of 60-70% to detect an increment of 10%, and a power between 80-85% to

28. Draper, N.R., and Smith, H. Applied Regression Analysis. New York: Wiley, 1966.

29. They are based on an error degrees of freedom of 92 (102-8 student-based PC's-Group-1).

detect an increment of 15%.[30] Power for the reduced sample, with the larger number of PC's is less, or conversely, the increment that can be detected at the .05 level with a given power is larger.

Finally, if a total multiple R^2 was not significant, no tests of its increments were made in order to protect the Type I error rate. If there was not sufficient reason to believe the amount of Y variance accounted for was greater than zero, there was little justification for analyzing increments. By the same token, if the increment for a group of sets (e.g., input sets of PC's) was not significant, no further analysis was made.

Table III-17 shows the results of the analyses of outcome variables when Group (MACOS, non-MACOS) was entered last into each analysis. As they should be, these results are essentially the same as those obtained by the analysis of covariance. The posttest variables for which Group still added a significant increment to remaining variance with input, process and climate variance partialled out, are AP (the MACOS course content questionnaire), and CAPS-1 (perceived ability of self as problem solver). WWA (posttest) was the exception. The Follow-up 1 outcome variables for which Group added significantly are Know and Interest. The Follow-up 2 outcome variables for which Group added significantly are AP1-4F (the Man and Animals part of AP), and

30. Actual increments represented by these percentages will depend on the total R^2. Consider the increments of 5 and 10% for Group or for Climate. Consider also four different total R^2's: .2, .4, .6, .8. The actual observed increments detected at the .05 level for the level of power indicated would be:

	Total R^2			
Increments	.2	.4	.6	.8
5%	4%	3%	2%	1%
10%	8	6	4	2

Table III-17

Increments of Proportion of Variance in Outcome Variables Associated With Input, Process, Climate and Group Variables (MACOS, Non-MACOS) Using Student PC's[1]

Period	Outcome Variable	Pre-post r^2	Total Multiple R^2	Increment from Input	Increment from Process	Increment from Climate	Increment from Group (M, N-M)
Posttest	AR	33**	60**	44**	2	1	13**
	STEP	81**	78**	76**	0	1*	0
	IDT	55**	62**	57**	3	1	1
	SS Ch	35**	26**	9	3	13**	1
	WWA	18**	36**	23**	12**	0	1
	WWB	16**	27**	18**	7*	2	0
	CAPS-1	22**	21**	11*	1	5*	4*
	CAPS-2	15**	27**	23**	3	1	0
	CAPS-3	36**	48**	42**	5**	1	0
	CAPS-4	26**	29**	27**	1	0	1
Follow-up 1	Skills	-[2]	13	0	8	1	4
	Know	-	25**	6	13**	0	6**
	Interest	-	22**	4	9**	3	6**
Follow-up 2	AP1-4F	32**[3]	55**	48**	2	0	5**
	SS Ch F	1[4]	10	6	2	1	1
	SS	-	20*	3	4	13**	1
	WWAF	1[5]	27**	24**	1	0	1
	WWBF	14**[6]	32**	27**	1	0	4*
	WWAPF	-	8	6	1	0	1
	WWBPF	-	16	15	0	1	0

*$p \leq .05$
**$p \leq .01$

1. Sample sizes for posttest outcomes: MACOS 55, non-MACOS 47. For FU-1 and FU-2 outcomes: MACOS 54, non-MACOS 43. Pre-post correlations, squared, are given for comparison with Multiple R^2. Note: incremental proportions may not add exactly to R^2 due to rounding. Decimals and leading zeroes have been omitted.
2. A dash indicates there was no pre-test for the variable.
3. Pretest was total AP. For pre AP1-4 and AP1-4F, $r^2 = .37$.
4. Pretest was SS Ch: SS Ch F was a modification of SS Ch.
5. Correlation is with WWA pre.
6. Correlation is with WWB pre.

WWB(the scale measuring reactions to people or groups with unusual customs).

It is apparent in Table III-17 that, in most cases, the set of variables classified as input (student pre achievement and attitude, and classroom characteristics) accounts for the overwhelming proportion of total variance of the variables. There are, however, several interesting exceptions. Social Studies Choices (SS Ch), the average number of times a class picked social studies in preference to 5 other subjects, showed a larger and highly significant increment due to classroom climate. Another instance was a measure employed in Follow-up 2: SS, a scale on which students made an absolute rating of how much they liked social studies. For SS, Climate (from the prior year) also accounts for the major portion of total variance.

Climate also contributed a small but significant increment to STEP posttest class averages, and to CAPS-1 posttest class averages. Process variables contribute significant increments to WWA (the posttest measure of reactions towards unusual customs or beliefs), WWB (the posttest measure of reactions towards people or groups having those customs or beliefs), CAPS-3 (the posttest measure of tolerance for ambiguity), Know (the Follow-up 1 measure of perceived advantage of having learned certain topics), and Interest (the Follow-up 1 measure of interest in social studies this year compared to last year).

It can also be seen in Table III-17 that for 7 of the 10 posttest measures, the total variance accounted for (total multiple R^2) by all PC's, plus Group, exceeds pre-post r^2 for the original variables. The exceptions are STEP, SS Ch and CAPS-1. In these cases it appears that by creating pretest composites there was some loss of variance accounted for by the individual pretest.

The loss is slight for STEP (2%); it is around 10% for the other two measures.

For the achievement measures (AP, STEP, IDT, AP1-4F), the input PC's and Group account for the majority of variance of the measures. For the rest they do not. That may be due in part to the lower reliability of the other measures, (i.e. the attitude measures).

1)) Contributions of Individual PC's

In Table III-17 the increments due to Climate and to Group are directly interpretable in the sense that there is only one variable in each of those sets (there is only one climate PC, and Group is a single dummy variable). For the cluster of variables called 'input', however, there are three sets of PC's, each set entered in a fixed order: pretest achievement, Ach; pretest attitude, Att 1, Att 2; and classroom demographic characteristics, Class 1, Class 2. For the group called 'process' there is one set containing two PC's: informal, unstressed group oriented classes, as rated by students, S Proc; and traditional, individual work oriented classes, S Proc. It is of interest to know which of those sets and variables was significant when entered into the analysis since it will delineate more specifically predictive relationships with particular outcome variables. For example, it is of interest to know if classroom demographic characteristics, classified in the analyses as inputs along with pretest achievement and attitude, account for a significant proportion of the total variance of, say, the MACOS course specific questionnaire, or the STEP test, after proportions of variance for preachievement and preattitude have been removed.

The procedure followed for these further analyses was to determine

whether, for a given dependent variable, the increment for input was significant. If it was, then the increment for each set, when it was entered, was tested for significance at the .05 level.[31] If the set was significant, then the F-test for the significance of the regression coefficient was examined for each variable in the set (one or two, depending on the set) as a means of identifying which (in the case of pairs of variables) was significant. The same procedure was applied to the analysis of the set of two process PC's.

As a final step, the F-tests of the regression coefficients for variables found to be significant when their entry was forced were examined after all variables had been entered. That is a test of the significance of the variable as a predictor if it were entered into a regression analysis last, after all other variables. It is interpreted here as further corroboration of the importance of the variable with respect to the dependent variable in question. It may be noted that the test of coefficient for Group by definition is the same, since it was always entered last. Since Climate was entered before Group, an examination of its regression coefficient if it were to be entered last is comparable to the final analysis of input and process variables. In all cases in which the increment for the Climate PC was found to be significant, it was also found to be significant if tested after all other variables.

The results of these further analyses are shown in Table III-18 for the

31. It is noted again that this further procedure was not applied if the total Multiple R^2 w[as] not significant. The reason for the stepwise testing within groups [of] sets of variables, or within a group, was to protect the Type I er[ror] rate.

dependent variables that showed significant increments for input or process or both. The table shows, for the specified dependent variable, the increment to R^2 added by the variable or variables in a set if the set was found to be significant, and the standardized regression weight (Beta). It also shows which variable (or variables) was significant within the set, and whether it was also significant if added last after all variables.

It may be seen in Table III-18 that with respect to input sets of variables, the significant sets and variables (PC's) are in all cases except one (CAPS-4, perceived creativity of self, class average), pretest achievement (predominantly) and attitude. In the case of CAPS-4, both class demographic characteristics PC's were significant. In both cases the regression weights were negative, suggesting that more positive performance of classes with respect to CAPS-4 at posttest would be associated with classes having higher proportions of minority and low-income students and small size (the opposite of the interpretations given for the two class PC's).

The process variables associated significantly with particular outcome variables depend on the variable, as does the direction of the relationship (the sign of the regression coefficient). With respect to WWA, both student-based process PC's have negative coefficients. With respect to three of the four remaining cases, the student process variable seem interpretable. The less students rated the class as traditional and non-group oriented, and the more positively the class was rated with respect to informality, group discussion oriented activities, the higher the ratings (on the average for a class) of tolerance of ambiguity (CAPS-3). The higher the rating of the class as traditional (S Proc 2),

Table III-18

Increments of Proportions of Variance and Standardized Regression Coefficients of Variables in Sets That Added a Significant Increment of Proportion of Variance to Outcome Variables[1]

		Significant Input Variables				Significant Process Variables			
		When Entered In Fixed Order			When Entered Last	When Entered In Fixed Order			When Entered Last
Period	Outcome Variable	Var.	Incr.	Beta	Beta	Var	Incr.	Beta	Beta
Posttest	AP	Ach	43	65**	56**				
	STEP	Ach	74	86**	81**				
		Att 2	1	10*	8				
	IDT	Ach	56	75**	64**				
	WWA	Ach	14	38**	25	S Proc 1	4	-22*	-25*
		Att 1	7	32**	33**	S Proc 2	8	-28**	-24*
	WWB	Ach	12	34**	16	S Proc 2	6	-26**	-18
		Att 1	5	28*	27*				
	CAPS-2	Att 2	17	41**	35**				
	CAPS-3	Ach	23	48**	12	S Proc 1	3	23**	23**
		Att 1	16	48**	46**				
		Att 2	2	15*	15				
	CAPS 4	Att 2	20	45**	45**				
		Class 1	23	-31*	-27				
		Class 2	27	-21**	-22*				
Follow-up 1	Know					S Proc 2	12	35**	25*
	Interest					S Proc 2	13	-27**	-13
Follow up 2	WWAF	Ach	21	46**	32*				
	WWBF	Ach	22	47**	26				
	APFU	Ach	2	67**	43**				

*p ≤ .05
**p ≤ .01

1. Sample sizes for posttest analyses: MACOS 55, non-MACOS 47; for FU-1 and FU-2, MACOS 54, non-MACOS 43.

Note: decimal points and leading zeroes for incremental proportions of variance and for standardized regression weights have been omitted.

the higher the ratings of Know in Follow-up 1. Higher ratings for Know mean students rated certain topics as not having been studied last year, and it would have been advantageous this year if they had. The less students rated a class as traditional (S Proc 2) the more they were apt to rate this year's social studies class in Follow-up 1 as less interesting than last year's class.

2)) Separate Analyses of MACOS and Non-MACOS Classes

Table III-19 shows the results of analyses of sets for each group of classes, MACOS and non-MACOS. The tables help clarify relationships of sets of variables in each group. It must be remembered, however, that there is a loss of power with these reduced sample sizes. For example, for the MACOS group, the power of the F-test to detect, at the .05 level, a significant increment of proportion of variance of 5% due to climate is between 30-50%; the power to detect an increment of 10% is 50-60%. Again, these values are taken without regard to the actual value of total R^2. The power of the F-test to detect an increment due to the two process variables of 10% is between 30 and 50%; the power to detect an increment of 15% is 60-70%. For the non-MACOS group, the power of the F-test to detect an increment of 15% for climate is 60-70%. For the non-MACOS group, the power of the F-test to detect a significant increment, at the .05 level, of 5% is 10-30%; of 10% about 50%; of 15%, power is between 60-70%.[32] For the two variable process set in the non-MACOS group, the power of the F-test to detect an increment of 15% is 50-60%. and less for smaller increments. Thus, one is not to be surprised that

32. Again, observed effect sizes represented by these percentages will depend on the size of the total multiple R^2.

Table III-19

Relationships of Sets of Student Based Input, Process and Climate Variables to Outcome Variables, by Group[1]

Outcome Variable	Total Multiple R^2		Increment from Input		Increment from Process		Increment from Climate	
	MACOS	Non-MACOS	MACOS	Non-MACOS	MACOS	Non-MACOS	MACOS	Non-MACOS
AP	46**	62**	46**	59**	0	2	0	1
STEP	79**	80**	77**	77**	0	0	1	2
IDT	65**	62**	62**	56**	1	-6	2	0
SS Ch	26	34**	14	17	8	5	4	11*
WWA	32*	55**	24**	27**	3	21**	5	7*
WWB	25	43**	22	27**	1	11*	3	5
CAPS-1	11	42**	8	21*	2	1	1	20**
CAPS-2	29*	41**	24**	32**	3	4	1	6
CAPS-3	42**	59**	36**	52**	7	4	0	2
CAPS-4	29*	33*	28**	33**	1	1	0	0
Skills	16	23	2	2	10	16	5	5
Know	20	21	11	5	8	15	1	1
Interest	15	14	8	8	6	2	1	1
AP1-4F	61**	53**	59**	47**	2	5	0	1
SS Ch F	16	18	14	9	1	4	0	5
SS	20	32	12	6	5	5	3	20
WWAF	28	36*	26	34*	2	2	0	0
WWBF	33*	45**	24**	42**	9	2	0	1
WWAPF	35**	41*	29**	27*	5	13*	1	0
WWBPF	17	29	12	21	2	1	2	7

*$p \leq .05$

**$p \leq .01$

1. Input: Ach, Att 1, Att 2, Class 1, Class 2; Process: S Proc 1, S Proc 2; Climate. Sample sizes for Posttest Outcomes (AP-CAPS-4): MACOS 55, Non-MACOS, 47. For FU-1 and FU-2 Outcomes: MACOS 54, Non-MACOS 43. Note: All figures are proportions of variance; decimal points and leading zeroes have been omitted. Incremental proportions may not add exactly to R^2 due to rounding. Increments for variables for which total R^2 was not significant were not tested.

there appear to be fewer significant increments at the .05 level of significance. Perhaps the more important point is that where there are significant increments due to process or climate, one should be particularly interested in them.

In Table III-19 it can be seen that input continues in both groups to be the typically overwhelming predictor of outcome. There are cases in which the non-MACOS group appears to show a significant increment due to input while the non-MACOS group does not. Close examination shows that those are all cases in which the overall R^2 was not significant in the MACOS groups; therefore increments were not tested.

Examination of Table III-19 suggests that there are some differences between MACOS and non-MACOS classes with respect to process and climate variables, even with increased effect sizes detectable with respect to given levels of power. The variables of interest are SS Ch, WWA, WWB, CAPS-1 and WWAPF, all attitude variables, although it must be noted that total R^2 was significant in the MACOS group only for WWA and WWAPF. Student related process variables contribute significantly (indeed, strongly) in the non-MACOS group to WWA; they also contribute in the same way for WWB. Climate contributes significantly to CAPS-1 in the non-MACOS classes. It should be recalled, of course, that significant interactions between groups were found for Climate and WWA. They were not found for the other three variables. For WWAPF, there is a significant increment due to process for the non-MACOS classes, not for the MACOS classes. Indeed, the fact that WWAPF R^2 reached significance in both group, but not in the combined groups, is in itself interesting. It suggests that the interaction effects found earlier are potent with this variable.

Table III-19A

Raw Regression Coefficients of PC's and Constants (Intercepts), for Each Outcome Variable, by Group[1/]

Outcome Variable		Ach	Att 1	Att 2	Class 1	Class 2	S Proc 1	S Proc 2	Climate	Constant
AP	M	34**	08	07	03	-06	01	02	-00	369
	N	34**	-01	09	02	-05	07	04	06	-410
STEP	M	50**	07	04	02	11	02	05	12	-075
	N	48**	09	04	-09	00	04	05	10	052
IDT	M	34*	12	04	09	08	03	07	13	019
	N	45**	04	03	-07	09	02	16*	-01	-162
SS Ch(x)	M	-08	21*	12	-01	-13	-08	-10	17	024
	N	-05	-02	13	12	10	-08	-15	24	-195
WWA	M	26*	21*	-03	-09	-04	-04	-13	-20	176
	N	02	19	-22	16	-13	-11	-19*	21*	-050
WWB	M	16	04	15	07	13	-05	00	14	089
	N	05	35**	-13	-10	-13	-01	-18	17	037
CAPS-1	M	-04	21	05	05	-11	-12	01	10	-157
	N	06	04	12	01	-17*	10	01	32**	362
CAPS-2	M	12	09	35*	-10	20	-11	-06	12	010
	N	16	15	19	-16	-14	-04	-02	15	013
CAPS-3	M	09	34**	18	00	-05	09	-10	06	-059
	N	04	33**	07	00	00	14*	05	11	-029
CAPS-4	M	13	-04	39**	-18	-02	-05	01	03	048
(x)	N	21	-00	41*	-14	-24*	-05	-04	-03	-123
Skills (x)	M	-11	-03	-00	16	13	03	27*	22	158
	N	-04	12	04	-06	12	-05	21*	-14	-335
Know (x)	M	10	02	15	-11	14	10	15	08	-316
	N	00	07	05	02	19	05	24*	-07	171
Interest	M	-11	-00	-07	15	-08	-03	-11	06	221
	N	-06	14	08	01	-05	08	-04	08	-302
AP1-4F	M	16	14	-09	22*	10	04	-07	-02	131
	N	36*	04	07	06	08	02	12	-08	-398
SS Ch F(x)	M	12	05	07	-11	-33*	-04	-08	-03	-029
	N	-03	-22	19	12	-12	05	-10	17	211
SS (x)	M	08	09	-02	-17	-19	-08	-07	14	-089
(x)	N	06	-24	-08	07	-06	-09	06	37**	070
WWAF (x)	M	25*	07	-22	01	-09	-07	-04	-04	084
	N	12	-02	11	25	04	04	-10	-01	-150
WWBF	M	14	04	-20	18	13	-15*	-02	-03	139
	N	17	18	27	13	-01	07	-10	-10	-266
WWAPF	M	-14	-17	-20	34**	21	04	09	-10	116
	N	-04	19	-10	-40*	-14	-06	-25	03	048
WWBPF(x)	M	29*	-04	-13	-23	07	-01	03	-14	114
(x)	N	47*	-10	-17	-44*	-04	08	-00	22	033

*p ≤ .05; **p ≤ .01.

Note: Significance levels refer to the F-test of the significance of the coefficient.

1. M = MACOS, N = Non-MACOS. Sample sizes for posttests: MACOS = 55; Non-MACOS = 47. Sample sizes for FU-1 and FU-2 variables: MACOS = 54; Non-MACOS = 43. Note: decimal points have been omitted from all coefficients and constants, although all are in decimal form due to the scaling of the PC's. (x) = variables for which the total multiple R^2 was not significant at the .05 level. See Table III-17.

Note: The regression coefficients and the constants in the final regression equation are shown.

Table III-19A gives the raw regression coefficients and intercepts for each PC for each dependent variable, by group.

The interpretation at this point is that: 1) there were outcome variables for which there appear to be differences between groups with respect to process or climate variables, even with the loss of power due to reduced sample sizes; 2) those variables are attitudinal, not achievement; and 3) process and climate variables appear to be of more importance for non-MACOS classes than MACOS classes.

b) Analyses with Group Entered after Input

Table III-20 shows results for the total samples when Group (MACOS, non-MACOS) is entered into the analysis following input (Ach, Att 1, Att 2, Class 1, Class 2). There are two changes for the increment due to Group from the case in which that variable was entered last. The increments due to Group are no longer significant at the .05 level for CAPS-1 and for WWBF (What Would You Think, Part B, in Follow-up 2), although in both cases the p-values marginally exceeded .05, as did the increment for WWA. There are several possible reasons for these differences. One is that they are simply chance fluctuations of significance for outcome variables that have previously shown marginal or fluctuating significance in other analyses. Another possibility is that there may be suppressor effects of some variables which, when removed, produce significance for the group variable.

The overall pattern of results from entering Group after input is very similar to the results of entering it last, after the variance of eight predictors has been partialled from the criterion variable variance. The only change with respect to process variables is associated with Interest (in Follow-up 1), for which the set of process PC's no longer contribute significantly. The only change with

Table III-20

Increments of Variance in Outcome Variables Associated With Input, Group, Process and Climate Variables Using Only Student-Based PC's[1]

Period	Outcome Variable	Pre-post r^2	Total Multiple R^2	Increment from Input	Increment from Group	Increment from Process	Increment from Climate
Posttest	AP	33**	60**	44**	16**	1	0
	STEP	81**	78**	76**	0	0	1*
	IDT	55**	62**	57**	1	3	1
	SS Ch	35**	26**	9	3	3	11**
	WWA	18**	36**	23**	3	10**	0
	WWB	16**	27**	18**	2	5*	2
	CAPS-1	32**	21**	11*	3	1	7**
	CAPS-2	15**	27**	23**	1	3	1
	CAPS-3	36**	48**	42**	1	5*	1
	CAPS-4	26**	29**	27**	1	1	0
Follow-up 1	Skills	–[2]	13	0	2	10	0
	Know	–	25**	6	10**	9**	0
	Interest	–	22**	4	13**	5	2
Follow-up 2	AP1-4F	32**[3]	55**	48**	6**	1	0
	SS Ch F	1[4]	10	6	0	2	2
	SS	–	20*	3	0	4	13**
	WWAF	1[5]	27**	24**	1	1	0
	WWBF	14**[6]	32**	27**	3	1	1
	WWAPF	–	8	6	1	0	0
	WWBPF	–	16	15	0	0	1

*$p \leq .05$
**$p \leq .01$

1. Sample sizes for posttest outcomes: MACOS 55, non-MACOS 47. For FU-1 and FU-2 outcomes: MACOS 54, non-MACOS 43. Pre-post correlations, squared, are given for comparison with Multiple R^2.
   Note: incremental proportions may not add exactly to R^2 due to rounding. Decimals and leading zeroes have been omitted.
2. A dash indicates there was no pre-test for the variable.
3. Pretest was total AP. For pre AP1-4 and AP1-4F, $r^2 = .37$.
4. Pretest was SS Ch; SS Ch F was a modification of SS Ch.
5. Correlation is with WWA pre.
6. Correlation is with WWB pre.

respect to Climate is associated with WWA, posttest, for which Climate no longer adds a significant increment. That is possibly because the interaction observed in the homogeneity of regression analyses has been absorbed by the group variable. Otherwise, while the absolute sizes of increments change somewhat in this analysis from what was previously found, the overall patterns appear to remain quite stable. That is taken to lend further confidence to the identification and interpretation of group effects and of process and climate relationships.[33]

Table III-20A shows for each outcome variable the standardized regression coefficients of input PC's at the point at which Group was entered, and of all predictors after all were in the equation.

c) Analyses with Teacher PC's Included

How do teacher variables contribute to the variance in outcome measures? Table III-21 gives the results of analyses with Group (MACOS 5, non-MACOS) added last. In these analyses, input includes in this order of entry: Ach, Att 1/Att 2, Class 1/Class 2, T Demo, T Psy 1/T Psy 2. That is, teacher background and attitude PC's were entered after student pretest and class PC's. Process includes in this order: T Proc 1/T Proc 2, and S Proc 1/S Proc 2. Process PC's based on teacher's ratings were entered before student process PC's.

Table III-21 shows largely the same patterns as the preceding analyses. The differences are at least in part due to reduced power (the sample sizes are smaller), as discussed earlier.[34] It is also possible that the reduced samples

33. Input variables in this analysis are of course identical in characteristics to what they were in the prior analysis.

34. In this case, with N = 81 and error degrees of freedom = 66, power for the .05 level to detect an increment for Group or Climate of 5% is between 30-50%; it is between 70-75% for an increment of 10% (cont'd, p. III-78).

Table III-20A
Standardized Regression Coefficients for PC's for Each Outcome Variable at the Step in Which Group was Entered Before Process/Climate PC's and for Each Predictor Entered[1]

Outcome Variables	Input						Process/Climate PC's		
	Ach	Att 1	Att 2	Class 1	Class 2	Group	S Proc 1	S Proc 2	Climate
AP	58**	05	10	04	09	-40**			
	56**	04	09	02	-08	-40**	07	07	04
STEP	83**	09	09	-04	04	02			
	81**	08	05	-06	05	06	07	08	14*
IDT	67**	09	06	02	05	-08			
	64**	07	05	00	07	-08	12	16*	09
SS Ch	-11	16	27**	05	-02	-18			
	-10	16	15	05	-00	-09	-13	04	40**
WWA	20	30**	-09	02	-03	-17			
	25	33**	-11	05	-05	-13	-25*	-24*	01
WWB	15	26*	09	05	01	-15			
	16	27*	04	03	-01	-07	-07	-18	15
CAPS-1	03	21	20*	-01	-19	17			
	03	20	11	01	-17	22*	-09	09	30*
CAPS-2	21	14	38**	-24	00	-08			
	23	15	35**	-24	-01	-04	-11	-10	10
CAPS-3	15	47**	17*	08	-02	-07			
	12	46**	15	-01	-04	02	23*	-07	12
CAPS-4	24	-02	44**	-30*	-22*	-08			
	26	-02	45**	-27	-22*	-10	-09	-03	-05
Skills (x)	-09	01	01	06	03	-15			
	-13	-01	00	11	09	-22*	07	35**	06
Know	14	06	10	09	15	32**			
	09	05	11	-08	19	26*	18	27*	-01
Interest	-13	08	12	24	-07	-36**			
	-14	07	09	17	-09	-26*	10	-13	13
AP1-4F	46**	07	-01	22	07	-24**			
	48	07	00	21	08	-25**	11	04	-05
SS Ch F (x)	08	-06	15	-02	-23	05			
	10	-06	10	-03	-24*	12	-06	-07	16
SS	10	-03	07	-15	-12	-03			
	12	-06	-04	-15	-13	09	-19	02	43**
WWAF	30*	09	-11	16	02	-11			
	32*	09	-12	17	01	-10	-09	-05	01
WWBF	23	15	-03	20	07	-17			
	26	16	-01	24	07	-21*	-13	-03	-10
WWAPF(x)	-13	-10	-17	07	05	-09			
	-12	-09	-16	07	04	-09	-03	-09	-06
WWBPF (x)	58**	-07	-12	-51**	-01	-07			
	58**	-08	-15	-49**	00	-07	-03	10	10

*p ≤ .05; **p ≤ .01.

Note: Significance levels refer to the F-test of the significance of the regression weight.

(x) = variables for which total multiple R^2 was not significant at the .05 level.

1. For each outcome variable, the top line shows the regression weight for input PC's and Group at step in which Group was entered. The second line shows the weights for each PC and for Group if each were entered last after all other variables. Posttest N = 102; FU-1 and FU-2 N= 97.

Note: decimal points have been omitted from all weights.

Table III-21

Increments of Proportion of Variance in Outcome Variables Associated With Input, Process, Climate and Group Variables (MACOS, Non-MACOS) Using Student and Teacher PC's[1]

Period	Outcome Variable	Pre-post r^2	Total Multiple R^2	Increment from Input	Increment from Process	Increment from Climate	Increment from Group (M, N-M)
Posttest	AP	33**	64**	46**	1	0	16**
	STEP	81**	79**	78**	0	1*	0
	IDT	55**	56**	53**	2	1	0
	SS Ch	35**	37**	13	9*	15**	0
	WWA	18**	50**	30**	15**	0	4**
	WWB	16**	33*	24**	7	1	1
	CAPS-1	32**	27	17	2	4	4
	CAPS-2	15**	40**	31**	7	1	1
	CAPS-3	36**	55**	49**	4	2	0
	CAPS-4	26**	29*	25**	2	1	0
Follow-up 1	Skills	-[2]	26	8	12	3	3
	Know	-	34*	8	18**	1	7**
	Interest	-	31*	8	17**	2	4*
Follow-up 2	AP1-4F	32**[3]	59**	55**	1	0	3*
	SS Ch F	1[4]	12	6	3	2	0
	SS	-	24	5	7	11	0
	WWAF	1[5]	37**	31**	6	0	0
	WWBF	14**[6]	38**	33**	4	0	1
	WWAPF	-	17	11	4	0	2
	WWBPF	-	19	17	1	0	0

*p ≤ .05

**p ≤ .01

1. Sample size for all outcomes = 81. MACOS 44, Non-MACOS 37.
   Pre-post correlations, squared, are given for comparison with Multiple R^2.
   Note: incremental proportions may not add exactly to R^2 due to rounding.
   / Decimals and leading zeroes have been omitted.
2. A dash indicates there was no pre-test for the variable.
3. Pretest was total AP. For pre AP1-4 and AP1-4F, $r^2 = .37$.
4. Pretest was SS Ch; SS Ch F was a modification of SS Ch.
5. Correlation is with WWA pre.
6. Correlation is with WWB pre.

are not simply random samples of the larger samples. That is, some differences may be due to the samples having characteristics that are a systematic, not random, part of the total group. The principal changes of interest center around WWA, which here is significant, and WWBF, which is not. Overall, however, patterns of increments appear similar to those found with larger samples.

The reason for doing the analyses with the reduced samples was to be able to examine the relationships of teacher variables, along with student ones, to different measures. The same analytic procedure as before has been followed, with the following exception. If a group of sets of variables was significant in the larger samples but not in the reduced samples, the group was also examined in the latter. The justification for this departure from the general principal followed in these analyses was to compensate for the loss of power due to reduced sample sizes. The justification of course rests on the assumption that the reduced sample was a random sample of the larger group. It was believed, however, that conclusions could be qualified to minimize the risks of misstatement or erroneous interest.

Table III-22 indicates the results of analyses of sets of input and process variables. It shows for each outcome variable analysed, the increment of variance added by the indicated PC, if its set was significant, and the regression coefficient of the PC when it was entered and in the final equation.

