DOCUMENT RESUME

ED 218 165 SO 014 034

AUTHOR Napier, John D.; Hepburn, Mary A.
TITLE Evaluation of a Locally Developed Social Studies Curriculum Project: Improving Citizenship Education.
PUB DATE 82
NOTE 36p.; Paper presented at the Annual Meetings of the American Education Research Association (New York, NY, March 19, 1982) and the National Council for the Social Studies (Detroit, MI, November, 1981).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Administrator Attitudes; Affective Objectives; *Citizenship Education; Cognitive Objectives; Elementary Secondary Education; Political Attitudes; Program Administration; *Program Effectiveness; Program Evaluation; Program Implementation; Social Studies; Teacher Education
IDENTIFIERS *Improving Citizenship Education Project

ABSTRACT

Evaluation results from the Improving Citizenship Education (ICE) Project are presented. The purpose of the ICE project was to design and test a model for improving the political/citizenship knowledge and attitudes of K-12 students by infusing citizenship education into an existing social studies curriculum. This evaluation examined the effectiveness of each of the five teacher support components of the change model (staff development, administrative support, cognitive and affective objectives, community resources, and materials) from the teacher's point of view (process evaluation). Also assessed was the impact of the curriculum change on student performance (product evaluation). Subjects consisted of 17 project and control elementary and secondary teachers, 345 secondary students and 214 elementary students. The measurement instruments for the process evaluation included a multiple-choice instrument and questionnaires. Project and control students were pre- and posttested using the Citizenship Knowledge Test and the Opinionnaire on Political Institutions and Participation. Teacher evaluation results indicated that they made a significant gain in knowledge from staff development, objectives were very helpful, principals were somewhat supportive, and the materials were very helpful. Product evaluation indicated that the ICE Project did significantly and practically affect the political knowledge and attitudes of secondary students; however, elementary students were not significantly affected. A re-evaluation of this project indicated that the ICE project model does significantly affect the knowledge and attitudes of elementary students. Discussion is included on conclusions, limitations, and implications. (NE)

EVALUATION OF A LOCALLY DEVELOPED SOCIAL STUDIES CURRICULUM

PROJECT: IMPROVING CITIZENSHIP EDUCATION[1]

John D. Napier and Mary A. Hepburn
University of Georgia

A paper presented to the Special Interest Group/Research in Social Studies Education at the Annual Meeting of the American Education Research Association (New York, NY, 1982) and at the Annual Meeting of the National Council for the Social Studies (Detroit, MI, November, 1981).

EVALUATION OF A LOCALLY DEVELOPED SOCIAL STUDIES CURRICULUM PROJECT: IMPROVING CITIZENSHIP EDUCATION

John D. Napier and Mary A. Hepburn
University of Georgia

It is an understatement to describe the literature on social studies project evaluation as sparse. Social studies journals and books have provided little information on systematic evaluations of social studies curriculum projects. Two systematic reviews in the recent literature (Wiley, 1977; Shaver, Davis, and Helburn, 1979) painted broad murals of the status of the social studies, and although informative, they contributed only summaries of or brief commentary on evaluation research on innovative social studies curriculum projects. Moreover, the reviews were based on evaluations which examined "new social studies" projects of the sixties and seventies. Not only were the projects dated, but evaluations often were based on impressionistic data or weakly designed studies.

The Elementary and Secondary Education Act (ESEA) of 1965 ushered in a new era of educational improvement projects in the social studies and other subject areas. The focus of the development effort was shifted from research and development centers of universities and educational corporations, where materials of the "new social studies" had been designed, to local school sites where curriculum improvement plans were to be designed and tested. Thus, the focus of evaluation shifted from research into the content and style of newly developed materials to outcomes in the classroom and processes in the school.

The social studies literature lacks evaluation reports or analyses of ESEA improvement projects. Some idea of the impact of the ESEA on the social studies can be gained from more general educational literature (Rand Corporation, 1978; Berry, 1979). But here too the synthesis is not comprehensive. The Rand study was based on a very small selection of local school projects. Also, it reported projects established under federal ESEA guidelines which preceded important amendments passed in 1977 and 1978. These amendments modified extensively the procedures for setting project objectives and designing projects to meet local needs. However, both the general literature and the social studies literature lack in depth studies

---

A paper presented to the Special interest Group/Research in Social Studies Education at the Annual Meeting of the American Education Research Association, New York, 1982.

of the social studies curriculum improvement projects developed under the most recent ESEA local development guidelines.