It may be seen that there are outcome variables for which the teacher background set does contribute a significant increment of variance, and that in four

---

again, these increment estimates are not observed increments, which will depend on the size of the total multiple R^2. For the set of process PC's (4 in all), power of the F-test is between 10 and 30% for an increment of 5%; between 50-60% for a 10% increment, and between 70-75% for a 15% increment. For a set of two PC's the power of the F-test for the .05 level of significance is between 30-35% to detect an increment of 5%, between 60-70% to detect one of 10%, and between 80-85% to detect an increment of 15%.

Table III-22

Increments of Proportions of Variance and Standardized Regression Coefficients of Variables in Sets That Added a Significant Increment of Proportion of Variance to Outcome Variables[1]

Period	Outcome Variable	Significant Input Variables: When Entered In Fixed Order — Var.	Incr.	Beta	When Entered Last — Beta	Significant Process Variables: When Entered In Fixed Order — Var	Incr	Beta	When Entered Last — Beta
Posttest	AP	Ach	40	63**	60**				
		T Psy 2	3	-19*	-15				
	STEP[2]	Ach	74	86**	83**				
	IDT	Ach	51	71**	62**				
	SS Ch					S Proc 1	5	-29*	-22
	WWA	Ach	11	33**	-02	S Proc 1	6	-31**	-30**
		Att 1	13	44**	50	S Proc 2	8	-27**	-27**
		T Psy 2	2	-15	-22*	T Proc 1	1	-11	-30**
	WWB[3]	Ach	10	31**	05				
		Att 1	9	36**	35**				
	CAPS-1	Ach	10	31**	19				
	CAPS-2	Ach	2	14	32*				
		Att 2	18	43**	35**				
		Class 1	5	-32*	-37*				
	CAPS-3	Ach	26	51**	28*	S Proc 1	3	23*	26*
		Att 1	9	34**	31**				
		Att 2	8	28**	25**				
		T Psy 2	5	-22*	-13				
	CAPS-4	Att 2	21	46**	46**				
Follow-up 1	Know Interest					S Proc 2	15	43**	38**
						T Proc 2	14	-43**	-26
Follow-up 2	AP1-4F	Ach	44	66**	49**				
		T Psy 1	5	-25**	-19*				
	WWAF	Ach	15	39**	28				
		T Psy 1	6	26*	22*				
	WWBF	Ach	16	40**	23				
		T Psy 2	6	-26**	-29**				

*p ≤ .05

**p ≤ .01

1. Sample sizes for all outcomes: MACOS 44, Non-MACOS 37.

2. Att 1 and 2 together added a significant increment of 2; neither PC significant alone.

3. Student process set significant; neither PC alone significant.

Note: decimal points and leading zeros for incremental proportions of variance and for regression coefficients have been omitted.

out of six such cases it was T Psy 2 (the PC related to scoring higher on Educational Scale VII Traditionalism, and Teachers at Work). The outcome variables involved are AP, WWA, CAPS-3, AP1-4F, WWAF and WWBF. The direction of association of T Psy 2 in all but one case is negative (that is, the sign of the regression coefficient is negative), suggesting that higher values of that PC predict a lowering of class average scores for the variables. The interesting case is AP. At post-test, here the implication appears to be that more conservative teachers are predictive of decreased scores on the MACOS questionnaire. For AP 1-4F in Follow-up 2, the teacher set of input PC's adds a significant increment of variance, but it is the other PC (T Psy 1) that is significant. The direction of predictive association is negative. For WWAF, T Psy 1, (general emphasis on many different objectives, not associated with traditionalism or progressivism) is a significant predictor and positive in direction of influence.

Teacher-based process PC's contribute significant increments of variance to WWA and to Interest (in Follow-up 1). The PC involved in WWA, along with both student process PC's, is T Proc 1. All process PC's have negative regression coefficients, suggesting a complex set of possible relationships. It will be recalled that this outcome variable showed an interaction of groups with respect to Climate. It is possible that by entering process variables in the analysis prior to entering Climate, they are are including some effects of Climate. T Proc 2 is associated with Interest in Follow-up 1, in a seemingly interpretable way. It would appear that the more teachers tended to rate emphasis on individual work, recall or comprehension, the more students tended to find this year's class more interesting than last year's. For other cases in which sets of process variables accounted for significant increments of variances it was student based process ratings, as

found in the earlier analysis.

4) Conclusions From Analyses of Covariance and Multiple Regression Analyses

Differences between MACOS and non-MACOS class means for twenty outcome variables have been examined by two major methods: multivariate and univariate analyses of variance and covariance, and multiple regression analyses. In addition, the associations of input, process and climate variables have been examined to identify potentially influential variables (PC's) with respect to the different outcome variables. The following are the main evaluation findings at this stage.

1) There is consistent evidence in all analyses that there were significant differences between MACOS and non-MACOS classes with respect to the following outcome variables:

Posttest

AP, the MACOS specific test; MACOS classes on the average scored significantly higher.

Follow-up 1.

Know, an average of summated ratings of whether certain subjects had been learned last year and how advantageous it was this year in social studies; non-MACOS classes on the average had higher scores, in the direction of indicating they had not learned them and wished they had, than MACOS classes;

Interest, a single scale indicating how interesting students found social studies this year compared to last year; MACOS classes on the average rated it in the direction of being less interesting this year.[35]

Follow-up 2

AP4-4F, the man and animals part of AP; MACOS classes on the average continued to score higher than non-MACOS classes.

35. This scale was repeated in Follow-up 2. It was not used in the preceding analyses since it was believed that Interest in FU-1 was based on the more immediate contrast for students. It was found that the relationship described here for FU-1 continued to hold in FU-2.

2) There is evidence, although it was not consistent in all analyses, that there were differences between the two groups with respect to the following outcome variables.

Posttest

WWA, an average rating of reactions to unusual actions, customs or beliefs; MACOS classes on the average tended to choose more positive reactions than non-MACOS classes. There was, however, non-homogeneity of regression for the two groups on this variable with respect to classroom climate. The interaction suggested non-MACOS classes that rated climate lower than the average for all classes gave WWA reactions that were lower than expected, compared to MACOS classes with climate ratings falling in the same range. The relationship (the regression slope) of climate to WWA was significant for non-MACOS classes, not for MACOS classes.

CAPS-1, a measure interpreted to refer to perceived ability of self as problem solver; non-MACOS class means tended to be more positive on the average than the MACOS class means.

Follow-up 2

WWBF, an average rating of reactions to people or groups having unusual beliefs or customs; MACOS classes showed some evidence of checking more positive reactions, on the average, than non-MACOS classes.

3) There is marginal evidence from one set of analyses, but not from another, that there was a slight difference between groups on the average with respect to Skills, an average of summated ratings of certain skills learned in social studies last year that may have been advantageous this year. Non-MACOS classes tended to give a little more positive average ratings.

4) There is consistent lack of evidence in all analyses of differences between MACOS and non-MACOS classes, on the average, with respect to the following outcome variables.

Posttest

STEP, a standardized test of social studies skills and knowledge.

IDT, the Interpretation of Data Test of interpreting and using ethnographic information

SS Ch, the class mean of the number of times social studies was chosen as preferred to 5 other subjects.

WWB, an average rating of reactions to people or groups having unusual customs or beliefs.

CAPS-2, an average measure for the class interpreted as interest in problem solving.

CAPS-3, an average measure for the class interpreted as tolerance of ambiguity in problems.

CAPS-4, an average measure for the class interpreted as perception of self as creative in thinking.

Follow-up 2

SS Ch F, the class mean of the number of times social studies was chosen as preferred to 3 other subjects.

SS, the class mean of ratings of how much social studies was liked this year per[illegible] not in relation to other subjects.

WWAF, an average rating of reactions to unusual customs, actions or beliefs.

WWAPF, an average rating of reactions to two examples of unusual behavior of a hypothetical peer.

WWBPF, an average rating of reactions toward a person having those behaviors.

5) There is consistent evidence that student pretest variables account for the overwhelming proportion of total variance with most measures. Classroom characteristics did not, with one or two exceptions related to CAPS, account for significant increments in the total variance of outcome variables. There is evidence that teacher background attitude characteristics were related to performance of classes on the MACOS questionnaire (AP, posttest, and again on API-4F, Follow-up-2); to reaction on WWA, posttest; and WWAF and WWBF ratings Follow-up 2; and to CAPS-3 (tolerance of ambiguity).

6) There is indication that student ratings of classroom activities or processes may be related to ratings of social studies preferences (SS Ch), WWA, and CAPS-3 (tolerance of ambiguity), posttest; and to Know and Interest ratings in Follow-up 1. There is indication that teacher-based ratings of emphases or activities may be related to student performance with respect to WWA, posttest, and ratings of Interest in Follow-up 1.

7) Ratings of classroom climate by students showed a relationship to posttest performance for SS Ch (preference for social studies), and for the STEP test, and for class average ratings on CAPS-1, ability of self as a problem solver. It was associated in Follow-up 2, with class average ratings of how much social studies was liked (SS).

8) Examination of increments of variance due to process or climate in MACOS and non-MACOS groups of classes separately suggest

a generally stronger relationship of those variables with non-MACOS classes than with MACOS classes with respect to attitude outcome measures.

c. Comparisons of Groups with Respect to Individual Process and Climate Variables

Regression analyses using principal components as predictors indicated relationships of process and climate ratings to certain outcomes. Further analyses were made to determine whether there were differences between MACOS and non-MACOS classes, on the average, with respect to those variables.

A multivariate analysis of variance, with Group as the independent variable and the 8 student based PC's as dependent variables, led to the conclusion that there was a significant difference between groups ($F_{8,93}=2.602$, $p \leq .013$). Results of the univariate tests are summarized in Table III-23. The two variables for which the individual F-test were significant were Climate ($p \leq .003$), and S Proc 2 ($p \leq .002$). S Proc 1 was not significant ($p \leq .086$). The other PC's (Ach, Att 1, Att 2, Class 1, Class 2) have p-values ranging from .2-.9. Results were similar when the analysis was done to include all teacher based PC's as well. The difference between groups was significant ($F_{13,71}=1.965$, $p \leq .037$). The individual variables that were significant were Climate ($p \leq .003$), S Proc 2 ($p \leq .005$), T Proc 1 ($p \leq .026$), TP roc 2 ($p \leq .001$), and T Psy 2 ($p \leq .011$). T Psy 1 was marginal ($p \leq .077$).

On which particular process and climate variables were groups different? An analysis of variance, with Group as the independent variable, was done using each process and climate variable individually as the dependent variable. The results for each variable are summarized in Table III-24. Table III-24 gives the means, standard deviations, sample sizes and differences for each variable.

Table III-23

Means, Standard Deviations, N's Differences Between Means and p-Values of Univariate F-tests (df 1,100) of Differences Between MACOS and Non-MACOS Classes on 8 Student-Based PC's

Principal Component (PC)	MACOS			Non-MACOS			Difference	p
	Mean	SD	N	Mean	SD	N		
Pre-Achievement								
Ach	.07	1.60	55	-.11	1.60	47	.18	.589
Pre Attitude								
Att 1	.11	1.26	55	-.17	1.63	47	.28	.324
Att 2	-.01	1.06	55	.10	1.13	47	-.11	.632
Classroom Characteristics								
Class 1	-.01	1.59	55	-.00	1.60	47	-.01	.976
Class 2	.17	1.21	55	-.17	1.68	47	.34	.237
Student Perceived Processes								
S Proc 1	.26	1.81	55	-.36	1.78	47	.62	.086
S Proc 2	-.49	1.67	55	.49	1.52	47	-.98**	.002
Classroom Climate	.44	1.33	55	-.47	1.66	47	.91**	.003

**$p \leq .01$.

Table III-24

Means, Standard Deviations (SD), N's and Differences and P-Values of F-Tests of Individual Student and Teacher Process and Climate Variables

Source	Variable	MACOS Mean	MACOS SD	MACOS N	Non-MACOS Mean	Non-MACOS SD	Non-MACOS N	Difference[1]	P
1. Students, Process Variables	Tchr Talk	2.56	.33	57	2.70	.34	51	-.14*	.029
	Speed (Pace of class)	2.08	.13	57	2.08	.17	51	.00	.925
	Listen	1.86	.46	57	1.72	.39	51	.14	.098
	Discussion	1.24	.22	57	1.44	.33	51	-.20***	.001
	Stress (Grades)	1.74	.39	57	1.53	.34	51	.21**	.005
	Compare	1.25	.24	57	1.46	.33	51	-.21***	.000
	Joking	1.72	.42	57	1.79	.46	51	-.07	.442
	Memory	3.49	.73	57	3.37	.67	51	.12	.387
	Translation	3.37	.54	57	3.47	.56	51	-.10	.374
	Interpretation	3.74	.51	57	3.77	.53	51	-.03	.789
	Application	3.49	.73	57	3.37	.67	51	.12	.387
	Analysis	3.50	.34	57	3.48	.43	51	.02	.806
	Synthesis	3.67	.70	57	3.20	.54	51	.47***	.000
	Evaluation	1.72	.42	57	1.79	.46	51	-.07	.442
	ODI (Opport. Discuss, Involve)	3.37	.59	57	3.38	.55	51	-.01	.907
	Test/Grade Stress	2.96	.43	57	3.08	.48	51	-.12	.181
2. Teachers, Process Variables	Emphasis on Affect	4.27	.73	56	3.49	1.02	49	.78***	.000
	Memory	2.76	1.02	55	2.80	.91	49	-.04	.842
	Comprehension	2.79	.99	56	3.49	.79	49	-.70***	.000
	Application	4.04	.93	56	3.51	.94	49	.53**	.005
	Analysis	3.63	1.02	56	3.15	.82	48	.48**	.010
	Synthesis	3.86	.86	56	3.27	.89	48	.59***	.001
	Evaluation	3.69	1.05	55	3.40	.84	48	.29	.119
	Indiv (Activities)	1.94	.43	57	2.02	.52	51	-.08	.410
	Group (Activities)	2.37	.45	57	2.26	.58	51	.11	.267
	PM (Perceptual Motor Activities)	1.99	.46	57	2.09	.55	51	-.10	.329
	Total Gp (Activities)	2.27	.43	57	2.19	.54	51	.08	.392

1. *p ≤ .05; ** p ≤ .01; *** p ≤ .001.

Table III-24 Continued

Source	Variable	MACOS Mean	MACOS SD	MACOS N	Non-MACOS Mean	Non-MACOS SD	Non-MACOS N	Difference[1/]	P
3. Students, Climate Variables	Satisfaction	1.89	.35	57	2.13	.34	51	-.24***	.001
	Apathy	2.29	.29	57	2.11	.32	51	.18*	.016
	Difficulty	2.40	.16	57	2.30	.21	51	.10*	.017

1. *p ≤ .05; *** p ≤ .001.

It also gives in each case the p-value of the F-statistic. The variables are listed according to source, students or teachers. In the case of student variables, the "raw score" was the class average. In case of teachers, the "raw score" was the teacher's rating, or average scale values for variables such as Individual, Group, PM and Total Group, in which sets of ratings were summed and averaged. For cases in which there were two or more teachers in a class, ratings were averaged.

Note that these analyses were based on all classes in the study. In the case of the teacher based ratings, there were a few omissions on some individual scales in both groups.[36]

Table III-24 shows the following significant differences between groups:

Student Perceptions

- MACOS classes, on the average, rated the teacher as talking more of the time than did non-MACOS classes (Tch Talk).
- MACOS classes were rated as having more discussion than were non-MACOS classes (Discussion)
- MACOS classes were rated as having less concern with grades than were non-MACOS classes (Stress, Grades)
- MACOS classes were rated as doing more comparing of things to find out how they are alike or different than were non-MACOS classes, on the average (Compare).
- MACOS classes, on the average, were rated lower on synthesis (making up new things from what was learned, such as stories, pictures, poems, plays, reports, etc.) than were non-MACOS classes (Synthesis)
- MACOS classes, on the average, were rated more positively on the classroom climate Satisfaction scale than were non-MACOS (Satisfaction)

36. Table III-7 listed the variables discussed here under S Proc 1, S Proc 2, T Proc 1, T Proc 2 and Climate. It gave the polarity of the variables and their correlations with the principal components.

- MACOS classes, on the average, were rated more favorably on the classroom climate apathy scale than were non-MACOS classes (Apathy)

- MACOS classes, on the average, were rated as less difficult than were non-MACOS classes (Difficulty)

Teacher Ratings

- MACOS teachers rated their curricula as emphasizing affective content (values, attitudes, emotions) to a greater extent, on the average, than non-MACOS teachers rated theirs (Affect).

- Non-MACOS teachers rated their curricula, on the average, as specifically aiming to develop student achievement on the comprehension level (understanding what is being communicated but not necessarily relating to other things) to a greater extent than did MACOS teachers (Comprehension)

- MACOS teachers rated their curricula as specifically aiming to develop student achievement on the application, analysis and synthesis levels to a greater extent, on the average, than non-MACOS teachers rated theirs (Application, Analysis, Synthesis).[37]

The results based on students' perceptions were consistent, in general, with information obtained from interviews with students and tape recordings of classes (presented in Section V of this report), particularly with respect to the lack of differences between groups on the scales indicating emphasis on different cognitive levels of activities (Memory, Translation, Interpretation, etc.) There was variation among classes within each group on a number of dimensions having to do with teaching strategies and student activities. But often similarity was observed in prevalence of activities reported between groups.

37. Definitions for these were: Application - using knowledge, methods, theories, etc., in new situations (to solve problems, for example); Analysis-breaking something down into its component parts and understanding the interrelationships of the parts; Synthesis, - putting things together to form a new entity such as a new idea, new hypothesis, or set of relationships.

There are two points of particular interest in the student-based ratings. Discussion was one of the two items making up the ODI scale; Stress (Grades) was one of the two items making up the Test/Grade Stress scale.[38] Taken individually, they discriminated between the two groups of classes; combined with the other item, the scales did not.

The differences between groups, on the average, with respect to the climate variables are noteworthy. It should be remembered, however, that those three variables produced only one significant principal component (PC), and it accounted for 82% of the variance among the three scales. Thus, the three scales, with the 5th and 6th grade students, may have reflected a generalized attitude toward their social studies classes. That in itself would still be a meaningful result. It is stated here as a caution against overinterpreting the implications of each variable separately.

The analyses of the individual process and climate variables, and of the PC's, support the conclusion that the MACOS classes were perceived or rated differently on a number of variables than were the non-MACOS classes, on the average. It should not be concluded or inferred, however, that all MACOS classes were rated absolutely higher (or lower) than all non-MACOS classes on any variable. There were many classes in each group that were rated higher (or lower) than many classes in the other group.

38. The items were included separately in these analyses, along with the scales, since they were considered especially important characteristics per se. The zero order correlation of Discussion with ODI is .56. The correlation of Stress (Grades) with Test/Grade Stress is .87.

Similarly, no inference can be drawn about any particular curriculum in the non-MACOS group. Statements cannot be made from the data presented here about the Taba curriculum, Holt Data Bank, the Harcourt, Brace, Jovanovich program, or others that were included in the non-MACOS group. Statements pertain only to the non-MACOS classes as a conglomerate.

Analyses of variance, with Group as the independent variable, were also done using four scales from Follow-up 2. The question asked was: "How does THIS YEAR'S social studies class compare to LAST YEAR'S social studies class on each of the following?

. Amount of reading out loud in class

. Amount of class discussion

. Amount of work you do alone

. Amount of art work, drawing, making things,[39]

Students rated those items on a 5-point scale (1 = a lot more; 2 = more; 3 = about the same; 4 = less; 5 = a lot less). Results were analyzed using class means and are shown in Table III-25.

39. These items had been selected for use in Follow-up 2 from a longer list used in Follow-up 1 because there had appeared to be differences between groups on them when analyses were based on responses of students as the unit of analysis. It was of interest to determine how those activities would be assessed by students after having had their present program for a year. In the present (FU-2) analyses, the units are class averages for each scale.

Table III-25
Means, Standard Deviations, Differences and P-Values
For Comparisons of Activities

	MACOS			Non-MACOS				
Activity	Mean	SD	N	Mean	SD	N	Diff	P
Reading	3.46	.67	56	3.42	.81	48	.04	.773
Discussion	2.85	.59	56	2.50	.60	48	.35**	.004
Work Alone	2.28	.41	56	2.36	.54	48	-.08	.388
Art, etc	3.53	.68	56	3.47	.67	48	.06	.685

**p≤.01

On the average, former MACOS classes, a year later, rated their current programs as involving less class discussion, compared to last year's class, than did the former non-MACOS classes. It may be seen, however, that with respect to absolute scale values, both groups means tended to fall to the left of the mid-point; that is, in the direction of saying more this year than last. Thus, a more precise statement would be that former non-MACOS classes, on the average, rated their current programs significantly more in the direction of having more class discussion, compared to last year's class, than did former MACOS classes. Although the data are not shown here, it was found that 6th graders (current 7th graders) were more likely to rate their current classes in the direction of less discussion than 5th graders (current 6th graders).

This finding was interesting in light of the tendency for MACOS students, on the average, to have rated their classes higher in amount of discussion than non-MACOS students during the MACOS year.

d. Classroom Climate Viewed as an Outcome

It was seen above (Table III-24) that there were significant differences between the two groups of classes with respect to classroom climate variables (Satisfaction, Apathy, Difficulty). Based on students' average ratings, MACOS classes were higher on the average in satisfaction, and lower in apathy and difficulty, compared to non-MACOS classes. It has also been seen in main analyses involving input, process and outcome variables, that classroom climate has been treated as a process, or predictor, variable.

Climate characteristics could also be viewed as outcomes related to prior variables. To examine this relationship, a fixed order regression analysis of each variable was made, using student and teacher PC's as predictors. The order of entry of sets of PC's followed the order used for all regression analyses. Group was entered last in order to examine its significance when all other variables had been taken into account. The total sample size for these analyses was 85 (46 MACOS and 39 non-MACOS classes).

Results are summarized in Table III-26. It can be seen that for satisfaction and apathy, the sets of variables that added significantly to the multiple R^2 were teacher characteristics and student process ratings. With respect to difficulty, it was pretest and teacher characteristics that added significantly.

Group did not add significantly after all other variables had been removed. Examination was made of when the partial correlation of Group with climate variables ceased to be significant if Group were to be entered into the equation next.

Table III-26
Increments of Variance of Climate Variables Associated With PC's (N=85)[1/]

Variable	Total Multiple R^2	Increment due to Pretests	Increment due to Class Characteristics	Increment due to Tchr Characteristics	Increment due to Tchr Process	Increment due to Student Process	Increment due to Group
Satisfaction	.36**	.05	.00	.14**	.05	.10**	.03
Apathy	.30*	.06	.00	.12**	.03	.07*	.01
Difficulty	.33**	.09*	.04	.10*	.06	.04	.00

1. Order of entry into analyses: Pretest (Ach, Att 1, Att 2); Classroom Characteristics (Class 1, Class 2); Teacher Characteristics (T Demo, T Psy 1, T Psy 2); Teacher based Process Variables (T Proc 1, T Proc 2); Student-based Process Variables (S Proc 1, S Proc 2). Group was entered last as a dummy variable, with MACOS = 1, non-MACOS = 2. Increments may not add to total multiple R^2 due to rounding.

*$p \leq .05$
**$p \leq .01$

after each preceding set. For satisfaction, the partial correlation for group became non-significant only after all other variables had been entered. For apathy and difficulty, it became non-significant after the teacher characteristics had been entered (T Demo, T Psy 1, T Psy 2). That suggested a stronger relationship of those variables (as a set) to the MACOS and non-MACOS groups with respect to apathy and difficulty than was the case with satisfaction.

The variables whose regression coefficients were significant after all variables had been entered were:

. For Satisfaction

Att 2
S Proc 2

. For Apathy

T Psy 2

. For Difficulty

Att 2

Finally, Table III-27 gives the zero order correlations between the three climate variables and the PC and Group variables.

Considering all these data and taking into account the polarity of the climate variables, the following appear to be relationships of predictors to climate variables viewed as criteria:

. the set of predictors as a whole accounted for significant proportions of variance of each climate variable;

. taken as sets, neither student pretest nor demographic (Class 1, Class 2) characteristics added significant increments to the variance accounted for in satisfaction and apathy; student pretest variables as a set did show a signi-

Table III-27
Correlations of Climate Variables with PC's and Group (N=85)[1]

	Satisfaction	Apathy	Difficulty
1. Ach	.01	.16	.23
2. Att 1	-.04	.17	.18
3. Att 2	-.22*	.18	.19
4. Class 1	.05	.10	.21*
5. Class 2	-.03	.06	.19
6. T Demo	.09	-.05	-.05
7. T Psy 1	-.05	-.01	-.08
8. T Psy 2	.36**	-.36**	-.33**
9. T Proc 1	-.23*	.19	.10
10. T Proc 2	.28**	-.28**	-.32**
11. S Proc 1	.06	-.01	.16
12. S Proc 2	.43**	-.34**	-.32**
13. Group	.33**	-.29**	-.25*

1. The climate variables were scaled such that for Satisfaction, the lower the score, the greater the satisfaction; for Apathy and Difficulty, the higher the score, the less the apathy and difficulty. For a sample of this size, correlations of approximately .21 and .28 are significant at the .05 and .01 levels respectively.

*$p \leq .05$
**$p \leq .01$

ficant increment for the difficulty variable;

- when added after student input variables, teacher characteristics as a set added a significant increment to the variance accounted for in all three climate variables, accounting for 40 to 50% of the remaining variance;

- student process variables, when added as a set after input and teacher process variables, added a significant increment to the variance accounted for in satisfaction and apathy (and accounted for 78-83% of the remaining variance); they did not add a significant increment with the difficulty variable;

- Group, added after all other variables, did not add significantly to what little variance remained for any of the three climate variables;

- when considered after all other predictor variables, Att 2 and S Proc 2 were significant predictors of satisfaction; the more students in a class, on the average, were interested in problem solving, and perceived themselves as potentially creative, and the less the class was perceived as having little discussion and as traditionally oriented, the better the class average score on the satisfaction scale;

- when considered after all other predictors, T Psy 2 was a significant predictor of class average scores on the apathy scale; the less the teacher tended to hold traditional views and to approve controlling behavior (as measured by ES VII Traditionalism and TAW scores), the less indifferent or apathetic the class average ratings by students on the apathy scale;

- when considered after all other predictors, Att 2 was a significant predictor of difficulty; the higher the class scored on interest in problem solving and on being potentially creative thinkers on pretest, the less it perceived the social studies curriculum as difficult;

- the individual teacher characteristic that was always the influential variable[40] when added as part of the set of three (T Demo T Psy 1, T Psy 2) was T Psy 2; climate variables were more

40. As measured by the significance of the F-test of the partial beta coefficient in the equation.

positive when teachers tended to be less traditional, as measured by scores on ESVII Traditionalism and TAW.

. the individual student process variable that was influential as a predictor of satisfaction and difficulty, when added as part of the set of two (S Proc 1, S Proc 2) was S Proc 1; the less the students perceived the class as traditional (emphasis on remembering facts, relatively less discussion, etc.), the more positive the perceived satisfaction and the less difficult the work was perceived to be.

The general conclusion drawn is that in the social studies classes examined here, classroom climate is strongly related to teacher attitudes and to how students perceive the operation of the class. It is also related to how a class of students perceives itself with respect to interest in problem solving and to perception of selves as potentially creative thinkers. It does not appear related to achievement levels or to demographic characteristics of classes. There are significant differences in classroom climate between MACOS and non-MACOS classes as a whole. Those differences are diminished when other factors are held constant.

e. Implementation of courses

Preceding analyses have treated (MACOS and non-MACOS) classes as if they were homogeneous with respect to content and amount of implementation. That was of course not the case for the group of classes labelled non-MACOS, which included a variety of different programs or curricula. It was not the case either with the group of classes called MACOS in this study. They too varied widely with respect to content and amount of implementation. It was apparent to the project staff during visits to classes, and from repeated interviews with teachers and students, that there was much variation among MACOS and non-MACOS classes

not only with respect to characteristics such as those depicted by the process and climate variables analyzed above, but also with respect to content, emphasis, amount of class time, and relationships to other aspects of the school's overall program. The interview material, discussed in Section V of this report, depicts variations among classes within both groups in much detail. In this sub-section, results primarily of quantitative analyses of variations in implementation, and their relationships to outcomes, will be presented.

1) Variations in Implementation

a) What were MACOS courses?

In keeping with the non-experimental, minimally intrusive design of this study, teachers (MACOS and non-MACOS) were not asked to keep logs or records as the school year progressed in order to minimize the impact of the project on what was actually being taught in social studies classes and certainly to avoid suggesting what should be taught and how. In the late spring of 1975 prior to the posttesting, course content forms were mailed to both groups of teachers. MACOS teachers were asked to fill out a detailed questionnaire (MACOS Course Checklist) specifying what units and lessons were covered, what length of time was spent on a given unit, what MACOS materials were used, and what supplementary materials and activities were included in the MACOS program. A second form, called "For Classes That Had MACOS and Other Programs," was also filled out by MACOS teachers. That form asked about other curricula, units, and materials and about approximate lengths of time spent on non-MACOS social studies units. Non-MACOS teachers filled out a similar checklist asking for the same information about their social studies program, although not in lesson detail. Again, these

three forms were collected late in the school year, giving teachers the maximum opportunity to report on what had actually been done in their classrooms during the year. Strong assurances were given to teachers that the checklists, while extensive, were not meant in any way to suggest that there was a preferred course content or time schedule.

Class time spent teaching social studies varied greatly from class to class, but on the whole, there was little difference between MACOS and non-MACOS groups. As Table III-28 shows, the two groups on the average spent about 3+ hours a week (shown in minutes) distributed over four days, on social studies.

Table III-28
Time Spent Teaching Social Studies
in MACOS and Non-MACOS Classes

MACOS Classes	Mean	SD	N
# of minutes per wk	196.0	59.7	53
# of days per wk	4.3	.96	
# of wks MACOS was taught	20.6	7.9	
# of weeks MACOS and other was taught	29.7	6.8	18
Non-MACOS Classes			
# of minutes per wk	186.6	60.3	51
# of days per wk	4.4	.91	
# wks spent on social studies	26.7	7.6	25

As reported on the course checklist, none of the MACOS classes spent more than 30 weeks of the school year on any kind of social studies.

Course content varied widely in MACOS classes and, as would be expected, in non-MACOS groups. In only four school districts was there evidence of a district wide policy for social studies curriculum. Otherwise, school buildings

and individual teachers appeared free to develop their own course direction, constrained primarily by budgetary considerations, state requirements and their own interests and aptitudes.

Variation in sequencing of units and lessons taught in MACOS was practically non-existent. That is, teachers did not vary the order of lessons within units from the sequence set by the developers. Nor did they vary the order of units (e.g., they did not go from the unit on baboons to the unit on herring gulls). Omission of lessons within the sequence, however, and in many cases, of entire units, was the rule rather than the exception. On the 55 classes for which there is information on MACOS implementation, only 14 teachers reported covering 100% of the 26 lessons on Man and Other Animals, and only 2 teachers covered 100% of the 36 Netsilik lessons. Thirty three classes were reported as teaching 90% or more of the Man and Other Animals lessons and 16 classes reported the same (90%) for the Netsilik. The total percentage of all lessons covered by the time of posttesting ranged from 16 to 100%.

The MACOS Course Checklist did not ask for reasons for excluding certain units and/or lessons, but a tally of omitted lessons did not reveal any particular pattern of omission. Even those lessons which might have been considered sensitive, such as the mating of herring gulls in the filmstrip "Herring Gulls", or the story, "Birth of a Hunter", or "Old Kigtak" (leaving an old woman to die on the ice), were omitted no more frequently than other non-controversial materials, such as "Chimpanzees: A Primate Contrast" or "How Netsilik Tools are Used". Three of the 55 MACOS teachers did report, however, that they had left out a sensitive lesson or booklet because of "...local publicity." In a small

number of other reported cases (3 or 4), principals and teachers reported making decisions as to whether to include possibly controversial materials and in all cases decided to do so.[41]

With respect to Man: A Course of Study, information from the supplementary form suggests three main types of curriculum variations:

MACOS as the primary curriculum, supplemented by MACOS related units.

MACOS and another social studies curriculum taught currently.

MACOS used to supplement other texts or program packages.

A small number of teachers (18%) reported supplementing MACOS units with MACOS-related activities or programs such as teacher-developed lessons on ecology, dissecting fish, map and graphing skills, simulation games, units on the modern Eskimo, kinship charts, etc. Other special units tied to MACOS dealt with evolution and natural selection, with health and reproduction, sharing and family life, contrasting Eskimo community life with, say, an African or native American tribe. These MACOS teachers mentioned in interviews that in their experience MACOS was uniquely suited to stimulate explorations in a variety of other subjects, for the most part topical or contemporary rather than historical.

Teachers using most of the MACOS units and lessons in tandem with another text, program package, or teacher developed materials, constituted the majority (about 2/3) of those classes called MACOS in this study. Frequently a unit of MACOS would be taught, then a unit on, for instance, the American

41. Further information about the handling of sensitive or controversial subjects, obtained from interviews with teachers, is reported and discussed in Section V.

Revolution or South America, or state and local history. Some classes started with the Eskimo unit and went on to other programs; some stopped at Eskimos and took up lessons on careers or, for example, SRA social studies materials. In all, 29 different texts and programs other than MACOS were mentioned explicitly by this group.

In the third type or category of classes, MACOS was used to supplement the primary curriculum - whatever it might have been. This group spent as little as 9 weeks and as much as 25 weeks on MACOS. Often only the Baboon unit, or the Salmon or Herring Gull units would be taught. Several teachers reported that they found the entire animal section too long; that students lost interest in the subject. And from this experience, they had planned from the start to cover only parts of the MACOS course (unbeknowst to the project at the outset). Geography, U.S. and World History, ancient myths, current events, Glasser's discussions, teacher and student devised projects on community awareness, family life, understanding onesself are but a small sample of the topics dealt with in these social studies classes.

In all, 51 different texts, kits and teacher devised programs were mentioned as the main tools for teaching social studies. Holt Data Bank, Concepts and Values (Harcourt, Brace, Jovanovich), Taba, and Silver Burdette's The Changing World; North and South America accounted for 50% of the programs used. Fewer non-MACOS than MACOS teachers reported using teacher devised materials. The most frequently mentioned supplementary materials were films and map skills lessons. Less frequently mentioned topics and materials were

current events, research techniques, group process, T.V., ecology and the Bicentennial.

2) Relationships of implementation to outcomes in MACOS classes

Three variables were selected as measures of implementation of MACOS in order to examine relationships of implementation to outcomes:

- .time, an estimate of the number of hours spent teaching MACOS, based on data reported by the teachers;
- .percentage of Man and Other Animals lessons taught by the time of posttesting (May 1975), as reported by the teacher;
- .percentage of Netsilik lessons taught by the time of posttesting, as reported by the teacher.

The percentage of lessons taught were based on the total number of lessons available in the Man and Other Animals part of the curriculum, or in the Netsilik part. Time was estimated in hours by taking the number of minutes of social studies per week and multiplying by the total number of weeks the teacher reported spending on lessons covered to date. The number was divided by 60 to give total hours.

A canonical correlation analysis was made of relationships between the implementation variables as predictors, and posttest, Follow-up 1, and Follow-up 2 sets of variables as criteria. Only the first canonical correlation for the set of posttest variables was significant ($R_c = .64$, $p < .02$, $N = 55$). Table III-29 gives the correlations of the variables in each set with the canonical variate.

Table III-29
Correlations of Implementation and Posttest Outcome Variables With Their Canonical Variates

Implementation Variables	Correlations with Predictor Canonical Variate
Time	.79
% Man and Other Animals Lessons	.83
%Netsilik Lessons	.68
Posttest Outcome Variables	**Correlations with Criterion Canonical Variate**
AP (Animals and People)	.19
STEP	-.01
IDT	.18
SS Ch	.37
WWA	.32
WWB	.31
CAPS-1	.02
CAPS-2	.45
CAPS-3	-.23
CAPS-4	-.15

(R_c = .64, $p \leq .02$)

Table III-29 shows that the three implementation variables all correlate highly with their canonical variate. On the outcome side, attitude variables correlate more strongly with the posttest canonical variate, compared to the achievement variables, SS Ch (Social Studies Choices), WWA (What Would You Think, part A, opinions about customs or beliefs), WWB (What Would You Think, part B, opinions about persons who would have such customs or beliefs); and CAPS-2 (interest in problem solving) all correlate more strongly with the variate than do AP (the MACOS course content test), STEP, or IDT (the Interpretation of Data Test).

These analyses used the total Animals and People (AP) score as one of the posttest and follow-up outcomes without taking into account pretests. Further

analyses were made of the relationship of amount of implementation of lessons in the Man and Other Animals and the Netsilik sections of the course to the two parts of the Animals and People test related to those sections; namely, AP1-4 and AP5-8. For these course content specific analyses, a different implementation variable was substituted for amount of time. MACOS teachers had been asked at the end of the year (at posttest) to look at the Questionnaire About Animals and People (AP) and to check any items they may have used with their students during the year for teaching or testing purposes.[42] Classes were coded according to whether the teachers indicated any use of some or all items during the year or no use was made of them. Of 54 MACOS classes for which there were data, about 1/3 had used some or all of them.

A multiple regression analysis was made for each of the following outcome variables:

. Animals and People, questions 1-4, (AP1-4), posttest;

. Animals and People, questions 5-8, (AP5-8), posttest

. Animals and People, questions 1-4 (AP1-4F), Follow-up 2.

The predictor variables in each case were:

. AP1-4 pretest;

. AP5-8, pretest;

. Percentage of Man and Animals lessons taught (relates to AP1-4);

. Percentage of Netsilik lessons taught (relates to AP5-8);

42. Many of the items in AP are to be found in the MACOS booklet, Evaluation Strategies, which is part of the total package. Teachers sometimes use items from it to introduce a unit or lesson, sometimes to evaluate students at the end of a unit. Some teachers do not use them at all. In order not to interfere with normal implementation of the course, no instruction was given to teachers during the year, about whether or not to use them.

. Items used (used = 1, did not use = 2)

. Percentage of 5th grade students in the class

The unit of analysis was, as with all preceding analyses, the class mean. Results of these analyses are given in Table III-30, which shows, for each outcome variable, the total proportion of variance accounted for (total multiple R^2) by the six predictors, and the standardized regression coefficients (Betas) that make up the final regression equation. Indication is also given of whether the regression coefficient is significantly different from zero. The significance test associated with each variable is the measure of whether the variable would add a significant increment of variance accounted for if it were entered into the analysis last after all other variables. It is thus a test of the significance of the standardized partial regression coefficient.

The results given in Table III-30 indicate that for the Man and Animals part of AP, (AP1-4), pretest and age (% 5th grade students) are the particularly important predictors of posttest and Follow-up 2 performance. On the other hand, for the Netsilik part of the instrument (AP5-8), the percentage of Netsilik lessons taught, along with pretest scores on AP1-4 were the more important variables. Age (% 5th graders) was less important in the total equation, and not significant if added after all other predictors. The Beta in this case was not significantly different from zero. Pretest class means for AP5-8 were not significant predictors of AP5-8 posttest scores. None of the implementation variables, with the one expection noted, was an important predictor of posttest or follow-up scores.

These results do not mean that amount of implementation of MACOS

Table III-30

Total Multiple R^2, Standardized Regression Coefficients (Beta) for Regression of AP Posttest and Follow-up 2 Scores on Pretest, Implementation and Grade Level Predictors Variables

Outcome Variable	Total Multiple R^2	AP1-4 Pre Beta	AP5-8 Pre Beta	%Man/Animal Lessons Beta	%Netsilik Lessons Beta	Items Used Beta	%5th Graders Beta
AP1-4 Posttest	.53**	.55**	.06	-.08	.07	-.11	-.26*
AP5-8 Posttest	.54**	.53**	.11	-.11	.25*	-.14	.17
AP1-4, Follow-up 2	.57**	.46**	.14	-.10	-.09	.01	-.31**

*p ≤ .05
**p ≤ .01

lessons is of no consequence. The Animals and People (AP) instruments provided the measure that most strongly and consistently differentiated between MACOS and non-MACOS classes. That is, the two groups differed reliably and significantly in performance on the course specific instrument whether or not adjustments were made for student, teacher, process and climate characteristics. The results do suggest that within the group of MACOS classes, pretest on the Man and Animalspart of AP (i.e., AP questions 1-4), and the age or grade level of students appear to be more important factors than just how many Man and Animals lessons were taught, or whether the teacher used any or all of the items at all during the year. They also suggest, however, that with respect to the Netsilik part of the test (AP5-8) it did matter what percentage of the units and lessons had been taught at the time students were posttested.

One reason that the AP2-4 pretest was an important predictor of class average posttest scores for AP5-8 as well as for AP1-4, while AP 5-8 pretest was not, may be that questions 1-4 are more like a general achievement test in format (and, possible content) than questions 5-8, and if so, students who could handle such formats were apt to be able to handle the less conventional formats of questions 5-8. One test of that hypothesis would be to examine the difference in correlations of AP1-4 and AP5-8 class average scores on pretest with the STEP class average pretest scores. If the correlations are significantly different, and if the AP1-4 pretest class averages correlate more highly, there would be support for the hypothesis. The correlation coefficients for the pretests for the total group of classes (MACOS and non-MACOS) were used, on the assumption

that there was no reason to regard MACOS and non-MACOS classes as different with respect to the AP pretest. The correlations were:

STEP and AP1-4: $r = .771$ ($N = 106$)

STEP and AP5-8: $r = .645$ ($N = 106$)

AP1-4 and AP5-8: $r = .582$ ($N = 108$)

The t-test for the significance of difference between dependent correlations was significant ($t = 2.372$, $df = 103$, $p \leq .05$).[43] It was concluded that AP1-4 class average scores correlate significantly more highly with STEP pretest class average scores than AP5-8 class average scores.

It is also possible, of course, that average reading level of the class was a factor. Therefore, a similar test was made of the correlations of AP1-4 and AP5-8 pretest averages and class average reading level (measured on a 5 point scale). The correlations were:[44]

Reading and AP1-4: $r = -.525$ ($N = 107$)

Reading and AP 5-8: $r = -.448$ ($N = 107$)

AP1-4 and AP5-8 pretest class scores did not correlate significantly differently with class average reading level ($t = -.354$, $df = 103$, $p > .05$). Thus, there appears to be some support for the hypothesis that ability of a group to handle general achievement test types of items may be more influential, even in the AP5-8 questions, than pretest performance per se, and that AP1-4 is tapping

43. The computational formula is that given by Cohen and Cohen. See: Cohen, Jacob, and Cohen, Patricia, Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. Hillsdale, N.J.: Lawrence Erlbaum Assoc., Publishers, 1975, page 53.

44. The reading variable was scaled such that the higher the scale value, the lower the class average reading level. The correlation of reading level with STEP pre was -.693.

that ability. Even if the correlations of the three variables have class average reading level partialled out, the differences in partial correlations of AP1-4 and AP5-8 class average pretest scores with class average STEP pretest scores are still significant ($t = 2.091$, $df = 103$, $p \leq .05$).[45]

From the preceeding analyses, it is concluded that:

- although the Man and Animals part of AP does differentiate between MACOS and non-MACOS classes (e.g., MACOS classes scored higher on AP1-4, on the average, than non-MACOS classes in Follow-up 2); it has more of the characteristics of a standardized achievement test (the STEP test) than the Netsilik part of the test (AP5-8);

- pretest class average scores and percentage of 5th grade students (essentially, the average age of the class) are more important predictors of posttest and Follow-up 2 class average scores for AP1-4 than the percentage of Man and Animals lessons completed within the group of MACOS classes;

- how a class scored on AP1-4 pretest and the percentage of Netsilik lessons completed are the more important predictors of posttest performance on the Netsilik questions (AP5-8)

- whether MACOS teachers used some or all of the AP items with their classes during the year has little influence on posttest on either part of the instrument or on follow-up class average scores.

45. The partial correlation coefficients are:
STEP and AP1-4: .62
STEP and AP5-8: .46
AP1-4 and AP5-8: .44

3) Item Responses for the Questionnaire About Animals and People (AP)[46]

The preceding analyses have focussed on relationships of amount of implementation to performance on the MACOS content questionnaire. Since that instrument, of all those used, consistently differentiated between MACOS and non-MACOS classes, as would be hoped, it is of interest to examine item responses made by MACOS and non-MACOS students. In some instances it is evident where MACOS students in particular did and did not end with a clear grasp of some parts of the course as reflected in the items.

For each item and choice, tables showing the percentages of MACOS and non-MACOS students (the student as the unit of counting, not the class) choosing each alternative in the posttest of the questionnaire were prepared. These item analyses are given in Table III-31, which shows each question and item, and the percentages of associated responses. In examining the items and item responses, it must be remembered that: 1) the individual responses came from groups of students who had been in the same classes; 2) as shown above, there was a range of implementation of both units among MACOS classes, and 3) the item distributions combine responses from 5th and 6th graders. With those qualifications, it is believed nevertheless that two further statistics will aid in examining items. One is the p-value of the Chi-square test of the significance of the difference between groups of the total response distributions for each item. The other statistic is the asymmetric lambda A (λA), with group (MACOS, non-MACOS) as the dependent variable. This statistic, as described

46. The item response distributions pertain only to students who responded; they do not include percentages of students who omitted the item.

earlier, is an aid in maintaining a perspective on the strength of predictive association between differences in response distributions between the two groups. It is the percent reduction in error of prediction of group membership, given responses to an item. One may find, for example, a significant chi-square, but little predictive association. Lambda A (λA) thus provides a guide to interpreting Chi-square p-values.

Sample sizes are not included in the tables to save space. They vary in each group, depending on the specific item. The following are the mean sample sizes, standard deviations, and actual maximum and minimum sample sizes for each group:

	M	SD	Max	Min.
MACOS	606.13	11.42	621	577
Non-MACOS	547.93	13.86	570	524

On occasion, comparisons will be made with results obtained in the formative evaluation of MACOS.[47] Item analyses of posttest (winter 1968) results were also made in that evaluation. In some cases results for MACOS students are similar to those obtained in the present study; in some cases they are quite different. There are several possible reasons for differences. One of course is that in some cases item wording and formatting were different in this study. Another is that the demographic characteristics of students (especially % 5th grade and % 6th grade) were different in the 1968 EDC formative evaluation study than in this one. The EDC study had more 5th graders, fewer th graders; it also had some 4th graders. Yet another possible reason is that programs were implemented

47. Hanley, Janet P., et. al., Curiosity/Competence/Community: An Evaluation of Man:A Course of Study. Social Studies Curriculum Program, Education Development Center, Inc., Cambridge, Mass., 1970, Vol. I, Section III.

differently in the two studies. This report has already shown that there were wide variations in implementation of the curriculum. The EDC report notes variations in implementation also. However, it does not give indication of the extent of variation found in the present study. Another possible reason for differences for the Man and Other Animals part of AP is that the EDC posttest was given at the end at the end of the unit (Winter, 1968); in this study it was given near the end of the year (May,1975), typically a semester after the unit. There are, of course, other possible reasons for differences in results.

1) Question 1

This question was described by EDC as related to the "ability to make comparisons and distinctions between man and other animals."[48] Five of the items (a, b, c, d, e - the latter modified in wording) are from the original instrument. The last four are not. Of the original five, the first two differentiate between MACOS and non-MACOS groups, although the indicies of predictive association are small. The percentages of correct responses of the MACOS students for item a (marry) are a little lower in the present study, compared to the EDC 1968 formative evaluation (80% in the present study compared to 90% for EDC); they are substantially lower for item b, (grow up without adult care: 52% compared to 81%). Item c (use a language) is probably the most interesting item in the whole set. The MACOS materials emphasize the fundamental difference between animal communication and human language. The course stresses language as one of the major shaping forces of human beings. Here there was no significant differentiation between groups in response distributions. There was a marked difference

48. Ibid., p. III-1.

Table 31
A Questionnaire About Animals and People

1. In the list below are 9 things that happen during the lifetimes of human beings or other animals. Some things happen just to human beings; some happen to other animals but not to human beings. On the line beside each question, write a

1 if it is true only for human beings

2 if it is true only for some other animals

3 if it is true for both human beings and some other animals

		1	2	3	X^2_p 1/	λ_A 2/
____a. marry	MACOS	80.0*%	1.6%	18.4%	.029	2
	Non-MACOS	77.5*	.4	22.2		
____b. grow up without adult care	MACOS	3.9	51.5*	44.6	.033	3
	Non-MACOS	3.0	44.9*	52.1		
____c. use a language	MACOS	36.6*	4.2	59.2	.287	1
	Non-MACOS	33.0*	5.6	61.3		
____d. protect themselves from enemies	MACOS	3.1	5.6	91.3*	.306	1
	Non-MACOS	3.5	7.7	88.7*		
____e. cooperate with each other	MACOS	16.1	5.2	78.7*	.722	0
	Non-MACOS	17.4	5.9	76.8*		
____f. have belief systems	MACOS	64.3*	7.5	28.2	.435	0
	Non-MACOS	60.7*	8.6	30.6		
____g. throw things	MACOS	32.0	8.8	59.2*	.007	6
	Non-MACOS	40.9	7.3	51.9*		
____h. have a social organization	MACOS	62.8	6.7	30.5*	.625	0
	Non-MACOS	64.1	7.7	28.3*		
____i. make symbols	MACOS	50.6*	8.7	40.8	.030	2
	Non-MACOS	57.1*	9.7	33.8		

*= answer scored as correct.
1. p-value of Chi-square.
2. Lambda A: index of predictive association of group from responses.

Table III-31 Continued

2. During their lifetimes, animals learn to do many things. They are able to do other things without learning. Read each sentence below. Then check one answer for each sentence showing whether or not you think the animal must learn the behavior or can do it without learning.

	MACOS	Non-MACOS	X^2_p 1/	λ_A 2/
Herring Gull (A Bird)				
a. find the edges of its territory				
* 1. Learns to find them	68.9%	62.2%	.018	4
2. Does not have to learn to find them	31.1	37.8		
b. crouch when in danger				
1. Learns to do it	37.3	44.4	.015	4
* 2. Does not have to learn to do it	62.7	55.6		
c. Recognize its chicks by spots on the head				
* 1. Learns to recognize them	48.6	51.4	.925	0
2. Does not have to learn to recognize them	51.0	49.0		
Baboon (A Mammal)				
d. know the alarm calls of other animals				
* 1. Learns to know them	58.2	67.1	.002	4
2. Does not have to learn to know them	41.8	32.9		
e. make sounds				
1. Learns to make them	30.9	38.4	.008	5
* 2. Does not have to learn to make them	69.1	61.6		
f. know its place in the dominance order of the troop				
* 1. Learns it	72.3	69.4	.290	0
2. Does not have to learn it	27.7	30.6		

* = answer scored correct.

1. p-value of Chi-square.
2. Lambda A: index of predictive association of group from responses.

Table III- 31 Continued

3. Beside each sentence in Column A, write in the number of the sentence from Column B which you think is the best answer. You should use an answer only once.

A

___a. a description of a structure

___b. a description of a function

___c. an example of dominance

___d. an example of learned behavior

B

1. The water is cold.
2. The puppy moved back from the food bowl when the adult male came toward it.
3. The ice kept the lemonade cold.
4. Herring gulls parents recognize their chicks by the spots on their heads.
5. She is pretty.
6. A tree has roots, a trunk, branches and leaves.

	Answers							
Questions Group	1	2	3	4	5	6	X^2p 1/	λA 2/
a. MACOS	10.7%	5.3	7.8	8.7	14.5	53.0*	.056	4
Non-MACOS	14.5%	7.2	8.7	7.7	9.6	52.4*		
b. MACOS	11.7	15.0	38.8*	14.8	6.8	12.9	.022	7
Non-MACOS	9.3	20.6	33.1*	19.3	6.3	11.4		
c. MACOS	7.8	53.3*	10.6	14.0	10.2	4.1	.000	29
Non-MACOS	11.0	18.6*	19.6	26.6	16.5	7.6		
d. MACOS	8.1	20.9	8.2	48.2*	8.2	6.3	.000	22
Non-MACOS	5.7	43.8	5.9	27.8*	8.8	8.0		

* = answer scored as correct.

1. p-value of Chi-square.
2. Lambda A: index of predictive association of group from responses.

Table III-31 Continued

4. On the line beside each word, write the number of the definition that best describes that word. You should use an answer only once.

____a. life cycle	1. a young human or other young animal
____b. offspring	2. a mammal and a primate
____c. juvenile	3. giving birth to young
____d. reproduction	4. the young of any animal
____e. human being	5. the pattern of being born, having babies, dying
	6. a jump to one side
	7. a delinquent or bad teenager
	8. the opposite of animal

	Answers									
Questions Group	1	2	3	4	5	6	7	8	X^2 p 1/	λ_A 2/
a. MACOS	9.3%	4.5	6.2	2.2	73.9*	1.7	.8	1.3	.001	10
Non-MACOS	14.9%	4.7	10.7	3.2	62.0*	2.4	1.3	.8		
b. MACOS	15.0	7.4	24.1	41.5*	4.1	5.0	1.7	1.2	.000	14
Non-MACOS	16.5	10.3	15.3	33.9*	3.8	15.7	1.7	2.7		
c. MACOS	36.3*	6.0	2.0	14.9	3.9	2.0	33.0	1.8	.000	27
Non-MACOS	13.3*	6.9	4.8	6.9	3.3	4.6	58.5	1.7		
d. MACOS	6.4	9.4	57.5*	7.8	6.6	5.2	3.5	3.6	.000	10
Non-MACOS	5.9	5.9	50.8*	7.8	15.8	5.3	4.8	3.6		
e. MACOS	5.9	14.9*	3.1	2.2	3.2	3.1	5.6	62.0	.106	2
Non-MACOS	8.6	15.9*	3.6	1.9	3.6	1.9	2.4	62.2		

* = answer scored as correct.

1. p-value of Chi-square.
2. Lambda A: index of predictive association of group from responses.

Table III-31 Continued

5. Several statements about the Netsilik Eskimos are listed below. What do you think each of the following things depend upon? You may decide that more than one answer could be given, but check only one answer for each statement. If you aren't sure, make a guess.

		Depends upon rules of Netsilik society	Depends upon what a person prefers	Depends upon what a person is able to do	X^2_p [1]	λ_A [2]
1. the kind of songs the Netsilik Eskimos make up	MACOS	33.4%	47.0*	19.6	.065	0
	Non-MACOS	27.1%	51.3*	21.6		
2. the Netsiliks' use of magic and other beliefs	MACOS	66.0*	17.4	16.6	.005	5
	Non-MACOS	56.6*	22.7	20.7		
3. the activities connected with the birth of a Netsilik baby	MACOS	51.1*	23.2	25.8	.000	8
	Non-MACOS	37.2*	29.7	33.0		
4. the friends Netsilik children make	MACOS	18.0	66.4*	15.5	.180	2
	Non-MACOS	19.5	61.5*	19.0		
5. choosing a song partner	MACOS	16.9	66.6*	17.3	.487	0
	Non-MACOS	16.1	63.4*	19.7		
6. surviving through a hard Arctic winter	MACOS	14.4*	16.4	66.4	.122	3
	Non-MACOS	22.1*	16.4	61.6		
7. the Netsilik Eskimos who act as leaders	MACOS	50.0	16.4	33.6*	.011	2
	Non-MACOS	53.7	20.5	25.8*		

* =answer scored as correct.

1. p-value of Chi-square.
2. Lambda A: index of predictive association of group from responses.

6. If a group of Netsilik Eskimos came to visit us, some things about our lives would seem familiar to them. Other things would seem different and strange or unfamiliar.

For each of the phrases below check whether you think the Eskimos would find it familiar and similar to their way of life, or different and unfamiliar to them.

	MACOS	Non-MACOS	X^2_p [1]	λ_A [2]
a. the fact that we use words to express our feelings and ideas				
* 1. similar, familiar	64.2%	62.9%	.691	0
2. different, unfamiliar	35.8	37.1		
b. the fact that different rooms in our houses have different functions (for example, we use one room for cooking, another room for sleeping)				
1. similar, familiar	28.2	29.2	.743	0
* 2. different, unfamiliar	71.8	70.8		
c. the way we feel when a friend makes fun of us				
* 1. similar, familiar	67.4	70.0	.365	0
2. different, unfamiliar	32.6	30.0		
d. the way most of us feel about dogs				
1. similiar, familiar	41.8	58.7	.000	17
* 2. different, unfamiliar	58.2	41.3		
e. the fact that our parents tell stories to us when we are young				
* 1. similar, familiar	77.2	72.1	.054	3
2. different, unfamiliar	22.8	27.9		
f. the fact that we often throw away unwanted objects or food				
1. similiar, familiar	25.7	30.7	.065	2
* 2. different, unfamiliar	74.3	69.3		

*answer scored as correct.

1. p-value of Chi-square
2. Lambda A: index of predictive association of group from responses.

7. Both Netsilik Eskimos and wolves hunt the caribou on the tundra. Below is a description of a hunter. Read the description and decide whether the hunter is a wolf, a Netsilik Eskimo or whether it could be either one.

> I sometimes hunt alone, but I often hunt in small groups. I often chase caribou towards others of my kind in hiding. I plan on using different methods for catching caribou depending on whether I hunt the caribou on land or in the water. Some of my kind may die during the year if not enough caribou are killed.

What am I? (Check the one best answer.)

	MACOS	Non-MACOS	X^2_p [1]	λ_A [2]
1. a Netsilik Eskimo	34.7*%	21.0 *%	.000	7.
2. a wolf	18.1	26.9		
3. the hunter could be either a Netsilik Eskimo or a wolf	47.3	52.0		

What was your reason for the answer you chose above? (Check the one best answer)

1. both Netsilik Eskimos and wolves do all the things listed above	39.4%	42.6%	.001	3
2. both Netsilik and wolves cooperate in hunting	21.0	23.4		
3. only man can plan which hunting methods will be most usefui in a particular place	28.3*	18.3*		
4. only wolves hunt in groups and chase caribou toward others of their kind in hiding	11.3	15.7		

*Answer scored as correct.
1. p-value of Chi-square.
2. Lambda A: index of predictive association of group from responses.

Table III-31 Continued

8. The Netsilik do some things mainly to meet needs of daily survival (physical needs). They do other things mainly to help explain and give meaning to their world, to feel more comfortable about life (spiritual needs). Think about whether the following activities meet mainly physical or mainly spiritual needs. Then check either physical or spiritual.

	MACOS	Non-MACOS	X^2_p 1/	λ_A 2/
a. The Eskimo woman covers the eyes of a fish with ashes.				
___ 1. physical	14.2%	24.8%	.000	9
* ___ 2. spiritual	85.8	75.2		
b. At the river camp the men first repair the weir (a trap made of stones to catch fish), which was damaged by winter ice.				
* ___ 1. physical	78.3	75.5	-.293	0
___ 2. spiritual	21.7	24.5		
c. The men are very careful not to repair their tools near the stone weir				
___ 1. physical	27.1	44.1	.000	14
* ___ 2. spiritual	72.9	55.9		
d. A young hunter makes up a beautiful song about the great caribou he is going to catch.				
___ 1. physical	30.8	31.2	.942	0
* ___ 2. spiritual	69.2	68.8		
e. A careful Eskimo always drips fresh water into a seal's mouth after it has been caught.				
___ 1. physical	23.4	40.0	.000	14
* ___ 2. spiritual	76.6	60.0		
f. A man can borrow the wife of his song partner when his own wife is unable to travel with him.				
* ___ 1. physical	57.9	43.4	.000	10
___ 2. spiritual	42.1	56.6		

*Answer scored as correct.