The purpose of this paper is to report the evaluation results from the Improving Citizenship Education (ICE) Project, which was initiated in 1977 under existing ESEA guidelines and modified to meet revised 1978 state guidelines for validation review. Besides documenting the effects of the ICE project on teachers and students, the paper examines the potentials of local social studies curriculum projects which employ the same type of curriculum change model.

Improving Citizenship Education Model

The ICE project (Fulton County, Georgia) tested a change model which had five teacher support components. The five areas of support were: 1) staff development - a teacher education program dealing with both the political science content and teaching strategies and materials; 2) administrator support - a series of meetings with principals of the project schools informing them and involving them in the project goals and methods; 3) specific learning objectives - cognitive and affective objectives defined to guide the planning and teaching of the project teachers; 4) community resources - persons, agencies, organizations, and sites in the local area were fully identified by topic area to increase the potential for first-hand learning experiences by teachers and students (also a group of community leaders served as an advisory board to the project staff); 5) materials - needed instructional materials were identified and provided for teachers (most needs were met with existing educational media, but some had to be developed by the staff).

The expressed purpose of the ICE Project was to design and test a model for improving the political/citizenship knowledge and attitudes of elementary and secondary students by infusing citizenship education into the existing social studies curriculum. Nine educational goals (Table 1) guided the project's initial effort. From these nine goals, eight cognitive content areas were defined by a teacher committee and project staff (Table 2). From the goals and content areas, the staff and evaluation team derived a set of cognitive and affective objectives to aid evaluation of the Project (Table 3). Further, the relationship of specific objectives to goals was defined (Table 4).

Evaluation Questions

Originally, the evaluation plan for the project called for a ten-weeks implementation; however, the demanf for more immediate information by the state department required a change to a six-weeks implementation in the winter school quarter following the development of the teacher support components. The evaluation had two purposes:

Table 1

Goals of the Improving Citizenship Education Project

1. Understanding and acceptance of the responsibilities of U.S. citizenship;
2. Understanding of the structure and operations of local, state, and national governments;
3. Understanding of the roles of individuals in the political decision-making processes and developing the skills for participating.
4. Understanding of the principles of individual rights and learning to make choices in the context of concern for the society as a whole;
5. Understanding of the U.S. legal system and a commitment to rule of law;
6. Knowledge of current public issues and skills for evaluating alternative choices in regard to these issues;
7. Understanding of the interdependence of the global community and the political processes in operation internationally;
8. Understanding of the means to participate in school, local, state, and national political processes;
9. Understanding of the need for government and community services and the ways to secure, utilize and contribute to these services.

Source: Margolis (1981)

Table 2

Content Areas of the Improving Citizenship Education Project

1. **National Government**-federalism, bicameral lawmaking. executive power and roles, judicial powers, separation of powers, checks and balances, policy-making processes, interest groups, lobbies, electoral college, two-party system, bureaucracy, national revenues and expenditures, U.S. citizenship.
2. **State and Local Government**-intergovernmental relationships, state and local lawmaking, state and local citizenship, services of state and local governments, financing services, executive leadership, bureaucracy, local interest groups.
3. **Democratic Principles**-popular sovereignty, limited powers of government, majority rule, minority rights, reserved powers, general welfare, constitutional republic.
4. **Politics**-political symbols, campaigns, elections, voting behavior, representation, public opinion, means and limits of political power, political socialization, political leadership.
5. **Law and Individual Rights**-rule of law, equality before the law, justice, the Bill of Rights, judicial principles, individual rights and responsibilities, limits of individual freedom.
6. **Global/International Studies**-interdependence, multi-national organizations, United Nations, treaties, balance of power, international interest groups, trade agreements, deterrence, detente, international assistance and foreign aid.
7. **Analytical Skills**-making informed decisions on public issues, value analysis, evaluating information and information sources, interpreting and evaluating quantitative data, distinguishing fact from opinion, synthesizing data from charts and graphs.
8. **Participation Skills**-conflict resolution skills, news awareness skills, registration and voting skills, skills for jury duty and participation in judicial process, party and interest group participation, career awareness, securing basic governmental services.