1. p-value of Chi-square.
2. Lambda A: index of predictive association of group from responses.

in percentage of correct responses between the MACOS students in this study and the ones in the EDC formative evaluation (37% in this study compared to 76% in the EDC 1968 national pilot study). MACOS and non-MACOS students in this study chose right and other responses at virtually identical rates. Item i (make symbols) in question 1 was added as a further check on the question of the distinction between language and signalling or communications. The terms 'create' or 'use' were considered in lieu of 'make'. It was believed that 'make', for 10-12 year olds, came closest to the distinctions being drawn between language and communications in MACOS. Here the item did differentiate slightly between MACOS and non-MACOS students, with non-MACOS students having a higher percentage of correct responses.

MACOS students, on the other hand, did appear to be the more knowledgeable about animal behavior and capabilities. Item g (throwing things) differentiated the most strongly of all items in the set between groups, in favor of MACOS students. It is not a specific item of MACOS written content- DeVore's journal makes no reference to it, no teacher's manual makes reference to it. The students' booklets on baboons, chimpanzees, animals of the African Savannah, and others make no reference to it. Possibly MACOS students were more attentive to some animal T.V. programs.

One item in question 1 over which there may be contention about the scoring is item h, 'have a social organization'. Bruner stressed social organization as a major humanizing shaping force. The MACOS student materials do not refer to animals having a social organization. They refer to troops, dominance, offspring, parent dependency, functions of animals in groups according to sex, age, aggresiveness,

and so on. The teacher's booklet on baboons explicitly refers to the social organization of troops of baboons. De Vore's article "Primate Social Life" in Seminars for Teachers (a MACOS teachers' booklet) refers more than once to the "social organization of baboon troops". De Vore, in his article, however, is careful to draw distinctions between human and animal social organization on a number of dimensions. The concept, in the MACOS materials, seems ambiguous. The discriminandum appears to be that human beings can vary their social organization (relationships); animals cannot, once they mature sufficiently to become active and reactive members of a species-specific group. Rebel baboons cannot, one is led to believe, change the fact of dominance structures within baboon troops. Rebel human beings may bring about a sequence of events that fundamentally alters dominance relationships (and others) between inter and intra-societal groups (e.g., the American Revolution). The forces operating are cultural, religious, economic, racial, or whatever, but not genetic.

The belief of this project was that the concept of social organization (a lack of purely individualistic or random behavior) was inherent in MACOS materials to teachers, if not to students. It was for that reason, assuming teachers noted the same thing, accepted it, and taught it, that item h was scored as it was.

2) Question 2

The wording of this question and the response format was modified in an effort to make it easier for poor readers. The change in response format did introduce ambiguity, as was pointed out by some students. For example, the choice "does not have to learn to do it' can mean: 1) is able to do it without learning how to, or

2) has no need to do it. The instructions for the question, however, are not ambiguous. The effect of ambiguity in response choice for the items, if one were uncertain about the concept (learned vs. innate behavior), or the instructions, or both, may be to make responses random. For the sample sizes involved here, the 95% upper and lower confidence limits for accepting the hypothesis that an observed percentage of correct responses is not different from 50% are: 1) for the MACOS group about 45-55%; 2) for the non-MACOS group 44-56%. There is one case in which the distribution for MACOS students fall within those limits (item c), and two cases for the non-MACOS students (items b and c). Otherwise, all distributions for both groups show percentages of correct responses significantly different from chance.

Of the four items for which the Chi-square for the difference between groups was significant at the .05 level, three (a, b and c) show the MACOS students having higher percentages of right answers. Curiously, for item d (baboons, know the alarm calls of other animals), the non-MACOS students had the greater percentage of right answers. The percentages for the MACOS students on posttest for this question are very close to those found on posttest in the EDC study. The formative evaluators noted that, along with the concept of language, applying the concept of learned versus innate behavior by identifying examples of each in herring gull and baboon behaviors was a problem area.[49] While 60-70% of the MACOS students got most of these items right, obviously 30-40% did not. However, a similar range of non-MACOS students also got four out of six of the items right. The average difference between groups in percentage of right answers for the three

49. Ibid., p. III-11.

items in which percentages exceeded chance and MACOS had the higher percentage is about 6%. As one can see in the table, while differences are significant, the strength of predictive association indices (λA) for the items are not very large (4-5% reductions of errors in predicting groups from responses). The course does give MACOS students on the whole an edge over non-MACOS students in ability to identify learned and innate behaviors, but it is not a large edge, as measured by this question.

3) Question 3

This item, appearing in the MACOS Evaluation Strategies booklet, was modified slightly. Response option 2 in Column B was changed from an example of baboon dominance behavior to dominance among dogs in order to test the ability of MACOS students to transfer a concept to another species. The instructions also included a statement to use an answer only once. While the MACOS formative evaluators, in commenting on a matching question, noted that instructions did not make it clear to students that an answer could be used more than once, the MACOS Evaluation Strategies booklet provides a key of single, presumably preferred choices for that question (modified) and for question 3 here. It therefore seemed appropriate that students should be told that.

Overall distributions of responses by the two groups did differentiate between them, more strongly for items b, c, and d than any of the items in the prior two questions. It can be seen, however, in the table accompanying question 3, that with respect to "right" answers, the concepts of structure and function (items a and b), non-MACOS students did about as well as MACOS students. MACOS students clearly had a strong edge over non-MACOS students in recognizing examples of dominance and learned behavior (items c and d). However, within the MACOS

group, the percentages of students correctly identifying the examples are only around 50%. Based on this form of assessment, the data provide further indication that the more complex and abstract concepts were not in general learned effectively.

4) Question 4

This is another matching question, testing use of course vocabulary. Four of the five items (a-d) show strong differentiations between the two groups of students, with MACOS students showing the higher percentages of correct choices. With the exception of item a (life cycle), however, the percentages of correct choices by MACOS students are not high. They are substantially lower than the posttest percentages of right responses found in EDC's formative evaluation. Item e (human being) was answered correctly, as defined by the MACOS scoring key, by only 15% of the MACOS students. They, like non-MACOS students, predominantly chose the definition 'the opposite of animal' for that term rather than 'a mammal and a primate'. The percentages of correct responses for that item in the EDC evaluation was also low (32%), but not as low as found in this study.

5) Question 5

This is the first of three questions on the Netsilik unit of MACOS. It asks students to classify behaviors according to whether they are most likely to depend on: 1) the rules of Netsilik society; 2) what a person prefers; or 3) what a person is able to do. Three of the six items (2, 3, and 7) show a significant differentiation between MACOS and non-MACOS students on posttest. Three do not. In the three cases in which there is a differentiation, MACOS students have the

higher percentages of correct responses, indicating more concrete knowledge (or understanding) of some reasons for certain specific behaviors. The indices of predictive association, however, for those three items were not very large, compared to other questions and items. Nevertheless, they were associated with significant differences between groups. Again it can be noted that the range of MACOS percentages of "right" responses is about 14-67%.

6) Question 6

This question explores understanding of Netsilik culture and its similarities and differences from ours. It asks students to indicate whether they thought that if a group of Netsilik Eskimos came to visit them, they would find certain structures, functions or attitudes similar or familiar to theirs, or different and unfamiliar. In this question, the response format was modified to make it more intelligible to a wider range of students than the original formats. Otherwise, items remain the same.

The two items that significantly differentiate between groups are: 1) the way most of us feel about dogs; and 2) the fact that our parents tell us stories when we are young. The other four items were answered correctly by non-MACOS students as frequently as by MACOS students.[50] None of the distributions in either group fell within the limits of chance.

7) Question 7

These two items were intended to measure the ability of students to reason from evidence based on knowledge gained from the course. There is no

50. It should be noted that Eskimos as a general cultural group were hardly unknown to many non-MACOS students.

question that MACOS students do better with them on posttest than the non-MACOS group, but the percentage of MACOS students giving correct responses to both parts of the question is a minority of the total group (35% to the first part of the question; 28% to the second part).

8) Question 8

This is a set of six items designed to assess students' understanding of the reasons for certain behaviors of the Netsilik Eskimos. It asks students to differentiate between physical and spiritual reasons underlying certain customs and behaviors.

It can be seen that except for the non-MACOS group for item c, proportions of correct responses are different from chance.[51] It may be concluded that: 1) in four of the six items MACOS students answered the question correctly significantly more often than non-MACOS students; 2) on the 4 items (a, c, e, and f), MACOS students got correct responses an average of 17% more often than non-MACOS students; 3) the indices of predictive association range from 9-14%; and 4) the items of question 8 were in all cases answered correctly by the majority of MACOS students.

The analyses of individual questions and item choices suggest several implications about MACOS, its implementation and effects, as measured by the questionnaire and procedures of this study. Students do learn many things about animals, animal behavior, reasons for behavior, Netsilik customs and reasons underlying them; and they do acquire and apply a vocabulary. There is indication as noted above, however, that some of the more difficult concepts of the course,

51. Consider MACOS two-tailed 95% confidence limits to be 45-55% and non-MACOS ones of 44-56%. These confidence bands apply to observed item sample sizes.

such as the distinction between animal communication and human language, are not getting across effectively. For example, question 1, item c, concerning whether animals as well as human beings use a language, not only did not differentiate between MACOS and non-MACOS students, but also was answered correctly by slightly over one-third of the MACOS students. The concepts of structure and function are others that, given this questionnaire as a means of assessing comprehension, were difficult concepts for many students. Further, as noted above in connection with question 4, item e, the majority of MACOS as well as non-MACOS students on posttest identified "human being" as "the opposite of animal". Finally, with respect to question 7, while proportionately more MACOS students than non-MACOS were able to recognize the concept of planning ahead in relation to circumstances as a difference between animal hunters (wolves) and human hunters (Netsilik Eskimos), the majority did not perceive it as it was expressed in the question.

Students in this study appeared to do best with factual material or with concepts for which there were readily observable or describable examples. The concept of dominance is an example, life cycle is another. Similarly, most students grasped the fact that some customs or behaviors of the Netsilik were based on certain beliefs about the world, or on spiritual needs as distinct from necessity (see e.g., question 8.)

While these analyses raise questions about the difficulty level of some of the course concepts and material for 10-12 year olds, several factors must be considered in tempering inferences and generalizations. One has already been stated: there was a wide range of amount of implementation of the course

in the MACOS classes in this study. Some of the variations in performance on particular questions and items may simply reflect variations in implementation: It has already been found that total scores for the Netsilik questions (AP 5-8) were significantly associated with percentage of Netsilik lessons covered (see Table III-30).[52] Second, the form in which questions and responses about abstract concepts was put may work to restrict the ability of students to demonstrate grasp and mastery of complex or abstract concepts. The questionnaire, like all such instruments, puts a heavy reliance on reading ability. Ambiguity of questions may have been a factor influencing choices. Furthermore, this form of assessment of knowledge and skills was found sometimes to be disliked and resisted by students. It was the experience of staff members of this project that eliciting information, descriptions and explanations from students in interview situations was often a more felicitous situation, judging from students' reactions, than administering the various forms, even when the latter were read aloud. Third, the instrument (AP) was administered last to students (following Study Choices - SS Ch, What Would You Think - (WWA and B, and CAPS). The responses may reflect a fatigue effect both in MACOS and non-MACOS groups.

There are other possible reasons that could account for or contribute to the relatively poor posttest performance of MACOS students, on the average, with some of the questions and items dealing with concepts and distinctions in the course. Nevertheless, the trend of results obtained in this study, particularly with respect to concepts such as language, and learned and innate behavior, is

52. MACOS formative evaluators found similar results. Cf. Hanley, Op. cit., p. III-31.

similar to that found in the formative evaluation of MACOS.[53] The consistency of trends raises the question of age appropriateness of some of the materials and subject matter. It has been seen that with respect to the Man and Other Animals questions (AP 1-4), when the class average score is used as the measure, the percentage of 5th graders in the class is a significant predictor of total score (cf. Table III-30). The higher the percentage of 5th graders, the lower the predicted posttest score for the class. That is not the case for the Netsilik part (AP5-8). It is the first part (AP 1-4) of the overall questionnaire that contains the questions on some of the more difficult concepts such as language, learned behavior, structure, function, human being, etc.

The conclusion is not that 5th grade students are incapable of learning such concepts and demonstrating it by being able to answer questions about them in the format used in this study. It is that as the course was implemented in classes in this study, the evidence is that the younger (5th grade) students in particular, and a number of MACOS students in general, were not successfully demonstrating mastery of some of the basic concepts.

The implication is that if MACOS is to be taught to upper elementary students, especially at the 5th grade level, and if understanding of some of the basic concepts as measured by instruments such as that used here (and as appear in the MACOS booklet, Evaluation Strategies), particular attention needs to be paid by teachers to lessons and methods designed to teach those concepts.

---

53. E.g., "One important conceptual area where the course fails to produce significant learning is that of symbolic language as a distinctly human phenomenon, and all other evidence points to this failure." Hanley, Op. cit., p. III-55.

In this regard, it is interesting to note that a recent study by Fitch, et. al., has called into question the adequacy of training provided to MACOS teachers.[54] That study found that teachers who had received the equivalent of a basic course in Anthropology (60 hours) plus 20 hours of training "consistent with normal MACOS-EDC recommended teaching procedures" used anthropological terms or concepts in teaching MACOS more accurately than did teachers who had only the MACOS training.[55] The same applied to the students of the teachers in those training groups. Similarly, the students of teachers who had the anthropology course training as well as the MACOS training did better on the MACOS achievement tests (used by the developers) than did students whose teachers had only the MACOS training. There were other findings, but the implication drawn by the researchers were that: "considering all results together, it appears that accurate use of the terminology of the field by the teacher is a crucial factor for student improvement. Apparently exposure to MACOS material alone is not sufficient to acquaint teachers adequately with the language of anthropology."[56]

Teachers in the present study were not tested (or observed) for knowledge or accuracy of use of anthropological or other disciplinary terms or concepts. Thus, no comment can be made supporting or not supporting the results of Fitch, et. al. The item analyses presented and discussed above do, however, point to problems in effective learning in some areas, as measured by the AP questions. Whether the

54. Fitch, Robert M., Haefner, John H., and Gonzalez, Nancie, Different teacher training and the teaching of "Man: A Course of Study.", Social Education, Vol. 41, No. 3, March, 1977, pgs. 242-246.

55. Ibid., p. 242.

56. Ibid., p. 246.

problems lie in teacher preparation, how much or in what way the course was implemented, the age appropriateness of the course, the means of measuring achievement, or other factors or combinations of factors cannot be determined from the analyses presented here.

4) Other Item Analyses

a. What Would You Think (WWA and B)

The questionnaire What Would You Think was designed to assess attitudes toward behaviors, beliefs or customs, and towards peoples or groups that might do or have them. It has been seen earlier that WWA and B have shown indications (though not consistently) of differentiating between groups on posttest and on Follow-up 2, though not on pretest. The scoring of the instrument was based on the rank ordering of responses from negative to positive by 5th and 6th grade students.

The instrument was intended to pertain to a major goal of MACOS; and as such it is of interest to examine distributions of responses of MACOS and non-MACOS students for each question. In this case, it is particularly of interest to examine pre-post responses, as well as to consider response choices of 5th and 6th grade students. Percentages of students choosing each response on posttest, as well as the percentage and direction of change from pretest are given in Table III-32. Table III-32 also shows for each question and response choice, the rank order of the choice. The lower the rank, the more negative the statement, as seen by samples of upper elementary students.[57] There are occasional ties, indicating that the computed scale values were sufficiently close as to make it not worth trying to distinguish

57. See Section IIB.

between them.[58] The table also gives the unweighted average change for a response from pretest to posttest across the two groups and grade levels. It should be noted that percentages within a group may not add to 100% due to rounding. For the same reason, amounts and directions of changes for all four choices within a group may not add to zero. The table is based on the total number of students responding to each item on both occasions (pre and post). N's are shown for each group for each question and column.

Consider the first question in Table III-32. Students were asked to read the situation described, and then check the statement in Column A that was most like their reaction, and similarly for Column B choices. Thus, for students (M-5), it may be seen that on posttest 18% chose the first item in Column A ("since it is part of their religion, the custom is alright"). The +4 under that indicates that this was an increase of 4% from the percentage of that group choosing that response on pretest (14% chose it then). The column indicates that the group of 5th and 6th grade students on whom the scaling of choices was based tended to rank that statement as the most positive (4) of the four in Column A. Looking across the four groups (M-5, NM-5, M-6, NM-6) for the first item, it can be seen that: 1) there was relatively little variation among groups in percentages of students picking that statement on posttest; 2) in all four groups there was an increase in percentage of students choosing it from pretest to posttest (all directions of change are positive, or increases); and 3) the unweighted average change from pretest to posttest for

58. As noted in Section II, each set of 4 responses was given a value, based on a scale with a mean of 5 and standard deviation of 2. The rank orders were obtained from those scale values.

Table III-32

Student Vs. Posttest Responses, and Changes From Pretest to Questionnaire What Would You Think, By Group and Grade[1/]

1. If you heard about a group of people in the Pacific Islands who each year would throw stones at one of their members until he was dead as part of a religious custom, what would you think? (Choose the one best answer for you in Column A, and then choose the one best for you in Column B.)

	Rank order of choice[2/]	% Students, Posttest, and Change from Pretest				Unweighted Average Change
		M-5	NM-5	M-6	NM-6	
Column A						
1. Since it is part of their religion the custom is alright.	4	18	16	21	18	
		+4	+2	+5	+4	+3.75
2. That is a bad custom even if it is for religious reasons.	2	34	33	34	33	
		+2	-4	-2	-8	-3.00
3. That is a horrible thing for any human being to do	1	24	27	20	23	
		-9	-4	-5	-1	-4.75
4. It is one of the most unusual customs I ever heard.	3	24	24	26	26	
		+2	+6	+3	+5	+4.00
	N=	556	551	527	485	
Column B						
1. It's hard to understand how people could do an awful thing like that.	2	29	27	21	26	
		-3	-6	-4	-2	-3.75
2. They are just people like we are, but they have different customs	4	28	25	32	27	
		+3	+4	+5	+4	+4.00
3. That custom is wrong and they should be made to change it.	1	17	21	16	17	
		-3	-4	-4	-6	-4.25
4. Some people have very strange customs, compared to our customs.	3	27	27	31	30	
		+3	+6	+2	+4	+3.75
	N=	553	556	524	483	

1. M = MACOS, NM = Non-MACOS; 5 = Grade 5; 6 = Grade 6.
2. The higher the rank, the more positive the statement.

Table III-32 Continued

2. If a friend of yours took your bike without permission, ran into a tree and bent the front wheel, what would you think? (Choose the one best answer for you in Column A, and then choose the one best for you in Column B.)

	Rank order of choice[2]	% Students, Posttest, and Change from Pretest				Unweighted Average Change
		M-5	NM-5	M-6	NM-6	
Column A						
1. I wouldn't mind; the bike can be fixed easily.	4	9	11	9	9	
		-2	0	-2	-3	-1.75
2. I would be very angry at my friend for doing that.	1.5	45	43	46	42	
		+2	0	+2	-1	+.75
3. That was bad luck for my friend and also for me.	3	12	9	11	13	
		+3	-1	+1	+2	+1.25
4. I think that was stealing and should not have been done.	1.5	34	37	33	36	
		-3	+1	-2	+2	-.5
	N=	545	545	521	476	
Column B						
1. Sometimes we forget to stop to think before we go ahead and do something.	3	43	38	39	36	
		+3	-1	+2	-2	+.5
2. Some people will do anything if they think they can get away with it.	1	20	26	23	25	
		-7	-1	-3	+2	-2.25
3. People usually have good reasons for doing things even if we don't know what they are.	4	13	13	15	16	
		+3	+4	-3	-2	+.5
4. Some people just don't take very good care of other peoples' things.	2	24	23	23	24	
		+1	-2	+4	+2	+1.25
	N=	534	536	513	477	

1. M = MACOS, NM = Non-MACOS; 5 = Grade 5; 6 = Grade 6.
2. The higher the rank, the more positive the statement.

3. If you heard that a group of people in South America tried to use magic to keep their enemies from hurting them, what would you think? (Choose the one best answer for you in Column A, and then choose the one best for you in Column B.)

Column A	Rank order of choice[2]	% Students, Posttest, and Change from Pretest M-5	NM-5	M-6	NM-6	Unweighted Average Change
1. I'd like to know more about how they do it.	4	44	38	43	41	
		+8	+4	-1	+6	+4.25
2. It won't work because there is no magic.	1	13	15	13	13	
		-4	-4	+1	-4	-2.75
3. Most people protect themselves as best they can.	3	27	27	28	22	
		+1	+3	+5	-4	+1.25
4. That's not a very smart way to protect themselves.	2	19	20	16	24	
		-3	-4	-4	+1	-2.50
	N =	547	531	519	477	
Column B						
1. They sound like foolish people to have such a silly custom.	1	15	14	12	14	
		-4	-3	-1	+1	-1.75
2. Those people may just be carrying on a very old custom.	3	41	37	41	42	
		+13	+6	+1	+8	+7.00
3. That group must not have up to date ideas like we do.	2	11	14	12	16	
		-4	-1	-2	+1	-1.50
4. It's good that different people can follow different beliefs and customs.	4	34	34	35	28	
		-3	-3	+2	-9	-3.25
	N=	532	526	515	473	

1. M = MACOS, NM = Non-MACOS; 5 = Grade 5; 6 = Grade 6.
2. The higher the rank, the more positive the statement.

Table III-32 Continued

4. If you heard that there was a country in which people often ate grasshoppers and earthworms, what would you think? (Choose the one best answer for you in Column A, and then choose the one best for you in Column B.)

Column A	Rank order of choice[2]	% Students, Posttest, and Change from Pretest M-5	NM-5	M-6	NM-6	Unweighted Average Change
1. Some people may eat them, but I wouldn't want to do that.	2.5	34	33	38	36	
		+3	+3	+3	+1	+2.50
2. I never thought that such things would be good to eat.	2.5	11	11	10	13	
		-1	-4	-1	+3	- .75
3. Yuck! It makes me sick just to think of eating them.	1	29	34	21	29	
		-9	-4	-9	-4	-6.50
4. That's no different from our country where people eat many things.	4	25	22	31	23	
		+6	+4	+7	+1	+4.50
	N=	544	530	519	473	
Column B						
1. I guess it must not do them any harm.	3	28	31	35	34	
		-2	+5	+2	+5	+2.50
2. I don't like people with such strange customs.	1	8	7	4	5	
		0	-2	+1	+1	0
3. They probably have good reasons for eating them.	4	48	43	48	47	
		0	-2	-3	-6	-2.75
4. They sound like a backward group of people.	2	15	19	14	14	
		+1	0	+1	0	+ .5
	N =	552	535	521	479	

1. M = MACOS, NM = Non-MACOS; 5 = Grade 5; 6 = Grade 6.
2. The higher the rank, the more positive the statement.

Table III-32 Continued

5. If you heard about a group of people in the United States who still firmly believe that the earth is flat, not round, what would you think? (Choose the one best answer for you in Column A and then choose the one best in Column B.)

Column A

	Rank order of choice [2]	% Students, Posttest, and Change from Pretest M-5	NM-5	M-6	NM-6	Unweighted Average Change
1. It would be interesting to know why they think that.	4	48	42	53	48	
		+6	+3	+7	-5	+2.75
2. It is hard to understand why they would believe that.	3	22	24	21	25	
		-3	+2	-2	+4	+ .25
3. Those people are crazy to believe a thing like that.	1	15	18	15	15	
		+1	+1	+2	+2	+1.00
4. That is a very strange thing for anyone to believe.	2	16	16	11	13	
		-1	-6	-7	-1	-3.75
	N=	527	532	515	469	

Column B

1. There are still some very backward people in our country.	2	14	18	10	12	
		-2	+3	-4	+2	- .25
2. They can go ahead and believe whatever they want.	3	33	29	33	35	
		0	-6	+2	+2	- .5
3. There may be good reasons why they have that belief.	4	43	39	46	42	
		+4	+5	+3	-3	+2.25
4. Some people are so dumb they will believe anything.	1	11	14	11	12	
		-2	-2	-1	0	-1.25
	N=	527	540	516	470	

1. M = MACOS, NM = Non-MACOS; 5 = Grade 5; 6 = Grade 6.
2. The higher the rank, the more positive the statement.

that response was +3.75%. The unweighted average change is simply the algebraic sum of the four change percentages for the statement, divided by 4. It should be noted that the statements under Column A and Column B for each question are listed in the order in which they appeared on the form.

With respect to question 1 taken as a whole, it may be seen that both for Column A statements and Column B statements the direction of change from pretest to posttest for all groups is toward the more positive (or at least less negative) statements. That is, if one compares the sign of the unweighted average percentages of change of percent of responses to the rank orders of the statements, it can be seen that statements with rank orders of 3 and 4 have a positive sign (more students chose those statements on posttest than on pretest). Statements with rank orders of 1 and 2 show negative signs; fewer students chose them on posttest.[59] The directions of change of response for all groups follow the same patterns in Columns A and B. On the whole, the percentage of students picking more positive statements increased from pre to posttest.

The resultant percentages of students in the four groups shown in the table for question 1 are generally very similar for each response choice. There are not obvious group or grade differences. Both groups showed small changes toward more positive choices on posttest for both parts of the question. It should

59. Statistical tests of differences of percentages (posttest or changes) were not undertaken. It is believed that the data should be examined impressionistically without attempting to make strong inferences based on statistical test criteria. One exception to this approach may be noted. With one exception, all posttest distributions of responses for each question and column differ significantly from a chance distribution. The exception is the distribution for non-MACOS 5th graders for Question 1, Column B.

not, of course, be inferred that all students who changed their choices did so in a positive direction. The figures presented here were obtained from cross-tabulations of responses of pretest vs. posttest for each group. Those cross-tabulations show changes in both directions as well as many students who did not change their responses at all. Consider, for example, the 5th grade MACOS students' responses in Column A for question 1. Not shown in the table, but obtainable from the pre-post cross-tabulation, is the fact that 35% of the students did not change their responses, 37% changed to a more positive response, and 28% changed to a more negative one.

Question 2 poses a situation involving oneself. The pattern of response (and changes from pretest) in Column A is generally away from the most positive reaction to the behavior (item 1) and toward a more negative or neutral one (items 2 and 3): There appear to be slight differences in the patterns of changes between MACOS and non-MACOS students. MACOS students were a little more likely to shift toward choosing a personal feeling of anger (item 2) and away from a moral condemnation (item 4) with respect to the behavior described. In Column B the net changes from pre to post are generally small, as they are in Column A. The overall shift is in a positive direction (away from choosing the most negative response (item 2), except for the non-MACOS 6th graders.

In question 3, the predominant reaction to the hypothetical custom on posttest was one of curiosity ("I'd like to know more about how they do it") in Column A. In Column B the general changes were away both from negative and

positive endorsing reactions toward a more factually oriented position (item 2: "these people may just be carrying on a very old custom").

Question 4, having to do with eating unusual food, has patterns of changes that may possibly reflect an effect of MACOS in that MACOS 5th and 6th grade groups overall show the largest net changes of choices from pre to posttest. In Column A, all four groups show a tendency to shift away from an initial reaction expressing personal distate (item 3) toward reactions other students had rated as more positive. The net shifts are more pronounced in the two MACOS groups than in the two non-MACOS groups. It is possible that for some students, at least, the experience of seeing the Netsiliks eating unusual foods may have diminished a purely personal reaction somewhat. In Column B the predominant response of all four groups is the most positive one (item 3: "they probably have good reasons for eating them"). In terms of average percentage of change, however, the shift generally is away from that reaction and toward a more neutral one (item 1: "I guess it must not do them any harm").

In question 5 the predominant posttest reaction of all groups for Column A responses is the most positive one (item 1: "it would be interesting to know why they think that"). The same is also true for Column B responses.

Overall for most questions the percentages of students in all four groups choosing particular responses are similar. The net changes from pretest to posttest in most cases are in a positive direction. In a few cases, they are toward a more neutral position. Comment has been made about the changes in percentages. It must be recognized, however, that such indices can only be viewed as suggestive.

They are, due, at least in part, to changes of opinions of students. In all cases, however, they are net changes. As noted earlier, the actual changes of responses by individuals within a group is not reflected in the net changes. Of interest here are overall patterns and relationships, not absolute values.

b. Study Choices

The variable SS Ch (Social Studies Choices) has been used as an outcome measure in analyses described earlier in this Section. That measure was the average for each class of the number of times social studies was chosen as preferred when it was paired with five other subjects: 1) arithmetic; 2) science; 3) reading; 4) spelling; and 5) English. The instrument from which the measure was derived used the method of pair comparisons, in which each subject, including social studies, was paired with each other subject once (for a total of 15 pairs).[60] The students were asked, for each pair, to pick the subject they would prefer if they had to make a choice. It is of interest to examine briefly how social studies compared in proportion of choices in relation to the other five subjects.

Scale values for each subject were determined using the method based on the assumption of a composite standard.[61] For purposes of comparability, normalized linear transformations of mean proportions for each group were converted to a scale with a mean of 5 and a standard deviation of 2. This was done because eight sets of scales were wanted: MACOS and non-MACOS 5th and 6th graders at pretest and posttest. The effect of the

60. $N \times N-1/2 = 15$, where N = No. of objects.

61. Guilford, J.P. Psychometric Methods (Second Edition) New York: McGraw Hill, 1954. See especially pages 168-177.

scaling was not to make the valid and absolute scale values for each group, but to spread proportions of choices on the same metric so that comparisons could be made. Other groups (and comparative evaluation models) could produce other values. The interest here is in the students in this study.

Table III-33 gives the scaled values of each of the six subjects for 5th graders, pretest and posttest, and for 6th graders, pretest and posttest, by group. The subjects are listed in the table from largest to smallest average scale value for all groups. Thus, arithmetic is listed first, reading second, etc.

Each distribution of scale values is read down a column. What is clear in Table III-33 is that, when paired with the rest of the field of choices, social studies ranked next to last in preference, (except with MACOS 6th graders at posttest) and English (language arts) last. What is not clear from the table is the consequence of the method. Many students may love English or social studies. Forced to choose between them and other subjects, as they were, they were somewhat more inclined to choose other subjects ("I really like sirloin, but if I have to choose between sirloin and filet mignon, I'll take filet mignon"). In effect, the method of pair comparisons forces winners and losers. The scaling procedure also distributes the proportions of choices (which are completely correlated with scale values) over a wider range than the absolute proportions, with the same parameters.

Given these conditions, it is clear that: 1) most subjects retain their rank order of preference when compared by students to the range of choices; 2) there are generally small differences in order of preferences between groups and grade

Table III-34
Scaled Pair Comparison Choices of Subjects,
by Group, Grade and Time of Testing*

	MACOS				Non-MACOS			
	Grade 5		Grade 6		Grade 5		Grade 6	
Subject	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Arithmetic	6.5	7.3	7.2	6.8	6.8	6.9	7.2	7.8
Reading	6.6	6.3	6.8	6.5	7.1	7.4	6.2	6.1
Science	6.3	5.9	5.2	5.0	5.2	5.7	6.2	5.4
Spelling	5.5	5.3	5.2	5.3	6.3	4.9	5.9	5.9
Social Studies	4.2	4.2	4.7	5.7	3.0	3.7	2.4	2.9
English	.9	1.1	1.0	.8	1.7	1.5	2.0	1.9

*N's: MACOS Gr. 5: 578, Gr. 6: 535; Non-MACOS Gr. 5: 572, Gr. 6: 502

levels in standard deviation terms; 3) most subjects retain their position relative to other subjects from grade level differences although science and spelling show a tendency to shift from pre to post.[62]

One is tempted to speculate about changes. Given the method, and understanding that the data in Table III-33 are based on the same 5th or 6th grade students from pretest to posttest, it is clear that, within the limits of accuracy (and assumptions) of scaling and rounding, for every gain there must be an equal loss. In fact, the algebraic sum of gains and losses from pretest to posttest, within rounding limits, is zero (as it must be, since each scale has a mean of 5).

While social studies usually ranked below other subjects than English both pre and post, it ranked relatively higher, even on pretest, for MACOS compared to non-MACOS students. This may have been due at least in part to the fact that MACOS had been started by students at time of pre-test and the choices reflected an early interest. MACOS 6th grade students showed a further increase in choice of social studies at posttest; MACOS 5th graders did not. Non-MACOS students showed some increase in choice of social studies at posttest at both grade levels. The shift of social studies upwards for MACOS 6th graders at posttest is large, relative to most other changes in Table III-33, suggesting a continuing impact of MACOS. It is not as large, however, as the downward shift in spelling seen for non-MACOS 5th graders.

A related series of questions was asked in Follow-up 2. Students were asked

62. The order of absolute proportions of choices is the same for each group, grade level and test period. All that the various linear transformations have done is to convert those proportions to scales with similar means and variances; they have not changed the order of absolute choices.

to rate how much they like social studies per se on a 4-point scale, ranging from 1, dislike very much, to 4, like very much. They rated arithmetic, science and English separately on the same scale. Students were also asked:

> "Of all the topics you have learned or studied about in school, whether in social studies or any other subject, which one would you really like to know more about, if you had the chance? Please describe it briefly."

A content analysis of responses was made, classifying responses as falling into the area of social studies vs. all other. The point biserial correlation between ratings of social studies and number of social studies subjects (one per student) was then computed. The point biserial correlations between ratings of liking social studies in Follow-up 2 and statements of subjects students would like to know more about (classified as social studies vs. all other) were: MACOS, $r_{pbi} = .20$; non-MACOS $r_{pbi} = .24$.

The means and standard deviations and N's, of ratings, and percentages of social studies statements were:

	Ratings			Statements
	M	SD	N	% Social Studies Statements
MACOS	3.18	.829	508	37%
Non-MACOS	3.24	.943	445	38%

With this absolute form of measure of liking of social studies, made in Follow-up 2 a year after MACOS, students in both groups rated social studies positively on the average.

The results, taken together, suggest that:

> . when students were asked to choose among pairs of subjects during the year MACOS was studied, social studies ranked low in preference on the average, for students in both groups, both pre and post;[63]

63. English or language arts was yet worse off; it consistently ranked last for all groups.

- the exception was for 6th grade MACOS students in which there was an increase in rank of social studies preference from 5th (next to last) at pretest to 3rd at posttest; the results suggest a more positive impact of MACOS on the 6th graders, compared to the MACOS 5th graders;

- non-MACOS students at both grade levels showed a relative increase in preference for social studies from pretest to posttest, although the rank of that subject did not change;

- students from both groups on the average rated liking of social studies positively to a similar extent on an absolute basis a year later in Follow-up 2.[64]

c. Skills and Knowledge

Two of the measures used as outcome variables have been the average of the summated ratings of eight questions aboust skills, and of seven questions about knowledge, asked of students in Follow-up 1. The specific questions asked were as follows:

> "A. We'd like to know what you studied in social studies last year, and if it's giving you an advantage in social studies this year.... Did you learn how to make or use maps in social studies? If you did not, circle NO. If you did learn to make or use maps last year, how much has it helped you this year? Circle 1 if it is a great advantage; 2 if it is some advantage; and 3 if it is of little or no advantage.
>
> "B. Now look at those items where you circle NO, you didn't study or do that last year. Are there some of those things you wish you had studied or done in social studies class last year because it would help you in social studies class this year? Put an X in the circle for those things."

The instructions in Part A were read aloud first; then each item was read aloud. Then part B was read to students. Table III-34 shows the percentages

64. Absolute and relative ratings obviously can give quite different information.

of responses for each item for 5th and 6th grade MACOS and non-MACOS students. It also gives the mean for each item based on a scale of 1 (great advantage) to 5 (didn't and wish I had). The last value was assigned to items that students circled NO and then put an X in the circle. The scale value 4 (didn't do or study) was assigned to items students circled NO but did not then put in an X. The table lists the skills items first, then the knowledge items. It may be seen that many of the distributions are bi-modal, but some are not. Considering only the means for items, the table clarifies why there were indications in earlier analyses of differences between groups for Skills and Know. While differences are not large, there are consistent relationships between MACOS and non-MACOS means at each grade level for skills items 1, 2, 3, and 4. For each of those items MACOS means are slightly higher than the non-MACOS means at the same grade level. Item 6 shows the reverse situation, although differences are very small. The remaining three skills items show mixed relationships. The four items on which MACOS students were slightly more likely to express a less positive opinion, as the scale is set up, are: 1) how to make or use maps; 2) how to make or use graphs; 3) how to find information in the library; and 4) how to write reports. The item for which non-MACOS students were on the average a little less positive than MACOS students is: how to support your ideas or opinions with evidence or facts. Of all items, it is the item concerned with map skills that shows the greatest difference between

Table III-34

Were Different Skills or Knowledge Studied Last Year and What Advantage Has That Been? Percentages of Responses and Means, by Group and Grade Level

Skill Items	Group and Grade[1]	1 Great Advantage	2 Some Advantage	3 Little or no Advantage	4 Didn't do or Study	5 Didn't and Wish I Had	Mean	N
1. How to make or use maps	M-5	14.0	29.0	15.1	17.9	23.1	3.1	351
	NM-5	19.9	36.5	26.3	8.3	9.0	2.5	312
	M-6	14.6	37.3	11.4	16.6	20.1	2.9	308
	NM-6	18.9	45.1	21.3	5.3	9.4	2.4	244
2. How to make or use graphs	M-5	7.5	18.2	16.7	32.3	25.4	3.5	347
	NM-5	8.0	19.2	24.0	29.2	19.6	3.3	312
	M-6	7.5	26.0	13.3	36.4	16.9	3.3	308
	NM-6	13.2	23.0	21.5	21.4	17.3	3.1	243
3. How to find information in library	M-5	24.1	29.6	11.2	27.6	7.5	2.7	348
	NM-5	29.1	28.2	12.6	21.0	9.1	2.5	309
	M-6	23.6	30.5	12.5	21.6	11.8	2.7	305
	NM-6	29.4	33.1	8.6	21.6	7.3	2.5	245
4. How to write reports	M-5	28.2	31.0	10.6	21.0	9.2	2.5	348
	NM-5	29.0	37.4	11.7	13.7	7.5	2.3	307
	M-6	24.0	37.7	15.0	15.0	8.3	2.5	300
	NM-6	31.3	34.1	14.2	12.2	8.1	2.3	246
5. How to tell the difference between facts and opinions	M-5	14.1	35.8	13.2	24.0	12.9	2.9	341
	NM-5	17.7	31.5	13.8	23.6	13.4	2.8	305
	M-6	18.2	41.3	11.6	19.8	9.2	2.6	303
	NM-6	15.1	37.6	11.8	23.3	12.2	2.8	245

1. M = MACOS, NM = Non-MACOS; 5 = Grade 5, 6 = Grade 6.

Table III-34 Continued

Skill Items	Group and Grade [1/]	1 Great Advantage	2 Some Advantage	3 Little or no Advantage	4 Didn't do or Study	5 Didn't and Wish I Had	Mean	N
6. How to support your ideas or opinions with evidence or facts	M-5	13.8	32.6	10.4	32.0	11.2	2.9	347
	NM-5	16.7	30.5	10.5	23.3	19.0	3.0	305
	M-6	17.4	37.0	11.1	25.2	9.2	2.7	305
	NM-6	17.2	33.6	10.2	27.0	11.9	2.8	244
7. How to work with other students in small groups	M-5	34.5	43.0	16.5	4.6	1.4	2.0	351
	NM-5	32.6	40.7	17.6	7.2	2.0	2.0	307
	M-6	31.0	47.9	14.9	4.6	1.7	2.0	303
	NM-6	23.6	38.6	26.0	9.8	2.9	2.3	246
8. How to look at all sides of a question before deciding what you think.	M-5	15.4	41.0	11.7	19.7	12.3	2.7	351
	NM-5	28.2	34.0	11.3	15.9	10.7	2.5	309
	M-6	18.6	40.2	14.7	19.6	6.9	2.6	306
	NM-6	21.8	36.2	13.6	20.2	8.2	2.6	243
Knowledge Items								
1. How people and their environment affect each other	M-5	26.0	36.9	17.7	14.0	5.4	2.4	350
	NM-5	23.7	32.8	17.5	16.9	9.1	2.6	308
	M-6	28.5	44.3	14.8	8.9	3.6	2.2	305
	NM-6	21.8	39.9	16.9	14.8	6.6	2.4	243
2. The history and/or customs of our country	M-5	15.8	28.4	11.7	23.8	20.3	3.1	349
	NM-5	24.5	42.3	11.9	11.0	10.3	2.4	310
	M-6	10.2	32.5	16.7	25.9	14.8	3.0	305
	NM-6	15.2	39.8	16.0	18.9	10.2	2.7	244
3. The history and/or customs of other countries	M-5	19.0	32.4	21.9	16.9	9.9	2.7	343
	NM-5	23.9	44.2	17.7	7.7	6.5	2.3	310
	M-6	16.9	34.9	19.2	15.0	14.0	2.7	307
	NM-6	33.6	35.7	22.1	5.7	2.9	2.1	244

1. M = MACOS, NM = Non-MACOS; 5 = Grade 5, 6 = Grade 6.

Table III-34 Continued

Knowledge Items	Group and Grade[1/]	1 Great Advantage	2 Some Advantage	3 Little or no Advantage	4 Didn't do or Study	5 Didn't and Wish I Had	Mean	N
4. Different beliefs people have	M-5	18.4	37.0	25.1	14.3	5.2	2.5	343
	NM-5	23.3	35.1	20.0	12.5	9.2	2.5	305
	M-6	16.8	42.6	21.8	12.9	5.9	2.5	303
	NM-6	18.9	42.6	22.1	11.5	5.9	2.4	244
5. How different animals behave and why they behave the way they do	M-5	25.9	32.0	30.8	7.6	3.8	2.3	344
	NM-5	6.5	15.5	7.1	37.9	33.0	3.8	309
	M-6	23.4	31.8	36.7	4.9	3.2	2.3	308
	NM-6	4.1	13.2	9.1	55.1	18.5	3.7	243
6. Similarities and differences in ways animals and people behave.	M-5	21.1	35.5	27.2	11.0	5.2	2.4	346
	NM-5	9.0	25.7	11.3	27.7	26.4	3.4	311
	M-6	17.3	36.9	35.3	6.5	3.9	2.4	306
	NM-6	5.3	23.4	16.8	43.0	11.5	3.3	244
7. Learning and understanding more about myself and people around me.	M-5	17.7	33.3	16.5	20.9	11.6	2.8	345
	NM-5	23.6	33.7	14.6	17.5	10.7	2.6	309
	M-6	21.2	29.8	14.6	22.2	12.3	2.8	302
	NM-6	22.2	38.7	8.6	22.6	7.8	2.6	243

1. M= MACOS, NM = Non-MACOS; 5 = Grade 5, 6 = Grade 6.

groups.[65] There is some evidence of content validity. In Section V it will be seen that non-MACOS teachers were more likely than MACOS teachers to emphasize map and globe skills.

Looking at the knowledge or content items, it may be seen that there are three items (1, 5,and 6) for which the responses of non-MACOS students on the average were less positive than those of MACOS students. There are 3 items for which the opposite is the case (2; 3, and 7). Item 4 (different beliefs people have) shows virtually identical average ratings for all groups. The differences in other items appear related to curriculum content. Somewhat more MACOS students, proportionately, at both grade levels, indicated they had not studied the history and/or customs of our country (item 2) or of other countries (item 3) and felt it would have been advantageous to them in their present social studies to have done so. Somewhat more non-MACOS students, proportionately, than MACOS students indicated the same with respect to the items dealing with animal behavior and similarities and differences of animal behavior and human behavior (items 5 and 6). MACOS students gave slightly more positive ratings than non-MACOS students at the same grade levels for item 1: how people and their environment affect each other. Non-MACOS students gave slightly more positive responses than MACOS students at the same grade level to item 7: learning and understanding more about myself and people around me.

65. It will be recalled that in earlier analyses, class means were used as the unit of analysis. Here, as with other item analyses, the individual student is the unit. It is possible in the item analyses for students from a small number of larger classes to impact upon the distributions and means more than would have been the case with unweighted class means.

Nearly all means shown in Table III-34 fall in the range of scale values 2-3: 'did study or learn how to do this and have found it some to little or no advantage this year.' The question did not ask students to indicate reasons why they responded as they did. The last two categories (4 and 5) have been treated as if they were continuation of a scale such as that defining the first three categories, although they are not psychometrically so, as the distributions make clear. There is an ambiguity in the last two categories "Didn't do or study" and "didn't and wish I had" could mean: 1) this was not covered in social studies last year, or 2) it was and I did not learn it for whatever reason.[66] Despite these and other limitations of the method, there do appear on the average to be some trends that seem related to the MACOS curriculum, as has been noted.

f. Attitudes of students during the year after taking MACOS

An important purpose of follow-ups with students from each of the classes in the study was to determine how students viewed their previous year's social studies course in retrospect. Follow-up 1, in October, 1975, focused on attitudes toward or perceptions of the previous year's course. It also sought to survey students' opinions of certain potentially emotionally charged topics or subjects students may have studied or learned about in their previous social studies course. In particular with MACOS students, there were questions about reactions to such things as seal and caribou killing, treatment of elderly Netsiliks, treatment of Netsilik children, hunger, sexual customs, foods eaten (fish eyes, raw liver) etc. In interviews conducted with students at the end of the prior school year, May 1976 (during posttesting), MACOS and non-MACOS students students had been asked if

66. Responses to all such items also may be influenced by response sets, halo effects, social approval sets and other extraneous factors.

they had studied or learned about customs or beliefs they felt were strange or wrong. Those interviews (reported and discussed in Section V of this report revealed a number of topics (in MACOS and non-MACOS classes) about which some students expressed negative opinions. Inquiry about students' opinions on a number of such topics was pursued both in Follow-up 1 (October, 1975) and Follow-up 2 (May, 1976, a year after students had taken MACOS). This inquiry was not intended as a clinical or psychological assessment, but as a survey of continuing opinions.

Some of the data collected in Follow-up 1 and 2 were used as measures of outcomes in analyses reported above. In those analyses, class averages were used as the unit of analysis.[67] In the opinion results to be reported here, the unit of analysis is the individual student. This is done in part because the form in which some of the questions were asked did not lend itself to forming averages for a class, and in part because it was thought that the results could justifiably be viewed as a survey of individuals, even though the individuals came from clusters (classes) the previous year that could make opinions of such sub-groups correlated.

1) Follow-up 1 Opinions About Emotional Topics

What were students' opinions in retrospect of a number of potentially emotional or value-oriented topics? The following topics were listed in the questionnaire used in Follow-up 1:[68]

67. In both follow-ups, the class was a 50% random sample from each MACOS and non-MACOS class of students who had been in the class all year (1975-76). The same students in Follow-up 1 were sought for Follow-up 2. Sample size were maintained by replacement on a random basis. Approximately 80% of the students in Follow-up 1 were also in Follow-up 2.

68. The topics listed were the result of a design conference with consultants in which the question of what subjects or topics to include was discussed at length.

. Cannibalism

. Different beliefs people have about what is right and wrong

. Killing animals

. Killing people

. Love

. Hatred

. Foods that different people eat

. Leaving people to die

. Different religious beliefs

. Different sexual customs

. Slavery

. Sharing and cooperation

. Starvation

. Treatment of young people

. Treatment of old people

Many of the topics were stated in a form that was not specific to a particular course in order to make it possible for students to respond in terms of the content of their particular social studies course. Students were asked the following questions about each topic[69]:

> . "There are people who are concerned about some of the things 5th and 6th grade kids may study in school. Last year in social studies did you study or learn about any of the following things? If you did please put a check mark in Column Number 1 just beside the ones you did study in social studies last year".

69. The questionnaires in Follow-up 1 and Follow-up 2 were read aloud to students to minimize reading problems. The only exception was that instruments contained in the Follow-up 2 questionnaire that had been part of the pre-post battery were administered as they had been on those occasions. Instructions were read aloud; items were not. In Follow-up 1, students were interviewed as a group after the questionnaire

."Were there any of those things you checked in Column Number 1 that really bothered or upset you? If there are any that bothered or upset you, and you wish you hadn't studied or seen or read about them last year, then put a check mark in Column Number 2 next to those topics you wish you hadn't studied."

."Even if you personally did not study these topics in social studies last year, are there any of them that you think 5th or 6th grade kids should not study in social studies? Check them off in Column Number 3."

These questions were formulated as a means of investigating concerns expressed by various groups and individuals (see brief review in Section I) about MACOS, and also to continue to examine opinions about certain topics expressed by some students in the posttest interviews, as noted above. The purpose of the first question was to determine the relative prevalence of opinions about whether a topic had or had not been studied or learned in last year's social studies class. The purpose of the second question was to obtain an indicator of the impact of a topic on students who said they had studied or learned about it. The purposes of the third question were twofold: 1) to assess whether studying a topic influenced opinions of students about its suitability for students their age; and 2) to assess whether, for students who said that they'd studied a topic, there were indications of emotional impact which appeared to influence further opinions about suitability. Again it should be noted that the inquiry was not intended to be a clinical or psychological assessment, but a continuing survey of opinions expressed by students over the period of the school year following the courses under study.

The three questions were responded to one at a time. That is, students went through the list responding to the first question before the second question

---

was administered. They were asked, among other things, questions about these topics and their opinions about them. Results of those interviews are given in Section VB-7.

was asked. The third question was asked only after the second had been completed. Students were monitored carefully to assure as accurate completion of the questions as possible with respect to the mechanics of marking answers. An error in responding, for example, was to check a topic in response to the second question if it had not been checked in response to the first question. The questionnaire was administered to groups of students and occasionally errors of the kind described went undetected. Overall error rates of that type in Follow-up 1 were low for each topic, ranging from 0-4%. Error rates had an even smaller range in Follow-up 2.

In sum, three questions were asked: 1) did you study or learn about this last year in social studies? 2) did it bother or upset you? 3) whether you studied it or not, what is your op nion about its appropriateness for students you age? Results were tabulated separately for former 5th and 6th grade students (present 6th and 7th graders) in order to reveal possible age level differences in opinions. Table III-35 a through o, summarizes results for Follow-up 1. The sub-tables give percentages of students responding to each of the three questions. The percentages are additive within a group (across a row) since the response of an individual student could be classified only in one of the seven possible categories shown. The tables are set up to permit ease of comparison between responses of MACOS and non-MACOS students at a given grade level. It is of course possible also to compare responses of 5th and 6th grade MACOS or non-MACOS students. In each table, the column labelled 'No Data' contains the percentages of students in the particular sub-group that completed the questions for that topic incorrectly, as described above. The 'No Data' category thus gives the error rate for the sub-group and topic.

Each sub-table in Table III-35 contains the following information, expressed in terms of percentages of students for the sub-group:

- the percentages of students indicating that they had not studied or learned about the topic in social studies the year before, and expressing no further opinion about that topic (Didn't Study, No Opinion);
- the total percentage indicating that the topic had not been studied, and should not, in their opinion, be studied by 5th or 6th grade students (Didn't Study, Shouldn't Study)
- the total percentage of students indicating they did not study or learn about the topic (the sum of percentages of Didn't Study, No Opinion and Didn't Study, Shouldn't Study)
- the percentage of students indicating they did study or learn about the topic last year in social studies and expressing no further opinion about it (Studied, No Opinion)
- the percentage of students indicating they studied the topic last year and it bothered them (Studied, Bothered)
- the percentage of students indicating they studied the topic last year, it bothered them, and 5th or 6th grade students should not study it (Studied, Bothered, Shouldn't Study);
- the total percentage of students who indicated they studied the topic and were bothered by it (the sum of percentages of Studied, Bothered and Studied, Bothered, Shouldn't Study)
- the percentage of students indicating they studied the topic, it did not bother them, but that in their opinion, 5th or 6th grade students should not study it (Studied, Not Bothered, Shouldn't Study);
- the total percentage of students indicating they studied a topic (the sum of percentages of Studied, No Opinion; Studied, Bothered; Studied, Bothered, Shouldn't Study; and Studied, Not Bothered, Shouldn't Study)
- the total percentage of students expressing the opinion that the topic shouldn't be studied by 5th or 6th graders, whether the respondents had studied it or not (the sum of percentages of Didn't Study, Shouldn't Study; Studied, Bothered, Shouldn't Study; and Studied, Not Bothered, Shouldn't Study)

. the percentage of students completing the questions erroneously (No Data);

. the approximate percentages of students in a particular category, for example, Didn't Study, who did or did not express further opinions about the topic (e.g., Didn't Study, No Opinions, or Didn't Study, Shouldn't Study divided by Didn't Study, No Opinion plus Didn't Study, Shouldn't Study[70]:

In each table percentages are rounded.[71] If the percentage category in any cell is not zero, but is less than .5(%), that fact is indicated by a dash (-). That condition indicates the presence of 1 respondent classified in that category. If the percentage is specifically zero, it is shown as 0 (%). Percentages added for all seven categories should total to 100%, but may not due to rounding.

Consider Table III-35, a , with respect to the three main questions asked. For MACOS 5th grade students (N=361), 69% plus 16%, or 85% indicated they had not studied or learned about cannibalism last year. Approximately 15% (13 plus 1 plus 1) indicated they had studied or learned about it, and an additional 1% completed the three questions incorrectly (No Data). The total is 101%, due to rounding errors. Approximately 2% of the total group indicated they had studied cannibalism and it bothered or upset them. Approximately 17% of the total group indicated by their response that they thought the topic should not be studied by 5th or 6th graders. The comparable percentages for the non-MACOS 5th grade (now 6th grade) students are: studied it, 83%, didn't study it, 17%; studied and bothered, 5%;

70. The within sub-category percentages are approximate due to rounding and also due to the No Data percentages. Since the No Data percentages are typically very low, the within sub-category percentages derived by this method are considered to be sufficiently accurate for general descriptive purposes.

71. A .5 or more rule was used for rounding up; less than .5 for rounding down.

Table III-35

Follow-up 1: Recollections of Studying Topics Last Year and Opinions About Them, by Grade and Group (Percentages of Students)[1]

Topic	Grade[2]	Group	Didn't Study No Opinion	Didn't Study Shouldn't Study	Studied No Opinion	Studied Bothered	Studied Bothered Shouldn't Study	Studied not Bothered Shouldn't Study	No Data	N
a. Cannibalism	5	MACOS	69%	16%	13%	1%	1%	-%	1%	361
		Non-MACOS	69	14	11	1	1	1	1	312
	6	MACOS	66	19	17	1	1	1	1	290
		Non-MACOS	68	13	17	-	1	1	-	248
b. Different beliefs people have about right and wrong	5	MACOS	25%	1%	70%	3%	1%	1%	1%	361
		Non-MACOS	14	1	80	4	1	1	1	312
	6	MACOS	21	-	76	-	1	1	1	290
		Non-MACOS	23	1	71	2	-	1	2	248
c. Killing Animals	5	MACOS	20	5	40	21	11	1	2	361
		Non-MACOS	36	10	25	16	8	2	4	312
	6	MACOS	23	8	46	13	6	2	2	290
		Non-MACOS	45	14	27	6	4	2	3	248
d. Killing People	5	MACOS	53	21	14	5	3	1	2	361
		Non-MACOS	49	16	13	11	5	3	4	312
	6	MACOS	49	19	19	4	4	3	2	290
		Non-MACOS	44	19	25	4	3	3	2	248
e. Love	5	MACOS	57	12	28	-	1	2	1	361
		Non-MACOS	61	7	20	1	1	1	0	312
	6	MACOS	58	6	35	0	0	1	-	290
		Non-MACOS	57	12	28	-	1	2	0	248

1. Percentages are rounded and may not add to 100 due to rounding. A dash (-) indicates a percentage of less than .5%; a 0 indicates no respondents in a cell.

2. Former grades - 5th grade students were presently 6th graders; 6th grade students were presently 7th graders.

Table III-35 Continued

Topic	Grade	Group	Didn't Study Didn't Study No Opinion	Didn't Study Shouldn't Study	Studied No Opinion	Studied Bothered	Studied Bothered, Shouldn't Study	Studied not Bothered Shouldn't Study	No Data	N
f. Hatred	5	MACOS	50	7	36	4	1	1	2	361
		Non-MACOS	39	8	40	6	3	2	2	312
	6	MACOS	46	10	40	3	1	-	1	290
		Non-MACOS	40	8	45	2	1	4	0	248
g. Foods that different people eat	5	MACOS	14	1	67	14	2	1	1	361
		Non-MACOS	14	3	74	5	2	2	1	312
	6	MACOS	17	2	71	9	-	-	-	290
		Non-MACOS	15	1	82	1	-	-	0	248
h. Leaving people to die	5	MACOS	29	7	39	16	7	2	1	361
		Non-MACOS	53	14	17	7	4	2	3	312
	6	MACOS	29	8	45	8	6	4	1	290
		Non-MACOS	53	16	22	4	2	2	1	248
i. Different religious beliefs	5	MACOS	19	1	74	3	-	1	1	361
		Non-MACOS	22	3	71	1	2	1	1	312
	6	MACOS	11	2	86	1	-	0	0	290
		Non-MACOS	15	1	83	0	1	1	0	248
j. Different sexual customs	5	MACOS	43	19	32	1	1	3	1	361
		Non-MACOS	52	20	23	1	2	1	1	312
	6	MACOS	49	14	29	-	-	5	2	290
		Non-MACOS	42	21	31	1	0	6	-	248
k. Slavery	5	MACOS	50	3	31	10	3	1	2	361
		Non-MACOS	21	2	47	17	8	3	1	312
	6	MACOS	56	4	29	5	3	2	2	290
		Non-MACOS	23	3	62	5	2	4	1	248

Table III-35 Continued

Topic	Grade	Group	Didn't Study, No Opinion	Didn't Study, Shouldn't Study	Studied, No Opinion	Studied, Bothered	Studied Bothered, Shouldn't Study	Studied not Bothered, Shouldn't Study	No Data	N
l. Sharing and cooperation	5	MACOS	31	1	65	1	–	1	–	361
		Non-MACOS	30	2	65	1	–	1	1	312
	6	MACOS	31	1	66	1	–	0	1	290
		Non-MACOS	37	2	59	1	–	–	–	248
m. Starvation	5	MACOS	31	3	45	17	3	1	1	361
		Non-MACOS	29	4	40	17	4	5	1	312
	6	MACOS	35	3	47	8	4	2	1	290
		Non-MACOS	37	4	49	8	–	–	1	248
n. Treatment of young people	5	MACOS	48	3	41	4	1	2	1	361
		Non-MACOS	56	5	29	4	3	1	2	312
	6	MACOS	36	3	55	2	1	1	1	290
		Non-MACOS	52	4	42	1	1	1	1	248
o. Treatment of old people	5	MACOS	36	3	47	9	3	2	–	361
		Non-MACOS	44	6	35	7	5	2	1	312
	6	MACOS	30	4	58	3	3	1	1	290
		Non-MACOS	46	2	47	1	2	2	1	248

studied and not bothered, 15%; should study this topic, 16%.

Still consider the same two groups for the topic 'cannibalism'. Of the 5th grade MACOS students who said they didn't study it, approximately 19% said it should not be studied (16% ÷ 69% + 16%, or 16 ÷ 85); 81% expressed no negative opinion about suitability (69 ÷ 85). Of the approximately 15% of the students who said they studied it, about 7% expressed the opinion that it should not be studied (1 ÷ 13 + 1 + 1, or 1 ÷ 15), while about 93% did not express that opinion. Of the approximately 15% of the students who said they studied it, about 13% (1 + 1 ÷ 15, or 2 ÷ 15) also said it bothered them, while about 87% did not express that opinion. The comparable approximate percentages for the non-MACOS 5th grade group are:

- didn't study, shouldn't, 17%; didn't study, no adverse opinion, 83%;
- studied, shouldn't 12%; studied, no negative opinion about suitability, 88%;
- studied, bothered, 29%; studied, not bothered, 71%.

The two groups of students, (MACOS and non-MACOS 5th graders) were quite comparable, for this topic, with respect to the three main questions, although the non-MACOS 5th graders who said they had studied cannibalism were a little more likely to say it had bothered them (in Follow-up 1) than were non-MACOS 5th grade students who also said they had studied it. Although a number of combined figures have been given to illustrate the information in the tables, the basic relations are evident in them. There were several ways in which the information could have been organized. It was felt that the form of presentation in Table III-35 provides the reader with the most generally useful information, and also provides the basis for presenting summary descriptive statements.

2) Summaries of Follow-up-1 Opinions by Topics

Taking the topics listed in Follow-up 1 in order (Table III-35a-o), the following are descriptive summary statements of how the project viewed the results.[72]

Cannibalism

Most MACOS and non-MACOS students at both grade levels said they had not studied it. A few students in all groups said they had and it bothered them. Non-MACOS 5th-graders who said they had studied it were a little more likely to have said it bothered them than were students in other groups.

Different beliefs people have about what is right and wrong

The majority of students in all four groups said they had studied them. Former 5th grade students both in MACOS and non-MACOS groups who said they studied them were slightly more likely to indicate being bothered than former 6th grade students in both groups. Slightly smaller proportions of students in all groups who said they studied them felt they were unsuitable as topics of study than the proportions of students saying they did not study them. With respect to both opinions, the proportion of students expressing negative opinions in all groups was small. This did not appear to be a potentially sensitive topic as gauged by students' opinions.

72. The statements are impressionistic. It was decided not to test the significance of differences for individual follow-up results (Follow-up 1, Follow-up 2), but to test the significance of changes in responses from FU-1 to FU-2.

Killing Animals

The majority of all students except non-MACOS 6th graders said in Follow-up 1 they had studied or learned about that. Proportionately more 5th graders both in MACOS and non-MACOS groups who said they had studied it also said it bothered them than did 6th graders who said they had studied it (approximately 44-47%, compared to approximately 26-28%). Proportionately more students in all groups who said they had studied it and were bothered by it were likely to say that the topic should not be studied than those who said they studied it and weren't bothered. Proportionately about the same number of students both in MACOS and non-MACOS 5th grade groups were likely to say the subject was inappropriate whether or not they said they had studied it, so there seemed to be no effect of studying, learning or seeing it or not. That is, curriculum in this case did not seem to change opinions with younger students, given only Follow-up 1 results. With 6th grade MACOS and non-MACOS students, opinions about suitability were somewhat more apt to be negative if students said they had not learned about killing animals (24-26%) than if they said they had (12-15%). This appeared to be a subject which had strong potential impacts on both groups (MACOS and non-MACOS), as gauged by Follow-up 1 responses. The impacts seemed greater on former 5th grade than on former 6th grade students in both groups.

Killing People

The majority of students in all groups in Follow-up 1 said they had not studied that (63-85%). This was a topic for which a proportionately greater number of non-MACOS than MACOS 5th graders said it bothered them (50% vs. 35%). Both 5th grade groups had greater proportions of students than 6th graders indicating that it bothered them. Proportions of students in both groups and grade levels were similar with respect to considering this topic unsuitable for students their age, regardless of whether they said they had studied it or not. Students in both groups at both grade levels were only very slightly more likely to judge the topic unsuitable if they said they had studied it and it bothered them. This appears to be a topic sensitive to grade level, not to curriculum.[73]

Love

Table III-35, e, makes it clear that most students in both groups did not connect this topic with anything they recalled studying or learning about in social studies the year before. Few students in any group or grade level claimed to have been bothered by it if they said they had studied it, or thought it should not be studied.

73. In what context might students in non-MACOS classes have learned about killing people? Current events (e.g., Vietnam), and history (battles) came up during Follow-up 1 interviews as frequently cited examples.

Proportionately somewhat more students thought the topic unsuitable if they said they had not studied it than if they said they had. This topic appeared to be one that was of little consequence in terms of impact to students who said they had studied (learned about) it in some sense in social studies the year before, and of slightly more consequence to students who said they hadn't.

Hatred

Sixth graders in both groups were more likely to say they had studied about that than 5th graders. Further, with the exception of non-MACOS 6th graders, students were a little more likely to express concern (bothered) about that topic if they said they had studied it than was the case with "love." It may be noted here that in administering the questionnaire, students frequently asked what hatred meant. It is concluded that this was not a particularly sensitive topic to students who thought they understood it and who said they had learned something about it, relative to students who said they had not studied it. Its negative connotation seemed to provide pause for thought among a few students.

Foods that different people eat

Here, in Table III-35, j, it is readily apparent that proportionately more students in both grade levels of MACOS students reacted a

little more strongly to this topic than non-MACOS students at comparable grade levels. The great majority of all students recognized this as a topic they had studied or learned in social studies the previous year, a point that lends face validity to results. MACOS students, in posttest and Follow-up 1 interviews, often mentioned the Netsiliks' eating habits; non-MACOS students seldom mentioned foods as a topic of interest or concern.[74] Proportionately few students who said they had learned about the subject thought it unsuitable for students their age, regardless of whether they said they were bothered by it. A somewhat greater percentage of students who said they had not learned about such things expressed the opinion that they would be inappropriate to study or learn. It is concluded that in Follow-up 1 the Netsiliks' eating habits continued to be vivid in the minds of many MACOS students, although students did not, for the most part, go on to express the opinion that the matter was age inappropriate.

Leaving People to die

This concept or topic was intended specifically to contrast MACOS with non-MACOS students, and the results in Table III-35, k,

74. See Section V5B6 and 7 for further information on this subject.

year before. Few students indicated that they thought the subject age inappropriate, whether they said they had studied it or not.

Different sexual customs

The majority of students in MACOS and non-MACOS classes said they had not learned about such things. If students in either main group (MACOS or non-MACOS) said they had studied about such matters, 6th grade students in both groups were slightly more likely to indicate that they thought they were age inappropriate (15-16% vs. 11%). Few students in either main group, regardless of grade level, said they were bothered by such matters if they said they had studied them. Students who said they had studied them and also said they were not bothered were a little more likely to say they were unsuitable than if they said they had been bothered by them Students who said they had not learned about such matters were proportionately somewhat more likely to express the opinion that this topic was grade unsuitable than students who said they had studied them (roughly, 22-33% vs. 11-16%, respectively).[75] This seemed to be a topic which, whatever students understood by it: 1) differentiated between students who said they had and hadn't studied such matters, no matter what the curriculum; and 2) evoked judgments about suitability from students who said they had studied about it that were

75. Again, interview data, discussed in Section VB-6 and 7 help clarify what these responses meant.

year before. Few students indicated that they thought the subject age inappropriate, whether they said they had studied it or not.

Different sexual customs

The majority of students in MACOS and non-MACOS classes said they had not learned about such things. If students in either main group (MACOS or non-MACOS) said they had studied about such matters, 6th grade students in both groups were slightly more likely to indicate that they thought they were age inappropriate (15-16% vs. 11%). Few students in either main group, regardless of grade level, said they were bothered by such matters if they said they had studied them. Students who said they had studied them and also said they were not bothered were a little more likely to say they were unsuitable than if they said they had been bothered by them Students who said they had not learned about such matters were proportionately somewhat more likely to express the opinion that this topic was grade unsuitable than students who said they had studied them (roughly, 22-33% vs. 11-16%, respectively).[75] This seemed to be a topic which, whatever students understood by it: 1) differentiated between students who said they had and hadn't studied such matters, no matter what the curriculum; and 2) evoked judgments about suitability from students who said they had studied about it that were

75. Again, interview data, discussed in Section VB-6 and 7 help clarify what these responses meant.

not dependent on possible emotional impact.

Slavery

The majority of non-MACOS students said in Follow-up 1 that they had studied this subject in social studies last year. The majority of MACOS students said they had not. If students in either group said they had learned about it, they were also apt to say the subject bothered them, 5th graders more so than 6th graders (29-33% compared to 10-21%). Proportions of students expressing the opinion that the subject should not be studied were similar for students who said they had or had not studied it. About the same proportions of students who said they had learned about slavery and were bothered by the topic said it should not be studied as students who said they studied about it and were not bothered. Overall, this appeared to be a topic which had an impact on some students if they said they remembered studying about it.

Sharing and cooperation

The majority of students in both groups at both grade levels said they recalled studying about this the previous year. It appeared to have little negative impact, as might be expected.

Starvation

The majority of students in both groups said they had learned about this in social studies the previous year. Of 5th grade

students in both groups who said they had learned about it approximately 30% also said it bothered them. Among 6th graders who said they had learned about starvation in social studies the previous year, approximately 20% of the MACOS students and 14% of the non-MACOS students also said it bothered them. For the most part, the proportions of students indicating that it was not an age appropriate subject were similar regardless of whether students said they had studied it or not. Also, for students who said they had studied it, there was no no obvious relationship between opinions about suitability and whether or not the subject bothered students. Overall, the subject of starvation seemed to have an impact on some students who said they recalled studying it, more so among 5th than 6th graders.

Treatment of young people

Table III-35, n, shows that except for MACOS 6th grade students, about 50% or more of the MACOS and non-MACOS students said they remembered learning about this topic in social studies the year before. There was a slight indication that 5th grade students in both groups were more apt to be bothered by the subject than 6th grade students. There was no strong indication that whether or not students said they had learned about it was associated with opinions of age appropriateness of the subject. There was

no indication that students both in MACOS and non-MACOS groups were more likely to consider it unsiutable if they also said they were bothered by the topic. This was not overall a topic about which students expressed negative opinions on the Follow-up 1 questionnaire.

. Treatment of old people

. Approximately 50% of the non-MACOS students said they had studied this topic last year; 60% or more of the MACOS students said they had. It was a subject which, if students said they had studied it, was more often said by 5th grade than 6th grade students to have bothered or upset them. Proportions of students expressing the opinion that it should not be studied were similar for groups that said they had or had not studied it. There was indication that this subject potentially had more impact on 5th graders from both groups than some of the other topics listed.

3) Opinions Expressed in Follow-up 2

Six of the 15 topics asked about in Follow-up 1 were included in Follow-up 2. The bases for selection were: 1) they were topics which appeared to have had more potential impact on MACOS and/or non-MACOS students, either generally or with respect to grade level; and 2) they were topics about which students were apt to express opinions when interviewed after completing the Follow-up 1 questionnaire. The topics included in Follow-up 2 were:

. Killing animals

. Leaving people to die.

. Foods that different people eat

. Starvation

. Treatment of old people

. Slavery

The purpose of surveying students' opinions about those topics which many students said they had studied (in some form) in social studies last year was to determine whether attitudes about them, as gauged by students' responses, remained the same, and if not, how they changed. The same questions as in Follow-up 1 were asked in the same way in Follow-up 2.

Approximately 80% of the students who were in Follow-up 1 were also in Follow-up 2. In order to assess stability and change, cross tabulation of their responses in Follow-up 1 versus Follow-up 2 was made for each of the 6 topics.

How stable were students' responses over time (from Follow-up 1 to Follow-up 2)? There are several issues underlying that question:

. Is there reason to believe that students' responses have a substantial basis (i.e., were proportions of consistencies and changes in students' responses from Follow-up 1 to Follow-up 2 other than chance occurrences)? If so, one would have more confidence in analyzing and interpreting consistent and changed responses.

. If there were substantial consistencies of or changes in responses, what were they (i.e. with respect to what topics and what questions)? For example, a marked shift by the same students from indicating that the topic 'killing animals' bothered and upset them to indicating in Follow-up 2 that it had not would support the hypothesis that the impact of the topic diminished over time.

. If there were substantive consistencies of or changes in responses, were they the same between groups (MACOS, non-MACOS) and between grade levels (5th grade, 6th grade)? For example, a differential change of opinions between MACOS 5th graders and MACOS 6th graders would suggest differential effects of time (and/or intervening experiences) related to age differences although other hypotheses are also tenable.

The first issue particularly needs further elaboration. Consider the question: "last year in social studies did you study or learn about /this topic/?"

There are four possible ways in which the responses of the same students could be classified if one considers responses for the same topic in Follow-up 1 and 2:

.Yes, yes

.No, no

.Yes, no

.No, yes

The first two possibilities are consistent responses from FU-1 to FU-2. The second two are changes. An indication of stability would be that the proportions of students giving the same response in Follow-up 2 as in Follow-up 1 are significantly different from the proportions giving a different response. More specifically, an indication of stability would be rejection of the hypothesis that the chance of giving consistent responses and of changing responses is 50%. Given the sample sizes involved here, that would mean essentially that the percentage of students in a particular group and grade level who maintained the same response from FU-1 to FU-2 should be equal to or greater than approximately 60% or equal to or less than approximately 40%.[76] Those limits were applied to an evaluation of the presumably factual question of whether or not a topic was studied last year in social studies, and to the question of whether students thought the topic was suitable or unsuitable for 5th and 6th grade students. The criteria appear to present a contradiction. If more than 60% of the students in a group changed responses to a question, that would appear to be a strong

76. Assume a sample size of 200. To accept the hypothesis that an obtained proportion is not different from .5, using a two-tailed test for the .05 level, those are the approximate upper and lower bounds. They are conservative for the larger sample sizes of some of the groups.

indication of instability. It was believed, however, that little could be said of consistency or change if the chances of both were 50/50.

The question of whether or not students continued to indicate that they were bothered (or not bothered) by a topic is contingent upon whether they continued to say they studied or learned about the topic. The points of interest in the case of this contingent question are the relative distributions of responses between groups and/or grade levels, and the amounts and direction of changes in responses for different groups of students. One should be conservative in interpreting such distributions. However, substantial differences and shifts of opinion, given that students continued to state that they studied the topic, are of interest. Shifts proportionately more in the direction of going from bothered to not bothered could suggest a diminishing of impact of the topic over time. Shifts proportionately more in the opposite direction would suggest an increase in impact over time, possibly because of further experience during the intervening time. Shifts that are proportionately equally likely in either direction are less substantively interpretable, although again the proportions of students giving consistent responses relative to those changing their response from Follow-up 1 to Follow-up 2 can provide information suggestive of the possible nature and direction of impacts.

Table III-36 through III-41 show for each topic how students, by group and grade level, responded to each of the three questions both in Follow-up 1 and Follow-up 2. With respect to the issue of stability of responses to the questions of whether students studied it and whether they thought the topic suitable or unsuitable, there was only one case in which the proportions did not meet the test criteria described above. For the topic "treatment of old people", the non-MACOS 6th grade group was equally

likely to give consistent and changed responses to the question of whether they had studied it (Table III-40a). Otherwise, all groups met or exceeded the chance criteria for those two questions for each topic.

The tables reveal interesting results. The substantive questions of the nature and direction of consistent responses and of changes, within and between groups and grade levels are generally clear. Consider, for example, Table III-36 which pertains to the topic 'killing animals'. A majority of MACOS students at both grade levels (67% and 61%) continued to say that they had studied that topic. Ten to fifteen percent of the MACOS students continued to say they had not. By comparison 20-25% of the non-MACOS students continued to say they had studied the topic in some form last year in social studies. When MACOS students changed their responses about studying the topic, it was proportionately more in the direction of going from not studied to studied than vice versa. In effect, the proportionate direction of change in response to the question in Follow-up 2 was from 'no' to 'yes'. The opposite was the case with non-MACOS students at both grade levels. The relative direction of change was from 'yes' to 'no'. Fifth grade non-MACOS students were more likely to change their responses than to remain consistent. Again, the proportionate changes were in the direction of 'no' (didn't study it).

Part b of Table III-36 shows, for the students in each group that consistently said they did study 'killing animals' last year in social studies, the percentages of students continuing to say that subject did or did not bother them, and the percentages of students changing their responses in effect from 'yes' to 'no' or from a 'no' to 'yes'. While the percentages of students continuing to say they were not bothered are roughly comparable in each group, there appears to be a grade level difference both

Table III-36

Responses of Students in Follow-up 2, Compared to Follow-up 1 to the Three Questions About the Topic Killing Animals, by Group and Grade

			In both Fu-1 and Fu-2		Changed from		
a. Did you Study it?	Group[1]	Grade	Studied	Not Studied	Studied to not studied	Not Studied to studied	N
	M	5	67%	10%	9%	13%	286
	NM		25	13	36	26	256
	M	6	61	15	7	17	244
	NM		20	46	21	13	209

			In both Fu-1 and Fu-2		Changed from		
b. Did it bother you?	Group	Grade	Not Bothered	Bothered	Bothered to not bothered	Not bothered to bothered	N
	M	5	43%	23%	22%	11%	192*
	NM		54	28	15	3	65*
	M	6	55	11	22	12	150*
	NM		51	17	12	20	41*

			In both Fu-1 and Fu-2		Changed from		
c. Should 5th 6th graders study it?	Group	Grade	Should Study	Shouldn't study	Should to shouldn't study	Shouldn't to should study	N
	M	5	69%	4%	14%	13%	286
	NM		68	6	14	11	256
	M	6	73	7	11	9	244
	NM		60	10	20	11	209

1. M = MACOS, NM = Non-MACOS

*Row percentages for this group are based on the number of students who consistently said (from Fu-1 to Fu-2) that they had studied the topic. Thus the N's are smaller. Percentages in all rows in the whole table may not add to 100% due to rounding.

Table III-37

Responses of Students in Follow-up 2, Compared to Follow-up 1 to the Three Questions About the Topic Leaving People to Die, by Group and Grade

	Group[1]	Grade	In both Fu-1 and Fu-2 Studied	In both Fu-1 and Fu-2 Not Studied	Changed from Studied to not studied	Changed from Not Studied to studied	N
a. Did you Study it?	M	5	49%	24%	17	10	286
	NM		9	59	22	10	258
	M	6	48	28	13	11	243
	NM		12	54	19	15	210

	Group	Grade	In both Fu-1 and Fu-2 Not Bothered	In both Fu-1 and Fu-2 Bothered	Changed from Bothered to not bothered	Changed from Not bothered to bothered	N
b. Did it bother you?	M	5	51%	23%	17%	9%	141*
	NM		57	17	9	17	23*
	M	6	66	4	18	13	111*
	NM		69	0	8	23	26*

	Group	Grade	In both Fu-1 and Fu-2 Should study	In both Fu-1 and Fu-2 Shouldn't study	Changed from Should to shouldn't Study	Changed from Shouldn't to should study	N
c. Should 5th/6th graders study it?	M	5	63%	6%	23%	7%	286
	NM		53	11	27	9	258
	M	6	63	9	20	8	243
	NM		51	10	30	9	210

1. M = MACOS, NM = Non-MACOS.

*Row percentages for this group are based on the number of students who consistently said (from Fu-1 to Fu-2) that they had studied the topic. Thus the N's are smaller. Percentages in all rows in the whole table may not add to 100% due to rounding.

Table III-38

Responses of Students in Follow-up 2, Compared to Follow-up 1 to the Three Questions About the Topic 'Foods Different People Eat' by Group and Grade

	Group[1]	Grade	In both Fu-1 and Fu-2: Studied	In both Fu-1 and Fu-2: Not Studied	Changed from: Studied to not studied	Changed from: Not studied to studied	N
a. Did you Study it?	M	5	72%	8	13	7	296
	NM		64	9	18	9	267
	M	6	73	10	8	9	247
	NM		73	3	11	12	215

	Group	Grade	In both Fu-1 and Fu-2: Not Bothered	In both Fu-1 and Fu-2: Bothered	Changed from: Bothered to not bothered	Changed from: Not bothered to bothered	N
b. Did it bother you?	M	5	71%	4%	19%	6%	214*
	NM		93	2	4	1	171*
	M	6	82	2	11	5	180*
	NM		94	1	1	4	158*

	Group	Grade	In both Fu-1 and Fu-2: Should study	In both Fu-1 and Fu-2: Shouldn't study	Changed from: Should to shouldn't study	Changed from: Shouldn't to should study	N
c. Should 5th/6th graders study it?	M	5	92	0	4	3	296
	NM		89	1	4	6	267
	M	6	96	0	2	2	247
	NM		94	0	3	2	215

1. M = MACOS, NM = Non-MACOS.

* Row percentages for this group are based on the number of students who consistently said (from Fu-1 to Fu-2) that they had studied the topic. Thus the N's are smaller. Percentages in all rows in the whole table may not add to 100% due to rounding.

Table III-39
Responses of Students in Follow-up 2, Compared to Follow-up 1
to the Three Questions About the Topic 'Starvation'
by Group and Grade

	Group[1]	Grade	In both Fu-1 and Fu-2: Studied	In both Fu-1 and Fu-2: Not studied	Changed from: Studied to not studied	Changed from: Not Studied to studied	N
a. Did you study it?	M	5	48%	20%	18%	14%	290
	NM		38	22	27	13	260
	M	6	40	29	21	10	247
	NM		44	27	16	12	214

	Group	Grade	In both Fu-1 and Fu-2: Not Bothered	In both Fu-1 and Fu-2: Bothered	Changed from: Bothered to not bothered	Changed from: Not bothered to bothered	N
b. Did it bother you?	M	5	58%	14%	18%	11%	139*
	NM		48	14	24	14	95*
	M	6	77	7	15	1	96*
	NM		74	8	16	1	85*

	Group	Grade	In both Fu-1 and Fu-2: Should study	In both Fu-1 and Fu-2: Shouldn't study	Changed from: Should to shouldn't study	Changed from: Shouldn't to should study	N
c. Should 5th/6th graders study it?	M	5	84%	2%	8%	6%	290
	NM		77	3	11	9	260
	M	6	84	3	6	7	247
	NM		81	0	13	6	214

1. M = MACOS, NM = Non-MACOS.

*Row percentages for this group are based on the number of students who consistently said (from Fu-1 to Fu-2) that they had studied the topic. Thus the N's are smaller. Percentages in all rows in the whole table may not add to 100 due to rounding.

Table III-40

Responses of Students in Follow-up 2, Compared to Follow-up 1 to the Three Questions About the Topic 'Treatment of Old People', by Group and Grade

	Group[1]	Grade	In both Fu-1 and Fu-2: Studied	In both Fu-1 and Fu-2: Not studied	Changed from: Studied to not studied	Changed from: Not Studied to studied	N
a. Did you study it?	M	5	38%	24%	12%	27%	293
	NM		16	46	20	18	260
	M	6	44	25	16	15	244
	NM		16	39	28	16	215

	Group	Grade	In both Fu-1 and Fu-2: Not Bothered	In both Fu-1 and Fu-2: Bothered	Changed from: Bothered to not bothered	Changed from: Not bothered to bothered	N
b. Did it bother you?	M	5	65%	5%	7%	22%	110*
	NM		66	5	15	15	41
	M	6	76	2	7	16	108*
	NM		77	6	6	11	35*

	Group	Grade	In both Fu-1 and Fu-2: Should study	In both Fu-1 and Fu-2: Shouldn't study	Changed from: Should to shouldn't study	Changed from: Shouldn't to should study	N
c. Should 5th/6th graders study it?	M	5	84%	1%	11%	4%	293
	NM		79	3	12	6	250
	M	6	83	1	12	4	244
	NM		84	1	11	4	215

1. M = MACOS, NM = Non-MACOS.

*Row percentages for this group are based on the number of students who consistently said (from Fu-1 to Fu-2) that they had studied the topic. Thus the N's are smaller. Percentages in all rows in the whole table may not add to 100 due to rounding.

Table III-41
Responses of Students in Follow-up 2, Compared to Follow-up 1 to the Three Questions About the Topic 'Slavery', by Group and Grade

	Group[1]	Grade	In both Fu-1 and Fu-2: Studied	In both Fu-1 and Fu-2: Not studied	Changed from: Studied to not studied	Changed from: Not Studied to studied	N
a. Did you study it?	M	5	23%	45%	21%	10%	284
	NM		67	11	8	14	264
	M	6	20	55	19	7	243
	NM		56	17	17	10	214

	Group	Grade	In both Fu-1 and Fu-2: Not Bothered	In both Fu-1 and Fu-2: Bothered	Changed from: Bothered to not bothered	Changed from: Not bothered to bothered	N
b. Did it bother you?	M	5	63%	3%	18%	15%	60*
	NM		53	13	24	11	178*
	M	6	79	6	13	2	48*
	NM		68	6	5	21	120*

	Group	Grade	In both Fu-1 and Fu-2: Should Study	In both Fu-1 and Fu-2: Shouldn't study	Changed from: Should to shouldn't study	Changed from: Shouldn't to should study	N
c. Should 5th/6th graders study it?	M	5	80%	2%	12%	5%	284
	NM		79	4	8	8	264
	M	6	84	2	8	6	243
	NM		77	4	14	5	214

*Row percentages for this group are based on the number of students who consistently said (from Fu-1 to Fu-2) that they had studied the topic. Thus the N's are smaller. Percentages in all rows in the whole table may not add to 100 due to rounding.

in MACOS and non-MACOS groups in percentages of students repeatedly saying the subject bothered them. Proportionately more 5th graders than 6th graders in both groups continued to indicate that the topic bothered them. The difference in percentages between MACOS 5th and 6th graders is significant ($X^2 = 7.909$, $df = 1$, $p \leq .01$); the difference for the non-MACOS groups, however, is not significant ($X^2 = 1.039$, $df = 1$, $p > .30$).[77] Students who changed their response from Follow-up 1 to Follow-up 2 in three of the four groups, were proportionately more likely to change from 'bothered' to 'not bothered'. Non-MACOS 6th grade students proportionately more often shifted their responses in the other direction.

The majority of students in all total groups indicated that they thought the topic was age or grade appropriate. Roughly comparable percentages of the total number of students in each group changed responses from FU-1 to FU-2 in each direction, although proportionately somewhat more non-MACOS 6th graders shifted from 'should' to 'shouldn't study it'.

The remaining tables can be interpreted in a similar fashion as a basis for arriving at conclusions about prevalences of topics studied within and between MACOS and non-MACOS groups, indications of impacts, and opinions about the suitability of various topics. There is, however, an issue that needs to be considered before such interpretations and conclusions are made. The issue is: do the percentages of students in Tables III-36 through III-41 shown as changing responses in one direction or the other from FU-1 to FU-2 adequately depict the significance of those changes? Consider in Table III-36c the percentages of students changing from 'shouldn't' to 'should',

77. Chi-square test here compared frequencies of responses classified as consistently said 'bothered' vs. 'all other responses'.

and 'should' to shouldn't' study 'killing animals'. For most groups, the percentages appear about the same for changes in each direction. Those percentages, however, are based on the number of students in a given category relative to the total group. Suppose students are classified according to their original response in Follow-up 1. Suppose further that the Follow-up 2 responses of each of those groups are classified as 'same' or 'different' compared to their Follow-up 1 responses. Suppose, finally, that one asks the question: "are there differences in proportions of change in response between the two groups?" For the MACOS 5th grade students in Table III-36c, a table based on that question would be as follows:

Group	Follow-up 1 Response	Follow-up 2 Response* Same	Different	N
MACOS Grade 5	Shouldn't Study	24%(12)	76% (38)	50
	Should Study	83 (197)	17 : (39)	236
				286

*Numbers in parentheses are frequencies.

This table shows that while the same absolute number of students in each group (as defined by their responses in Follow-up 1) changed responses, students who originally said 'shouldn't' were proportionately far more likely to reverse their responses in Follow-up 2 than were students who originally said 'should'. The Chi-square test for that table is highly significant ($X^2 = 71.186$, $df = 1$, $p \leq .01$).[78]

78. The frequently used test of change proposed by McNemar tests the hypothesis that the number of changes in one direction is not different from the number of changes in the other direction. In the present example, the McNemar $X^2 = .013$, $df = 1$, $p > .90$. It does not take into account the fact that the two originally defined groups were of very different sizes (50 vs. 236). If group sizes were equal, the McNemar X^2 for change and the one computed for the illustrative table would be more comparable. For the McNemar X^2, see McNemar, Quinn, Psychological Statistics. New York: John Wiley Sons, 1949. The X^2 computed for the illustrative table was, in conventional notation:

$$X^2 = \frac{N(|AD-BC| - .5N)^2}{(A+B)(C+D)(A+C)(B+D)}$$

In effect, in response to the question: "even if you personally did not study these topics in social studies last year, are there any of them that you think 5th or 6th grade kids should not study in social studies?", students were much more likely, proportionate to their numbers, to go from 'yes, I think they should not, to 'no, I wouldn't say that about killing animals,' than from 'no' to 'yes'.

To aid in interpreting directions of change for all topics and questions, the change X^2 used above was computed. Since three questions were asked of each topic, and since there was interdependence or overlap between questions and respondents in different categories of response to different questions, it was decided that a conservative approach to decisions about significance of amount and direction of change was appropriate. Thus, to maintain a topicwise Type I error rate of .05, the Bonferroni procedure was employed to establish a critical value for the Chi-square test of each question within a topic. Thus, a p-value of .05/3, or .0167 was taken as the significance level needed for each of the three tests individually for a given group and grade level. The associated critical value of the X^2 is 5.731. The Chi-squares for each topic, question and group are given in Table III-42. The table gives the direction of proportionate change, as described above, whenever the Chi-square is significant. The table also gives the results of the test of the hypothesis that MACOS and non-MACOS students were proportionately equally likely to change their responses (in either direction).[79] In the event that one or more cells had an expected frequency of less than 5, no test was made. The situation that always produced that condition was that the number of students in one or both initial groups (defined by

79. For those tests, the Bonferroni procedure was not applied to determining significance. The rationale was that while the questions were related within a topic, the groups were independent.

Table III-42

Chi-Squares of Tests of Differences in Proportions of Students Changing Their Responses From Follow-up 1, and Direction of Change[1]

Topic	Question	MACOS vs. Non-MACOS[2] X^2 5th gr.	MACOS vs. Non-MACOS[2] X^2 6th gr.	MACOS[3] 5th Grade X^2	MACOS[3] 5th Grade Direction	MACOS[3] 6th Grade X^2	MACOS[3] 6th Grade Direction	Non-MACOS[3] 5th Grade X^2	Non-MACOS[3] 5th Grade Direction	Non-MACOS[3] 6th Grade X^2	Non-MACOS[3] 6th Grade Direction
Killing Animals	1. Did you study it?	.173	5.739	55.151*	Yes	19.052*	Yes	13.592*	No	17.304*	No
	2. Did it bother you?	4.467*	.008	15.609*	No	33.889*	No	7.817*	No	NT	-
	3. Should 5th/6th graders study it?	.130	5.595*	71.186*	Yes	33.908*	Yes	39.173*	Yes	11.057*	Yes
Leaving people to die	1. Did you study it?	1.233	4.989*	.563	-	1.49	-	80.688*	No	27.185*	No
	2. Did it bother you?	.058	.063	11.865*	No	36.925*	No	NT	-	NT	-
	3. Should 5th/6th graders study it?	1.694	4.787	9.04 *	Yes	6.502*	Yes	1.428	-	.669	-
Foods that different people eat	1. Did you study it?	3.506	2.094	23.014*	Yes	37.464*	Yes	24.852*	Yes	63.733*	Yes
	2. Did it bother you?	24.433*	8.133*	14.875*	No	NT	-	NT	-	NT	-
	3. Should 5th/6th graders study it?	.684	.106	NT	-	NT	-	NT	-	NT	-
Starvation	1. Did you study it?	3.103	.096	6.199*	Yes	1.829	-	NT	-	.237	-
	2. Did it bother you?	1.744	.027	22.737*	No	NT	-	14.912*	No	NT	-
	3. Should 5th/6th graders study it?	3.800	2.443	NT	-	NT	-	57.732*	Yes	NT	-
Treatment of old people	1. Did you study it?	.013	9.450**	27.779*	Yes	2.695	-	17.436*	No	23.693*	No
	2. Did it bother you?	.032	1.62	NT	-	NT	-	NT	-	NT	-
	3. Should 5th/6th graders study it?	.729	.038	NT	-	NT	-	NT	-	NT	-
Slavery	1. Did you study it?	7.159*	.144	27.885*	No	41.461	No	55.876*	Yes	3.088	-
	2. Did it bother you?	.001	1.878	NT	-	NT	-	39.762*	No	NT	-
	3. Should 5th/6th graders study it?	.293	1.522	NT	-	NT	-	63.834*	Yes	NT	-

1. NT means no test was made because a cell had an expected frequency of less than 5.
2. For Chi-squares in these groups,
   *p ≤ .05
   **p ≤ .01
3. *p ≤ .0167 (see text)

responses in Follow-up 1) was small. That often was the case especially with respect to the questions about being bothered by a topic, and about its suitability. For most of the topics, most students were not bothered, if they said they studied it both in Follow-up 1 and Follow-up 2; similarly, most students did not express the opinion that the topic was unsuitable in either follow-up.

The following are the main findings derived from the data in Table III-36 through III-42.[80]

a) Differences in prevalences of topics studied

- More MACOS than non-MACOS students at both grade levels studied 'killing animals,' 'leaving people to die,' and 'treatment of old people', according to Follow-up 1 and 2 responses.
- More non-MACOS than MACOS students at both grade levels studied 'slavery'.
- Similar percentages of students in both groups studied or learned about 'foods that different people eat' and 'starvation.'

These differences are consistent with content differences between MACOS and other curricula. None of the 6 topics listed, however, was completely unique to MACOS or to other social studies programs. Judging by the small percentages of non-MACOS students who consistently said they had studied 'leaving people to die' (9-12%), that topic comes closest to being unique when stated to students in a generalized form. About 50% of the MACOS students felt they recognized it as something they had read, seen, or heard about last year in social studies; few non-MACOS students did.

80 The results pertain to the approximately 80% of the students who were in Follow-up 1 and Follow-up 2. The results presented in the tables are for about 1,000 students who made no errors in completing questions in both follow-ups.

- When compared with each other, MACOS and non-MACOS 5th or 6th grade groups in four cases changed their responses from Follow-up 1 to Follow-up 2 at proportionately different rates with respect to whether a topic had been studied. Within different groups, there were even more frequent significant relative shifts of responses. Significant relative changes of responses were mainly non-MACOS students changing more to saying they had not studied the MACOS related topics ('killing animals,' 'leaving people to die,' and 'treatment of old people'); MACOS students shifted responses to 'no' with respect to a predominantly non-MACOS related topics, 'slavery'; and all groups changed to 'yes' with respect to the general topic, 'foods that different people eat.'

The trends in changes are interpreted as suggesting that some students initially reacted to affective qualities of the topics, as listed, and in Follow-up 2 reacted more to the implicit substantive content of topics in relation to memories of the prior year's social studies class.

b) Suggestions of the potency of different topics

- For each of the 6 topics, nearly all MACOS and non-MACOS 5th and 6th grade groups had at least a few students who consistently (in Follow-up 1 and Follow-up 2) said they had studied the topic and it bothered them. The one exception was non-MACOS 6th grade students with respect to 'leaving people to die.' None of the 26 students who consistently said they had studied it also said both times that it bothered them. The majority of those 26 students (69%) consistently said 'not bothered.'

- For all topics and groups for which there were large enough samples to make tests of significance of relative direction of change of response, significantly larger proportions of students changed from 'bothered' to 'not bothered' than in the other direction, given the relative sizes of those groups.

- With two exceptions, the majority of students in all groups consistently said 'not bothered' for each topic. The exceptions were MACOS 5th grade students with respect to the topic 'killing animals', and 5th grade non-MACOS students with respect to 'starvation.'

. The topics for which the largest percentages of students in all groups consistently indicated 'not bothered,' if they consistently said they had studied it, were 'foods that different people eat' and 'treatment of old people'.

. The topics for which the smallest percentages of students in all groups indicated 'not bothered' were 'killing animals' and 'leaving people to die'. Relative percentages of students indicating 'not bothered' for 'starvation' appeared related to grade level rather than group (MACOS or non-MACOS).

. Of the cases in which there were large enough sample sizes to test, only two showed a significant grade level difference within MACOS or within non-MACOS groups. Significantly larger percentages of 5th grade than 6th grade MACOS students consistently said they had been bothered by 'killing animals' and 'leaving people to die'. There were indications of greater impact of other topics on 5th grade students than on 6th graders in both groups. The differences in percentage either were not significant, however, given sample sizes, or not testable by the methods and ground rules employed.

. While topics varied in potency, sometimes in relation to general curriculum groups, sometimes according to grade level, sometimes both, the general trend from Follow-up 1 to Follow-up 2 was a diminishing potency, as suggested by consistencies and changes in responses to the question about being bothered.[81]

. The maximum absolute fraction of students in any group that were consistent in saying they were bothered by a topic, if they also were consistent in saying they studied it, was relatively small (no more than 11%).

c) Opinions about suitability of topics

. The largest percentages of students in both groups and grade levels consistently indicating they thought 5th and 6th graders should (or at least could) study pertained to the topics 'foods different people eat' (89-96%), 'treatment of old people', 'starvation', and 'slavery' (all the latter in the range of 77-84%).

81. The reader is directed again to Section V B7, in which a report of interviews with students following administration of the Follow-up 1 instrument is given. The data reported and summarized here bear on that report, and vice versa.

. Topics which the smallest percentages of students consistently indicated as suitable (did not indicate 'shouldn't study') were 'killing animals' (60-73%) and 'leaving people to die' (51-63%)

. With respect to 'killing animals', all groups and grade levels were significantly proportionately more likely to shift from 'shouldn't' to 'should study', given their original positions in Follow-up 1.

. With respect to 'leaving people to die', MACOS 5th and 6th grade students who originally indicated 'shouldn't study' were significantly more likely to shift in Follow-up 2 to 'should' than those who started with 'should' were likely to shift to 'shouldn't'. That was not the case for non-MACOS 5th and 6th graders, who were equally likely to shift in both directions. Few members of those two groups, however, consistently said they had studied the topic (23 5th graders and 26 6th graders).

. When any group changed its opinion from Follow-up 1 to Follow-up 2, and a test could be made of the significance of proportionate change in one direction or the other, all significant changes were in the direction of 'should study', rather than 'shouldn't study'. This is interpreted as further evidence of the diminishing of impact of any topic over time, possibly influenced by intervening learning, on non-MACOS as well as MACOS students.

d) General impressions

. Emotional reactions or negative opinions about various topics appeared to fade or diminish from Follow-up 1 to 2.

. Some topics or issues faded more than others, suggesting less centrality of some topics on the one hand, and the effects of mode of presentation on the other. A case in point is 'foods different people eat'. In Follow-up 1, former MACOS students were still much aroused (positively or negatively) over Netsilik Eskimos eating fish eyes and the like. By Follow-up 2, a year after the course, few former MACOS students were willing to indicate concern on the questionnaire. The predominant shift was from bothered to not bothered. More serious topics, such as 'leaving people to die', 'slavery', 'starvation', and 'killing animals' appear to have had more lasting effects with more students.

. Short term impacts appear easier to achieve with 5th graders than 6th graders.

- No inference from these data can be drawn about psychological effects beyond the substance of the questions asked. If students said they were bothered and wished they hadn't seen or read or heard about such things, that is a statement, not a diagnosis. Students differ in maturity and perspective. Furthermore, there is good reason to be bothered about some matters. Bothered does not necessarily connote 'dammaged', however defined.

- Considering the general thrust of social studies toward involvement of students' attention and interest in important issues and problems, the data suggest it is not easy to engage young students' minds and emotions in an enduring fashion, no matter what the curriculum and mode of presentation.

4) Opinions of Last Year's Class in Follow-up 2

Overall, what did students think of last year's social studies class a year later? Their global opinion was sought in Follow-up 2. In the context of several questions designed to help students recall last year's class, they were asked:

> "Still think about LAST YEAR'S social studies class - what you studied and how you studied it. If you had a younger brother or sister coming along, would you recommend that he or she take the same social studies course that you had last year?"

Students responded on a 4-point scale ranging from 'definitely no' to 'definitely yes'. The responses of all students in Follow-up 2 were tabulated, not just those of students who also had been in Follow-up 1. The results are given in Table III-43 for each group and grade level.

Table III-43 shows that the majority of students in all four groups responded on the positive side of the scale. The two questions of further interest, however, are:

- were MACOS students at a particular grade level more positive in their attitudes toward last year's class than non-MACOS students at the same grade level?

- was there a difference in attitude between 5th and 6th grade students

Table III-43

Percentages of Students Responding to Question About Recommending that a Sibling Take the Same Social Studies Course[1]

Grade 5	Definitely No	I Don't Think So	I Guess So	Definitely Yes	No Data[2]	N
MACOS	11%	12%	47%	30%	0	356
Non-MACOS	20	19	37	23	-	309
Grade 6						
MACOS	6%	11%	41%	42%	-	311
Non-MACOS	15	17	49	20	0	276

1. Percentages may not add to 100% due to rounding.
2. A dash (-) indicates less than .5%.

in the MACOS group or in the non-MACOS group?[82]

Table III-44 gives the means and standard deviations for the four groups of students. It also shows the point-biserial correlation coefficient (r_{pbi}) for each comparison made in order to provide a measure of strength of relationships between group membership and ratings. It can be seen in Table III-44 that the average ratings of MACOS students were more positive than those of non-MACOS students, more so for 6th grade students than 5th grade students. Chi-square tests were made of the differences between distributions of responses for each pair of groups for which a point-biserial correlation coefficent was computed. All Chi-squares were significant at or beyond the .05 level. While the difference between each pair of distributions of ratings is significantly different from chance, the point-biserial correlations indicate that with one exception, the correlation between group membership and ratings is small. The exception is MACOS, grade 6 compared to non-MACOS, grade 6. There the point-biserial correlation is .25, substantially more strength of association than shown by other groups.

As a means of maintaining perspective on these results, one can also examine the index of predictive association for each pair of variables.[83] Suppose one asks, in each case: What would be the reduction in error of prediction of recommendation ratings, knowing group and grade level? What would be the

82. It should be remembered that students identified as 5th and 6th graders here were in fact finishing 6th and 7th grade, respectively, at the time they were administered this survey.

83. Cf. Hayes, William L., Statistics for Psychologists. New York: Holt, Rinehart and Winston, 1963. pp. 606-610.

Table III-44

Means and Standard of Ratings of Recommendations of Last Year's Class, by Group and Grade, and Point-biserial Correlations of Between Groups by Grade, and Between Grade Levels Within Groups

Group	Grade	N	Mean	SD	Point biserial Correlation Coefficients for: Grade 5 MACOS/Non-MACOS	Grade 6 MACOS/Non-MACOS	MACOS Grade 5/6	Non-MACOS Grade 5/6
MACOS	5	356	2.9	.93	.08	.25	-.07	.03
Non-MACOS	5	308	2.6	1.05				
MACOS	6	310	3.2	.86				
Non-MACOS	6	276	2.7	.94				

percentage of reduction of error of prediction of group membership, knowing ratings? What would be the percentage of reduction of error in prediction of one or the other variables (classification, ratings), knowing both variables? The measure used to answer the first question is called λA (lambda A). The measure used to answer the second question is called λB. The measure used to answer the third question is called λAB. The first two are known as 'asymmetric' measures of predictive association; they are like regression coefficients. That is, λA may not be the same as λB. The latter is known as a 'symmetric' measure; it is similar in that respect to a simple correlation coefficient, in that the correlation of A with B is the same as the correlation of B with A. Table III-45 gives the asymmetric and symmetric indices of predictive association for different prectors and criteria.

The information in Table III-45 essentially confirms the information in Table III-44. There are differences between groups and grade levels in ratings of recommendations of last year's social studies class. The differences are statistically significant, but modest with respect to strength of association as measured either by correlation coefficients or by indices of predictive association. The difference in ratings is most pronounced between MACOS and non-MACOS 6th grade students, with MACOS 6th graders giving the most positive ratings compared to any other groups and non-MACOS 6th graders giving less positive ratings. Considering MACOS and non-MACOS 6th graders, if one knew only ratings students made, the reduction in error in predicting group membership would be 15%. That is the largest reduction of error of prediction of any

Table III-45

Percent Reduction of Error of Prediction of Recommendation Ratings from Knowledge of Group, or Group from Knowledge of Ratings

Group and Grade Level[1]	Prediction of Ratings from Group λ A	Prediction of Group from Ratings λ B	Joint Prediction λ AB
M5, M6	0%	8%	4%
NM 5, NM 6	0	7	3
M5, NM 5	1	12	5
M6, NM 6	1	15	7
M, NM	0	12	5
Grade 5, 6	0	1	1

1. M = MACOS; NM = non-MACOS; 5 = 5th grade; 6 = 6th grade.

combination of groups. However, the index of predictive association for predicting group membership at the 5th grade level is nearly as large (12%). It is therefore not surprising that the index for predicting group membership, (MACOS, non-MACOS) regardless of grade level, from a knowledge of ratings is also relatively high (12%).

From these analyses it is concluded, with respect to general opinions about last year's social studies class held by students a year after the class, that:

- the majority of students in all groups held positive opinions;
- former 5th and 6th grade MACOS students, on the average, had more positive opinions than non-MACOS students at those grade levels;
- the differences in opinions between grade levels of students within each group were smaller than differences between groups at each grade level;
- MACOS 6th grade students, on the average, held the most favorable opinions of all groups of students;
- Non-MACOS 5th grade students, on the average, held the least favorable opinions;
- differences between MACOS and non-MACOS students in degree of positive opinions were statistically significant but functionally small with respect to overall strength of association of opinion with group (i. e. -MACOS-, non-MACOS).

g. School and District Effects

1) Districts

How did districts compare with each other with respect to performance on various outcome measures? Were there interactions of district and curriculum? That is, did MACOS and non-MACOS classes perform differently in different districts? To examine those questions, a two-way analysis of covariance was done, with posttest measures as dependent variables, and district and group (MACOS, non-MACOS) as independent variables. The covariates were pretest and percentage of 5th grade students. The tests were univariate F-tests. The unit of analysis for each measure was the class mean. The results are summarized in Table III 46, which gives the p-values of the tests of main effects and interactions for each posttest variable.

It may be seen in Table III-46 that with respect to the MACOS vs. non-MACOS variable, results are similar to those obtained in earlier analyses of covariance and multiple regression (see Section III C 3b). That is, MACOS classes on the average were higher on posttest than non-MACOS classes on AP (the MACOS specific instrument). They were not significantly different from non-MACOS classes on other measures with the exception here of WWA and B, the questionnaire about attitudes towards customs or beliefs, and people. That difference will be discussed below.

There are district effects for three variables: AP, IDT, (the Interpretation of Data Test) and WWA. There are significant district-treatment interactions also for three posttest measures: STEP, SS Ch, (social studies choices), and CAPS-3

Table III-46
P-Values of Univariate F-Tests of MACOS Effects, District Effects and Interactions Between MACOS and District for Posttest Measures Adjusted for Pre-test and % 5th Graders

Posttest Variable	MACOS Effect (df = 1, 76)	District Effect (df = 14, 76)	Interaction (df =14, 76)
Animals and People (AP)[1]	.001***	.014**	.471
STEP+	.336	.541	.043*
IDT	.249	.002**	.575
SS Ch	.087	.845	.015*
WWA[2]	.009**	.039*	.081
WWB[2]	.046*	.676	.227
CAPS-1	.189	.103	.299
CAPS-2	.158	.796	.899
CAPS-3	.325	.096	.033*
CAPS-4	.116	.680	.844

+The error degrees of freedom for this measure are 73.

*$p \leq .05$
**$p \leq .01$
***$p \leq .001$

Note: As in most tables presenting p-values, those values that are .05 or less are denoted by asterisks as a visual aid in identifying variables of particular interest.

1. MACOS classes on the average scored higher than non-MACOS classes.
2. MACOS classes on the average had more positive reactions than non-MACOS classes.

(tolerance of ambiguity in problems). It may be noted that for those posttest variables for which there were significant main effects ($p \leq .05$), there were not significant interactions.

District effects were analyzed by application of Tukey's WSD test at the .05 level for least significant differences between all pairs of means. This procedure led to an identification of districts whose adjusted posttest means differed significantly. Such an identification provides the opportunity to examine relationships between significantly different districts and other variables, such as metropolitan status, size, policy, amount of implementation of MACOS, etc.

Such analyses have not been pursued in detail. It will simply be noted that the districts that were significantly different for the MACOS posttest (AP) were 12 (low) and 9 (high); for the Interpretation of Data Test (IDT), district 12 (low) differed from districts 6, 8, 9 and 10 (highs), and district 1 (low) differed from district 10. However, for WWA, the districts that differed were 4 and 8 (lows) and 15 (high). In effect, the same district that was low for IDT (#12) was also low for AP, while that was not the case for the other low district for IDT (#1). Furthermore, one of the districts that was relatively high for IDT (#8) was low for WWA. The district that was high on AP (#9), was near the top on IDT. It was about in the center of the distribution on WWA posttest. There was thus some consistency of standing of high and low districts within sets of measures (e.g. the achievement measures vs. the attitude measure), but not across sets.

The three variables for which there were significant interactions (STEP, SS Ch and CAPS-3) also of course indicate district effects. As before, examination

of adjusted contrasts (intercepts) showed that districts with the high and low contrasts for the STEP were not the same districts with high and low contrasts for the attitude measures (SS Ch and CAPS-3). Once again there appears to be an independence of achievement and attitude in the small sets of classes standing for different districts in this study.

The finding of significant differences for WWA and WWB requires further analysis. Earlier multivariate analyses, described in Sections III C 3b 1) and 2) showed evidence that MACOS classes, on the average, tended to react more positively on posttest than non-MACOS classes to WWA. However, there was no indication of a differentiation between groups on posttest on WWB.

There are several possible reasons for the disparity between the analyses of districts and earlier analyses of classes. One is that earlier analyses dropped several classes for which there were missing data, even considering analyses using only student-based PC's. The present analyses employ all 108 classes, except for STEP, for which N = 105. Thus, there may be sample variations that result in the difference. A second possibility is that the analyses of district effects used as covariates the percentage of 5th graders in classes and the class pretest mean of each posttest variable. The main analyses presented earlier used principal components (PC's) as covariates. Differences in covariates could influence results.

One means of investigating possible explanations is to compare the p-values of F-tests of analyses of covariance based on the same sample sizes and covariation, but not contrasting districts, with the p-values of analyses of covariance using principal components with similar sample sizes and different sample sizes.

to the district analyses. The p-values will reflect variation in means, variances regressions and sample sizes. Patterns of consistent or varying results can bear on choices among competing hypothesis. Table III-47 gives the p-values of F-tests of analyses of covariance of posttest variables for:

- MACOS vs. non-MACOS effects in district analyses (from Table III-47), N = 108;
- MACOS vs. non-MACOS, with pretest and % 5th graders as covariates, N = 108;
- MACOS vs. non-MACOS, with 8 student PC's - input, process and climate - as covariates (from Table III-47), N = 102;
- MACOS vs. non-MACOS, with 8 student PC's as adjusters (from Table III-47, N = 81.

It may be seen in Table III-47 that the two-way analyses of covariance (the MACOS by District analyses, Column 1) and the one-way analyses of covariance (Column 2) produced very similar results. The only marginal difference of interest is with WWB, which for the district analysis has a p-value of .046 and for the one-way analysis has a p-value of .058. Column 3, with nearly the same sample size as analyses represented in Columns 1 and 2, shows different results. Overall, the results in Column 3 (analyses of covariance using student-based PC's as covariates) are more similar to results in Column 4 (same covariates, reduced sample size) than to the results in Columns 1 and 2. WWA and B do not achieve significance at the .05 level when posttest class means in the two groups are adjusted for input, process and climate PC's. CAPS-1, on the other hand, is significant at the .05 level in Column 3, and nearly so in Column 4.

Table III-47

P-values of F-tests of MACOS Effects
in Analyses of Districts, and of Analyses of Covariance
Using Different Covariates and Sample Sizes

Posttest Variable	1 MACOS vs. Non-MACOS in District Analyses (N=108)[1]	2 MACOS vs. Non-MACOS Posttest Analyses (N=108)[2]	3 MACOS vs. Non-MACOS Using Student PC's (N=102)[3]	4 MACOS vs. Non-MACOS Using Student PC's (N = 81)[4]
AP	.001	.000	.000	.000
STEP	.336	.396	.310	.548
IDT	.249	.125	.244	.611
SS Ch	.087	.227	.367	.981
WWA	.009	.025	.161	.090
WWB	.046	.058	.509	.666
CAPS-1	.189	.215	.035	.065
CAPS-2	.158	.100	.720	.538
CAPS-3	.325	.225	.839	.889
CAPS-4	.116	.527	.295	.300

1. Covariates: pretest and % 5th graders. (Note: for STEP, N = 105).
2. Covariates: pretest and % 5th graders. (Note: for STEP, N = 105).
3. Covariates: student input, process and climate PC's.
4. Covariates: student input, process and climate PC's.

It will be recalled from Section III C 3b 1) above that WWA showed heterogeneity of regression with Climate for both the large and reduced sample (Figures III-2 and 3).[84] It will also be recalled that WWA and WWB both showed significant increments in proportion of variance associated with process variables in multiple regression analyses entering Group last; CAPS-1 showed a significant increment associated with Climate (See Table III-17; also Table III-20, in which Group was entered after input). STEP was also associated with Climate, as was SS Ch (social studies choices). However, unlike WWA, WWB and CAPS-1, for which only 52-67% of the total posttest variance accounted for was associated with input, 97% of the STEP variance accounted for was associated with input. Thus, there was little variance left to account for in STEP once input (primarily Ach, or pretests) was removed.

It thus appears that the differences between results in Columns 1 and 2, on the one hand, and Columns 3 and 4 on the other, may be viewed as attributable primarily to the use of different covariates rather than to variations in sample sizes. This does not mean that variations in sample sizes (and therefore sample compositions) have no influence on results. They can and do effect results. The interpretation here is that for the variations under consideration, method of analysis appears to be the more potent factor.

It is of interest to note that the posttest variables that were influenced importantly by changes in method of analysis (or in sample sizes or both) are variables in the attitude set, not in the achievement set. No major conclusions

84. It may be noted that there were no heterogeneous regressions for any analysis in Columns 1 and 2 in Table III-47.

about differences between MACOS and non-MACOS groups of classes with respect to AP, STEP and IDT have been influenced by method of analysis, including which variables have been used as statistical controls.

2) Schools

A design goal of the study was to have only one class (MACOS or non-MACOS) in a particular school building. The resultant samples made that goal unattainable, as described in Section IIA. There were 108 classes in 76 different schools, or an average of 1.4 classes per school. Table III-48 shows for each group the number of schools with 1, 2, 3 or 4 classes per school. With the exception of two pairs of classes in two schools, the classes that were grouped in a given school were either MACOS classes or non-MACOS classes, but not both types in the same school. The question of primary concern is: has the nesting of some classes in schools influenced main results?

To explore that question, it was hypothesized that if the posttest and follow-up measures for which there pretests were averaged for the classes that were in the same building, an analysis of covariance of posttest and outcome measures of all available classes would show results similar to results of other analyses. The covariates for these analyses, as with the analyses of districts, were pretest (averaged for classes within a given building) and % 5th graders (averaged for the same classes). The sample size after averaging and after deletions for missing data, was 74 (MACOS = 38, non-MACOS = 36).

Table III-49 gives the p-values of the F-tests of the analyses of covariance of independent classes for each posttest and Follow-up 2 measure, and also for three of the analyses whose p-values were listed in Table III-47.

Table III-48
Number of Schools With One Class Per School
and With More than One Class, by Group

Number of Classes Per School	MACOS	Non-MACOS
1	29 (51%)*	27 (53%)*
2	4 (14)	10 (39)
3	4 (21)	0 (0)
4	2 (14)	1 (8)

*Percentage of total classes in group.

The reason for repeating the figures from that table is for ease of comparison with the results for the analysis of classes averaged within schools.

It may be seen in Table III-49 that the p-values of tests for classes averaged within schools (Column 1) follow quite closely the results of analyses of covariance of posttests (Column 2), as did the MACOS effects in the district analyses. The most striking difference is that WWA and WWB show a stronger treatment effect when independence of units (classes) is brought about (within the limits of nesting in districts). It may be seen that the similarity of results extends to the Follow-up 2 measures as well as to posttests. It follows from the similarity of results in Column 1 to those in Column 2 that Column 1 results differ from those in Colum 3 and 4 (in which student PC's were the covariates) in predictable ways.

The analyses of covariance for which p-values of F-tests are shown in Column 1 of Table III-49 showed one noteworthy difference from those in Column 2, and a similarity to those in Columns 3 and 4. The analysis of WWA, when classes within schools were averaged, showed a heterogeneity of regression, as it did in the analyses using student PC's. However, when each covariate for the classes averaged analysis (Column 1) was tested, neither covariate alone, (pretest, % 5th graders) showed significant heterogeneity.[85]

85. For the homogeneity of slopes test for pretest: $F_{1,70} = 2.453$ $p \leq .12$; for % 5th graders, $F_{1,70} = .750$, $p \leq .39$.

Table III-49

P-values of F-tests of Analyses of Covariance With Classes in Schools Averaged, and of Analyses of Covariance Using Different Sample Sizes and Covariates

Outcome Variables	1 MACOS vs. Non-MACOS, Classes in Schools Averaged (N=74)[1]	2 MACOS vs. Non-MACOS Posttest Analyses (N=108)[2]	3 MACOS vs. Non-MACOS Using Student PC's (N=102)[3]	4 MACOS vs. Non-MACOS Using Student PC's (N=81)[4]
AP	.000	.000	.000	.000
STEP	.766	.396	.310	.548
IDT	.100	.125	.244	.611
SS Ch	.243	.227	.367	.981
WWA	.002	.025	.161	.090
WWB	.003	.058	.509	.666
CAPS-1	.681	.215	.035	.065
CAPS-2	.115	.100	.720	.538
CAPS-3	.611	.225	.839	.889
CAPS-4	.428	.527	.295	.300
Follow-up 2				
AP1-4F	.019	.011	.003	.013
SS Ch F	.698	.318	.293	.309
WWAF	.571	.366	.434	.701
WWBF	.181	.210	.041	.110

1. Covariates: pretest and % 5th graders.
2. Covariates: pretest and % 5th graders (Note: N's for Follow-up 2 are 98 except for (SS Ch, for which N=97).
3. Covariates: student input, process and climate PC's.
4. Covariates: student input, process and climate PC's.

As a further means of examining the effects of some classes being in the same schools, a reanalysis of the multivariate ANCOVA was done. That analysis, it will be recalled, tested the difference between MACOS and non-MACOS classes on 20 criterion variables simultaneously, using all student PC's as covariates. In this analysis, PC's of classes within the same schools were averaged, as well as outcome variables. The question of interest was: to what extent were major findings replicated?

Table III-50 in Column 1 gives the results in terms of p-values of the univariate tests of each dependent variable. The overall difference between groups continued to be significant: $F_{20,37}=2.252$, $p \leq .016$.[86] Table III-50 also gives for purposes of comparison, the p-values of univariates F-tests using:

1. Posttest scores adjusted for pretest and % 5th graders;
2. Posttest and follow-up scores using student PC's as covariates (larger sample);
3. Posttest and follow-up scores using student PC's as covariates (reduced sample).

Columns 3 and 4 give p-values for analyses similar to those using classes within schools averaged (Column 1); only for different sample-sizes and classes within schools not averaged. Column 2 is repeated from Table III-49, Column 2.

It is of interest to note in Table III-50 that most main results are repeated by the analysis described above. AP, AP1-4F, WWA and interest continue to show significant differences the two groups of classes. Variables such as CAPS-1 and

86. For this analysis, there were 35 MACOS classes and 31 non-MACOS classes.

Table III-50

P-values of Univariate F-tests of Analyses of Covariance of Classes Within Schools Averaged, Using Student PC's as Covariates

Outcome Variables	1 MACOS vs. Non-MACOS Classes Within Schools Averaged Using Student Background and Process Climate PC's (N=66)	2 MACOS vs. Non-MACOS Posttest Scores Adjusted for Pretest and % 5th Graders (N = 108)	3 MACOS vs. Non-MACOS Using Student PC's (N=97)	4 MACOS vs. Non-MACOS Using Student PC's (N=81)
AP	.001***	.000***	.001***	.001***
STEP	.972	.396	.277	.528
IDT	.293	.125	.346	.605
SS Ch	.584	.227	.497	.916
WWA	.040*	.025*	.079	.018*
WWB	.179	.058	.595	.412
CAPS-1	.106	.215	.064	.063
CAPS-2	.940	.100	.795	.411
CAPS-3	.756	.225	.619	.895
CAPS-4	.119	.100	.103	.446
Know	.055		.046**	.009**
Skills	.142		.011**	.115
Interest	.039*		.013*	.056
AP1-4F	.010**	.011*	.002**	.029*
SS Ch F	.510	.218	.293	.600
SS	.936		.459	.581
WWAF	.812	.366	.296	.978
WWBF	.174	.210	.030*	.318
WWAPF	.308		.423	.235
WWBPF	.973		.504	.940

*p ≤ .05
**p ≤ .01
***p ≤ .001

Note: as in most tables presenting p-values, those values that are .05 or less are denoted by asterisks as a visual aid in identifying variables of particular interest.

and Know approach significance at the .05 level but fall short of it. WWBF does not approach significance in this analysis.

Table III-51 gives the means and standard deviations of student PC's and outcome variable of MACOS and non-MACOS groups of classes for the three main multivariate analyses of covariance whose results are shown in Table III-50 (Columns 1, 3 and 4). These data are given as a further basis for assessing the comparability of results using different samples of classes from the two groups. In Table III-51 it may be seen that there are variations in means and deviations for different variables for different sample sizes. In most cases the variations do not appear to be severe. Particularly encouraging i he relative consistency of the PC's for the three major samples shown. With the exception of the MACOS Class 2 mean for the reduced group (N=81; the group of classes with complete student and teacher data sets), input PC means for both groups do not differ significantly from zero. One of the main consequences of losing classes that may be seen in some of the criterion or outcome variables is the reduction in variance (e.g., in the MACOS groups, SD's in the three samples for WWA range from .93 to .77; they stay essentially constant in the non-MACOS samples in this case). Overall, Table III-51 gives the appearance of substantial consistency in means and standard deviations among covariates, and among the achievement related measures (AP, STEP, IDT) for the three samples of classes. There is more variation in these two statistics for process and climate covariates, and for some attitude measures.

3) Implications

The overall results of district and school analyses are taken here as encouraging. Even with drastically (and non-specified) reduced sample sizes,

Table III-51

Means and Standard Deviations of Covariates (Student PC's) and Unadjusted Outcome Variables of MACOS and Non-MACOS Groups of Classes, by Samples Analyzed

	Large Group(N=97)				Reduced Group(N=81)				Classes Averaged(N=66)			
	MACOS		Non-MACOS		MACOS		Non-MACOS		MACOS		Non-MACOS	
Variable	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
1. PC's												
Ach	.07	1.62	.01	1.59	.12	1.51	.01	1.53	-.03	1.67	.03	1.45
Att 1	.14	1.26	-.01	1.58	.12	1.28	-.06	1.49	.09	1.16	.07	1.56
Att 2	-.00	1.07	.08	1.10	-.04	1.07	.20	1.03	-.12	1.02	.17	1.09
Class 1	.02	1.58	.11	1.61	.05	1.59	-.00	1.58	-.01	1.70	.08	1.62
Class 2	.19	1.21	-.01	1.58	.26	1.14	.20	1.51	.24	1.16	.06	1.55
S Proc 1	.27	1.83	-.17	1.73	.14	1.75	-.01	1.84	.03	1.72	-.13	1.80
S Proc 2	-.48	1.68	.51	1.57	-.41	1.66	.59	1.58	-.56	1.41	.47	1.48
Climate	.44	1.34	-.57	1.68	.44	1.38	-.64	1.69	.48	1.24	-.80	1.65
2. Outcomes												
AP	.41	.94	-.42	.80	.40	.89	-.46	.82	.33	.94	-.47	.80
STEP	.03	1.03	.04	.89	.09	.90	.03	.89	.02	.95	.02	.83
IDT	.12	.98	-.05	.94	.14	.94	-.01	.95	.16	.89	-.00	.85
SS Ch	.15	.88	-.17	.96	.06	.84	-.18	.99	.24	.83	-.23	1.04
WWA	.17	.93	-.28	1.02	.19	.89	-.29	1.04	.32	.77	-.29	1.03
WWB	.16	.90	-.16	1.08	.23	.77	-.07	1.06	.34	.74	-.17	1.01
CAPS-1	-.15	1.05	.14	.88	-.20	.98	.09	.82	-.26	1.01	-.00	.79
CAPS-2	.14	1.13	-.04	.82	.17	1.20	-.09	.87	.07	.95	-.11	.84
CAPS-3	.10	1.06	-.03	.81	.13	1.14	-.09	.79	-.13	.91	-.06	.80
CAPS-4	.04	.92	-.19	.86	-.01	.85	-.08	.87	.00	.77	-.13	.79
Know	-.29	.90	.33	1.06	-.31	.90	.34	1.13	-.40	.80	.28	1.02
Skills	.15	1.08	-.15	.86	.10	1.11	-.12	.86	.22	1.08	-.08	.88
Interest	.29	.90	-.38	.93	.34	.89	-.39	.93	.47	.91	-.28	.88
AP1-4F	.22	.91	-.28	1.03	.21	.95	-.25	1.01	.15	.89	-.22	.95
SS Ch F	-.07	.99	.09	1.03	-.08	1.00	.04	.97	-.20	.93	-.07	.98
SS	-.04	.91	-.09	1.09	.04	.86	-.12	1.03	-.12	.78	-.33	1.06
WWAF	.08	.95	-.16	1.04	.12	.94	-.10	.96	.10	.89	-.02	.86
WWBF	.14	.85	-.24	1.14	.20	.85	-.15	.95	.20	.75	-.07	.90
WWAPF	.06	.94	-.14	1.12	.17	.89	-.13	.87	.08	.76	-.11	.82
WWBPF	.03	.93	-.16	1.10	.05	.87	-.08	1.07	-.01	.88	-.16	.92

one continues to see results that for the most part conform to (or are similar to) results obtained otherwise. Means and standard deviations of criterion variables and covariates remained similar, although these are not shown here.

With respect to districts, there is indication of an interaction between treatment (MACOS, non-MACOS) with respect to posttest class means for STEP, Social Studies Choices (SS Ch) and CAPS-3 (Tolerance for Ambiguity). None of these three measures showed any consistent main effect of treatment. But there is indication that results one may obtain using those variables as measures of performance with MACOS and other social studies courses may differ in different districts. As noted above, however, detailed analyses of the characteristics of districts and particular programs within them have not been undertaken.

The results of analyses of differences between MACOS and non-MACOS classes when independence of classes is achieved by analyzing only classes (and averaged within-school classes) in different buildings support the hypothesis that results obtained by other analyses were not unduly biased by the inclusion of classes nested in the same building. They do not do so unequivocally, however. There is the alternative hypothesis, for example, that changes in results from other analyses could also be the result of changes in the characteristics of samples. The relative stability or similarity of main results with different sample sizes, however, is taken as encouragement for interpretations made of earlier analyses.

IV. RELATIONSHIPS BETWEEN INPUT, PROCESS, CLIMATE AND OUTCOME VARIABLES

A. Canonical Variate Analyses

In order to examine relationships between sets of variables more extensively a series of canonical correlation analyses was performed. The objectives of these analyses were to be able to examine relationships between sets of input and process variables to sets of outcome variables, and to each other, and to identify possible variables within sets that appeared related to significant canonical variates. Thus, for example, if there were significant canonical correlations between input and process, it would be of interest to see which predictor variables (input PC's) correlated highly with the predictor canonical variate. Similarly, it would be of interest to know which process and climate PC's correlated highly with the related criterion canonical variate. Such information would bear on the question of what seem to be relationships of input characteristics to process characteristics.[1]

The following analyses were made:

- Input and Process/Climate ⟶ Outcomes 1, 2 and 3[2]
- Input ⟶ Outcomes 1, 2, 3
- Input ⟶ Process/Climate
- Process/Climate ⟶ Outcomes 1, 2, 3

In each analysis, canonical relationships were computed for the total sample, MACOS only, and non-MACOS only.

---

1. The question of interactions is not addressed by such analyses. The canonical correlation model gives that weighted linear combination of variables that best predicts a weighted linear combination of criterion variables.
2. Outcomes 1, 2, and 3 are, respectively, posttest, Follow-up 1, Follow-up 2.

The main results are summarized in Table IV-1, 1-4. This series of tables shows, for each set of analyses, the significant canonical correlations for each set of analyses for the whole sample and for MACOS and non-MACOS classes separately. It can be seen that there were typically two significant pairs of canonical correlations for the total group and that there was never more than one for MACOS and non-MACOS. The structure of the predictor variate is given at the top of each sub-table; the criterion variate structure is at the bottom. The numbers given for each variable are the correlations of the variables with the canonical variate, i.e., the canonical variable loadings.

It should be noted from the outset that caution is needed in generalizing from these analyses. A cross validation study of these variables could result in several changes: the PC's could change; correlations between them could change; the canonical correlations could drop to non-significance; all could happen. Canonical correlations, like other correlations, are subject to sampling errors, and especially with small samples to overfitting that produces high correlations that shrink upon cross validation.[3] It is unlikely, however, that canonical correlations involving pre and post achievement tests would shrink to non-significance. But correlations among other predictors could change, thus changing the structures of predictor and criterion variates.

The consequence of this brief discussion is that while statements will be made about the data in Table IV-1, they will be in the nature of hypotheses rather

3. For a recent discussion of the problems, and an approach to dealing with them, see Weinberg, Sharon L., and Darlington, Richard B., Canonical Analysis When Number of Variables is Large Relative to Sample Size. Journal of Educational Statistics, 1976. Vol. 1, 4, 313-332.

than conclusions from cross validated results. Further, as an arbitrary ground rule, no attempt will be made to comment on variable loadings of less than ± .50 for these small samples of classes. In addition, no comment will be made about the second pair of canonical variates (CV-2) for the total group.

In Table IV-1,1a, it can be seen that for the total group, there were two significant canonical correlations, each with its pair of canonical variates. In looking at the loadings (correlations) of predictor PC's with the first predictor variate, (the input/process variate under the first canonical variate, CV-1) it appears that the maior variables are pre-achievement (Ach), pre-achievement related attitude (Att 1), and older, more affluent students (Class 1). On the 1st year outcomes side of CV-1, the major variables loading on the canonical variate are the post-achievement measures (AP, STEP, and IDT), as well as CAPS-3 (tolerance of ambiguity). In effect, when all input and process variables (PC's) are related to posttest variables, the first canonical variate (CV-1) appears to be dominated by achievement-related variables.

Looking across at the MACOS and non-MACOS group in Table IV-1,1a, one can see that only the first CV was significant. The predictor variate structure, related to total group CV-1, for both groups is similar, although S -Proc 1 appears to take on a little more significance in the MACOS group. On the outcome side for both groups, achievement measures have the highest loadings. But in the non-MACOS group, CAPS-1 and 2 appear to load on the variate as well. In effect, for the non-MACOS group, posttest performance, on achievement measures and self attitude scales appear more generalized.

Table IV-1, 1b shows relationships of input and process to Follow-up 1 outcomes. The relationships are weak; the canonical correlations are relatively small, and for the non-MACOS group, not significant at the .05 level. In the MACOS group of classes there is the indication that the less teacher emphasized individual work (T Proc 2) the more students subsequently felt they missed learning some content and skills that would have been useful to them in their present social studies class. But they also found the present class less interesting.[4]

In Table IV-1, 1c, which shows input and process variables in relation to Follow-up 2 outcomes, there appears once again to be in CV-1 a generalized relationship of pre-achievement and achievement related attitude (Ach, Att 1), and age and affluence of the class (Class 1) to follow up performance on the MACOS course related instrument (AP1-4F) and the two more positive attitudes toward customs and people measures (WWAF and WWBF). The relative influence of informal, relaxed classes that students perceived as not emphasizing particular types of activities (S Proc 1) is noticeable in the predictor variate (given the .50 criterion) in the non-MACOS group of classes, but not in the MACOS group.

Table IV-1, 2a-c shows relationships of input and outcome variables. The canonical variate structures strongly parallel those found when process and climate PC's were included as predictors. The implication is that, particularly for achievement measures, or attitude measures that are correlated with achievement ones, pretest tends to be the predominant predictor of outcome. This was noted in other analyses described in Section III. Again it can be noted

4. It will be recalled the Skills, Know and Interest were all scaled such that a low value was positive and a high value negative.

that the first class PC (Class 1) is strongly related to the predictor variate, along with Ach and Att 1.

Table IV-1, 3 suggests relationships between initial characteristics of classes and teachers on the one hand, and what teachers did or emphasized, as reported by teachers (T Proc 1, 2) and perceived by students (S Proc 1, 2 and Climate). It also suggests that some interesting differences between these MACOS classes as a group and the non-MACOS group (which was, it should always be remembered, a conglomeration of various curricula). Taken at face value, the overall group pattern leads to the following statement of relationships: the older, more affluent and higher achieving the class (Ach, Class 1) and the less traditional the teacher (T Psy 2 negative), the more informal, non-grade stressed, group and discussion oriented the class (S Proc 1) as rated by students, and the better the perceived climate (higher satisfaction, less apathy, less perceived difficulty of the work).

There are suggestions of variations in this pattern between the MACOS and non-MACOS group of classes. In the MACOS group, the most heavily loading variables are Class 1 (older, more affluent classes) and younger, more non-traditional teachers (T Demo, T Psy 2 negative). In the non-MACOS classes, the structure suggests that pre-achievement level of classes (Ach), and pre-achievement related attitude (Att 1) load noticeably. Traditionalism also loads negatively as in the MACOS classes, but years of teaching experience (T Demo, which can be taken as a proxy of age) has no relationship to the variate. Class 1 does not load noticeably on the predictor variate in the non-MACOS classes as a group, but does

in the MACOS group. Whether students on the average perceived the class as informal and group oriented (S Proc 1) correlates strongly with the process/climate variate in the MACOS group, but not in the non-MACOS classes. Good classroom climate was the most highly correlating variable in the non-MACOS classes under consideration here, with lack of individual grades emphasizing work also of importance.

In Table IV-1, 4a, b and c, relationships between process/climate and outcomes are shown. These are correlations that have not had input partialled out (i.e. held constant). They thus do not uniquely depict relationships with input held constant.

Table IV-1, 4a shows relationships to 1st year outcomes (posttests). In the total group the structure of relationships in the first pair of canonical variates (CV-1) suggests that emphasis by the teacher and the curriculum, as rated by the teacher, on group activities, on affect or feelings, and higher order cognitive skills (T Proc 1) and good classroom climate, as rated by students, are influential as predictors. On the outcome side the structure of the variate is similar to that found when both input and process variables were included (Table IV-1, 1a) with one striking exception. SS Ch (attitude towards social studies) here loads highly on the outcome variate. The correlation between sets of variables for the MACOS group of classes was not significant, and therefore is not interpreted here. But the relationship was significant for the non-MACOS classes. There it can be seen that for that group of classes it is the non-achievement related outcome variables, SS Ch, CAPS-1 (perceived ability of self as problem solver), and CAPS-2 (interest in problem solving), that correlate more

strongly with the outcome variate.

The marked differences in structures of the first outcome variate between the total group and the non-MACOS sub-group suggests as one possibility that the relationships of process and climate characteristics to achievement and attitude measures operated differently in the non-MACOS classes than in the MACOS classes.

Table IV-1, 4b suggests that for the total MACOS and non-MACOS group, good climate characteristics and relative absence of individual, recall-oriented instruction (S Proc 2, T Proc 2) are predictive of finding next year's social studies class less interesting (Interest, positive). In the MACOS group, marked lack of emphasis on individual work and remembering (T Proc 2, negative), was apt to be associated also with samples of students from these classes feeling, on the average, that they had not been adequately prepared in certain skills and knowledge that would have been helpful the following year (Skills and Know, positive).

Finally, Table IV-1, 4c suggest that while Climate may not have had predictive relationship to Follow-up 2 outcomes in the MACOS classes, poor classroom climate in the non-MACOS classes and emphasis on individual, non-group oriented work (T Proc 2) appears predictive of negative attitudes a year later towards social studies (SS Ch F, SS) and towards the behaviors or beliefs described in WWAF. In the MACOS classes, S Proc 1 is predictive of performance on the MACOS specific instrument (AP1-4F). That relationship is not found in the non-MACOS group.

Table IV-1, 1a.

Summaries of Canonical Correlations and Canonical Variate-Correlations by Analyses and Group

Analysis	TOTAL GROUP CV-1 Correlations Rc 1 = .92***	TOTAL GROUP CV-2 Correlations Rc 2 = .70**	MACOS CV-1 Correlations Rc 1= .92*	NON-MACOS CV-1 Correlations Rc 1 = .97***
	Input/Process	Input/Process	Input/Process	Input/Process
1. Input and Process vs. Outcome				
a. First year outcomes (posttest)				
	Ach .94	-.25	.94	.89
	Att 1 .71	.21	.56	.72
	Att 2 .20	.50	.20	.31
	Class 1 .66	-.17	.63	.63
	Class 2 .15	.14	.07	.19
	T Demo -.10	.10	-.21	.05
	T Psy 1 -.12	-.03	.05	-.28
	T Psy 2 -.17	-.26	-.02	-.22
	S Proc 1 .48	-.43	.56	.42
	S Proc 2 -.09	-.42	-.08	.06
	T Proc 1 .18	.39	-.07	.36
	T Proc 2 -.03	.06	.25	-.24
	Climate .32	.61	.17	.35
	1st Yr Outcomes	1st Yr Outcomes	1 Yr Outcomes	1st Yr Outcomes
	AP .74	-.07	.70	.79
	STEP .96	-.13	.94	.95
	IDT .81	-.19	.81	.73
	SS Ch .16	.72	-.09	.25
	WWA .35	.26	.32	.22
	WWB .42	.40	.38	.36
	CAPS 1 .38	.26	.29	.56
	CAPS 2 .31	.59	.19	.51
	CAPS 3 .73	.26	.65	.79
	CAPS 4 .18	.35	.08	.34

*p≤.05
**p≤.01
***p≤.001

NS = Not significant.

Note: Sample sizes in all analyses in Tables IV-1, 1-4 are: Total Group, 81, MACOS 45, non-MACOS, 36.

Table IV-1, 1b.

Summaries of Canonical Analyses

Analysis	TOTAL GROUP CV-1 Correlations Rc 1 =.54*	TOTAL GROUP CV-2 Correlations Rc:NS	MACOS CV-1 Correlations Rc 1=.76*	NON-MACOS CV-1 Correlations Rc:NS
1. Input and Process vs. Outcomes b. Follow-up 1	Input/Process	Input/Process	Input/Process	Input/Process
	Ach .11	-.02	-.05	.41
	Att 1 .08	.03	-.01	.19
	Att 2 .27	-.10	.25	.07
	Class 1 .09	.17	-.01	.30
	Class 2 -.07	-.29	-.31	.21
	T Demo .02	-.31	-.11	.41
	T Psy 1 -.20	.11	-.14	-.52
	T Psy 2 .16	-.29	.29	-.33
	S Proc 1 .32	-.03	.23	.39
	S Proc 2 -.41	-.73	.18	.40
	T Proc 1 -.17	.26	-.15	-.13
	T Proc 2 -.48	-.55	-.52	-.20
	Climate .26	.56	.20	.05
	FU-1 Outcomes		FU-1 Outcomes	FU-1 Outcomes
	Skills .68	-.29	.56	.10
	Know .51	-.80	.60	.86
	Interest .43	.88	.58	-.21

*$p \leq .05$

NS = not significant. Correlations with variate shown, but offset to right.

Table IV-1, 1c.

Summaries of Canonical Analyses

Analysis	TOTAL GROUP CV-1 Correlations Rc 1 = .79***	TOTAL GROUP CV-2 Correlations Rc 2 = .59*	MACOS CV-1 Correlations Rc 1 = .83**	NON-MACOS CV-1 Correlations Rc 1 = .89**
1. Input and Process vs. Outcomes. c. Follow-up 2	Input/Process	Input/Process	Input/Process	Input/Process
	Ach .88	-.13	.80	.75
	Att 1 .63	.04	.47	.50
	Att 2 -.12	-.00	-.34	-.05
	Class 1 .78	-.22	.78	.73
	Class 2 -.01	.04	.01	-.03
	T Demo .01	.35	-.22	.02
	T Psy 1 -.17	.16	.04	-.06
	T Psy 2 -.34	-.30	-.04	-.34
	S Proc 1 .45	-.51	.38	.71
	S Proc 2 -.04	-.07	.05	.42
	T Proc 1 .10	.41	-.14	-.04
	T Proc 2 -.08	-.15	.17	-.31
	Climate .21	.53	-.25	.07
	FU-2 Outcomes	FU-2 Outcomes	FU-2 Outcomes	FU-2 Outcomes
	AP-F .92	-.24	.85	.78
	SS Ch-F .04	.01	.06	-.02
	WWAF .51	.27	.57	.41
	WWBF .68	.27	.65	.53
	WWAPF -.07	-.21	.48	-.66
	WWBFP .17	.26	.40	.08
	SS .11	.63	-.13	.01

*p ≤ .05
**p ≤ .01
***p ≤ .001

Table IV-1, 2a.

Summaries of Canonical Analyses

Analysis	TOTAL GROUP CV-1 Correlations Rc 1=.92***	TOTAL GROUP CV-2 Correlations Rc 2=.61*	MACOS CV-1 Correlations Rc 1 =.90**	NON-MACOS CV-1 Correlations Rc 2 = .97***
2. Input vs. Outcomes a. First Year Outcomes (Posttest)	Inputs	Inputs	Inputs	Inputs
	Ach .95	-.28	.96	.91
	Att 1 .72	.23	.64	.72
	Att 2 .17	.83	.16	.27
	Class 1 .68	-.31	.64	.66
	Class 2 .13	.12	.01	.18
	T Demo -.07	.07	-.17	.08
	T Psy 1 -.10	-.03	.04	-.28
	T Psy 2 -.16	-.25	.02	-.22
	1st Yr Outcomes	1st Yr Outcomes	1st Yr Outcomes	1st Yr Outcomes
	AP .73	-.05	.73	.80
	STEP .95	-.14	.92	.94
	IDT .80	-.14	.81	.71
	SS Ch .11	.48	-.08	.20
	WWA .41	-.05	.44	.28
	WWB .41	.21	.40	.35
	CAPS-1 .36	.21	.30	.55
	CAPS-2 .28	.65	.17	.48
	CAPS-3 .71	.45	.64	.78
	CAPS-4 .18	.64	.04	.33

*$p \leq .05$
**$p \leq .01$
***$p \leq .001$

Table IV-1, 2b.

Summaries of Canonical Analyses

Analysis	TOTAL GROUP CV-1 Correlations Rc 1: NS	TOTAL GROUP CV-2 Correlations Rc 2: NS	MACOS CV-1 Correlations Rc 1: NS	NON-MACOS CV-1 Correlations Rc 1: NS
2. Input vs. Outcomes b. Follow-up 1	Note: Correlation of variables with canonical variates not listed because the canonical correlations did not reach significance.			

NS = not significant.

Table IV-1, 2c.

Summaries of Canonical Analyses

Analysis	TOTAL GROUP CV-1 Correlations Rc 1 = .76***	TOTAL GROUP CV-2 Correlations Rc 2 = .49*	MACOS CV-1 Correlations Rc 1 = .82***	NON-MACOS CV-1 Correlations Rc 2 = .78*
2. Input vs. Outcomes c. Follow-up 2	Inputs	Inputs	Inputs	Inputs
	Ach .91	.06	.88	.88
	Att 1 .64	.15	.54	.64
	Att 2 -.11	-.24	-.26	.03
	Class 1 .78	-.05	.80	.82
	Class 2 .01	-.13	-.05	.05
	T Demo -.15	-.03	-.17	-.19
	T Psy 1 -.15	.80	.08	-.10
	T Psy 2 -.34	-.25	-.08	-.36
	FU-2 Outcomes	FU-2 Outcomes	FU-2 Outcomes	FU-2 Outcomes
	AP(1-4)F .94	-.29	.90	.84
	SS Ch F .08	-.10	.12	.06
	SS .10	.04	.02	.02
	WWAF .54	.66	.59	.64
	WWBF .62	.24	.63	.72
	WWAPF -.04	-.42	.32	-.31
	WWBPF .15	.27	.33	.09

*p ≤ .05
**p ≤ .01
***p ≤ .001

Table IV-1, 3

Summaries of Canonical Analyses

Analysis	TOTAL GROUP CV-1 Correlations Rc 1 = .67***	TOTAL GROUP CV-2 Correlations Rc 2 : NS	MACOS CV-1 Correlations Rc 1 = .74**	NON-MACOS CV-1 Correlations Rc 1 = .85**
3. Input vs Process	Inputs	Input	Inputs	Inputs
	Ach .59	.59	.41	.51
	Att 1 .48	.38	.26	.51
	Att 2 .17	.02	.23	.48
	Class 1 .61	.61	.61	.34
	Class 2 .07	-.00	-.24	-.09
	T Demo -.37	-.08	-.51	-.09
	T Psy 1 -.08	-.23	.13	-.06
	T Psy 2 -.74	.64	-.67	-.52
	Process	Process	Process	Process
	S Proc 1 .76	.44	.82	.31
	S Proc 2 -.38	.32	-.46	-.05
	T Proc 1 .39	-.05	.43	.25
	T Proc 2 -.46	.76	-.25	-.49
	Climate .59	-.40	.51	.79

*$p \leq .05$
**$p \leq .01$
***$p \leq .001$

NS = not significant. Correlations with variate shown, but offset to right.

Summaries of Canonical Analyses

Analysis	TOTAL GROUP CV-1 Correlations Rc 1 = .67***	TOTAL GROUP CV-2 Correlations Rc 2 = .59**	MACOS CV-1 Correlations RC 1: NS	NON-MACOS CV-1 Correlations Rc 1 = .83***
4. Process vs. Outcomes	Process	Process		Process
a. First Year (Posttest)	S Proc 1 .15	.97	.53	-.47
	S Proc 2 -.37	.13	-.39	-.30
	T Proc 1 .53	-.16	.47	.51
	T Proc 2 -.08	-.19	.13	.01
	Climate .89	-.08	.83	.82
	1st Yr Outcomes	1st Yr Outcomes	1st Yr Outcomes	1st Yr Outcomes
	AP .55	.55	.50	.08
	STEP .52	.60	.58	.30
	IDT .39	.60	.49	.07
	SS Ch .79	-.29	.58	.65
	WWA .26	-.10	-.14	.40
	WWB .53	.05	.36	.44
	CAPS-1 .43	-.12	.25	.57
	CAPS-2 .45	-.20	.23	.61
	CAPS-3 .54	.58	.70	-.00
	CAPS-4 .22	-.28	.03	.39

*$p \leq .05$
**$p \leq .01$
***$p \leq .001$

NS = Not significant. Correlations with variate shown, but offset to right.

Table IV-1, 4 b

Summaries of Canonical Analyses

Analysis	TOTAL GROUP CV-1 Correlations Rc 1 =.45***	TOTAL GROUP CV-2 Correlations Rc 2 =.38*	MACOS CV-1 Correlations Rc 1 =.51*	NON-MACOS CV-1 Correlations Rc: NS
4. Process vs Outcomes b. Follow-up 1	Process	Process	Process	Process
	S Proc 1 .18	.54	.43	.24
	S Proc 2 -.80	.57	.15	.83
	T Proc 1 .20	-.52	-.35	-.58
	T Proc 2 -.65	-.49	-.84	.37
	Climate .77	.16	.47	-.29
	FU-1 Outcomes	FU-1 Outcomes	FU-1 Outcomes	FU-1 Outcomes
	Skills -.19	.83	.73	.93
	Know -.45	.72	.59	.64
	Interest .97	.23	.60	.08

*$p \leq .05$
**$p \leq .01$
***$p \leq .001$

NS = not significant. Correlations with variate shown, but offset to right.

Table IV-1, 4 c.

Summaries of Canonical Analyses

Analysis	TOTAL GROUP CV-1 Correlations Rc 1=.49**	TOTAL GROUP CV-2 Correlations Rc : NS	MACOS CV-1 Correlations Rc 1 = .68**	NON-MACOS CV-1 Correlations Rc 2 = .65*
4. Process vs. outcomes c. Follow-up 2	Process	Process	Process	Process
	S Proc 1 .79	.31	.71	-.13
	S Proc 2 .06	-.37	-.19	.03
	T Proc 1 -.42	.41	-.26	.08
	T Proc 2 .13	-.18	.35	.82
	Climate -.18	.92	.21	-.70
	FU-2 Outcomes	FU-2 Outcomes	FU-2 Outcomes	FU-2 Outcomes
	AP(1-4)F .62	.71	.80	-.18
	SS Ch F -.10	.38	.14	-.48
	SS -.55	.70	-.13	-.61
	WWAF -.06	.34	-.13	-.54
	WWAB -.03	.32	-.25	-.19
	WWAPF -.08	-.14	.02	.23
	WWBPF -.01	.18	-.12	-.36

*$p \leq .05$
**$p \leq .01$
***$p < .001$

NS = Not significant.