Source: Margolis (1981)

Table 3

Specific Objectives for the Improving Citizenship Education Project

1.0 Know and apply specific facts, basic concepts, and processes related to government and politics.

- 1.1 Identify specific facts, processes, and basic concepts of **national government.**
- 1.2 Identify specific facts, processes, and basic concepts of **state and local government.**
- 1.3 Identify specific facts, processes, and basic concepts of **democracy.**
- 1.4 Identify specific facts, processes, and basic concepts of **politics.**
- 1.5 Identify specific facts, processes, and basic concepts of **law and individual rights.**
- 1.6 Identify specific facts, processes, and basic concepts of **global affairs.**
- 1.7 Identify **participation skills** related to government and politics.
- 1.8 Utilize **analytical skills** with government and political data and issues.

2.0 Demonstrate commitment to democratic institutions, principles, and processes.

- 2.1 Express commitment to **democratic institutions.**
  - 2.11 Identifies magnitude of agreement with statements about political institutions as others should view them.
  - 2.12 Identifies magnitude of agreement with statements about political institutions as self views them.
- 2.2 Express commitment to **community democratic processes.**
  - 2.21 Identifies magnitude of agreement with statements about community democratic participation as others should view them.
  - 2.22 Identifies magnitude of agreement with statements about community participation as self views them.
- 2.3 Express commitment to **school democratic processes.**
  - 2.31 Identifies magnitude of agreement with statements about school democratic participation as others should view them.
  - 2.32 Identifies magnitude of agreement with statements about school democratic participation as self views them.

Source: Margolis (1981)

Table 4

Relationship between Goals and Specific Objectives of the Improving Citizenship Education Project

Specific Objectives	Goals 1	2	3	4	5	6	7	8	9
1.1	X								X
1.2	X	X							X
1.3			X						
1.4	X							X	
1.5	X			X	X				
1.6	X						X		
1.7	X					X			
1.8	X		X					X	X
2.11	X				X				X
2.12	X				X				X
2.21	X			X				X	
2.22	X			X				X	
2.31	X							X	
2.32	X							X	

1) to evaluate the effectiveness of each of the five teacher support components of the change model from the teachers' point of view (process evaluation), and 2) to assess the impact of the curriculum change model on student performance (product evaluation).

Specific process evaluation questions were:

1. How much did the teachers learn about teaching procedures in the staff development sessions?
2. How supportive did the teachers feel their administrators were during implementation of the citizenship project?
3. How useful did teachers find the delineated objectives of the program?
4. How helpful did the teachers find the community resources identified for the project?
5. How helpful did the teachers find the curriculum materials provided or developed for the project?

Specific product evaluation questions were:

1. Was there a significant and practical difference between the political knowledge of project and control students?
2. Was there a significant and practical difference between the political attitudes of project and control students?

The linking rationale (Rutman, 1977) generated by project staff and evaluators clarified the connection between processes and products and was stated thus:

> Teachers will take part in a staff development project which will make them aware of the specific objectives, cognitive content areas, and goals of the project. In addition, teachers will be instructed in the use of various activities and teaching strategies to achieve the specific objectives, and informed of a variety of useful materials (i.e., community resources and located or developed materials). Teachers will be given administrative support by principals familiar with the goals and objectives of the project. With this background and support, teachers will spend more time teaching the content topics, employ a greater number of activities and strategies, and use a greater number of materials in their classrooms to achieve the specific objectives. Therefore, students will be exposed to an improved citizenship education program, and as a result students will increase their achievement of the specific objectives of the project.

It should be noted that two related elements were evaluated. One, the curriculum change model, a process, was evaluated by operationalizing citizenship education as defined by the specific objectives of the Project. Two, the effectiveness of "citizenship education" with students as defined by the same objectives was simultaneously examined. The data collected allowed for the evaluation of both elements.

Additionally, it should be clarified that because of differing requirements for the teacher education portion of the process and the differing instrumentation needed to measure product, elementary and secondary programs were evaluated separately using the same evaluation design.

Procedures for Evaluation

Subjects

Teachers. Originally, there were 12 secondary and 7 elementary teachers designated for the formal evaluation of the ICE project. One teacher failed to complete the teacher education program and another left the school system. Therefore, the final sample of project teachers consisted of 10 secondary (3 eighth; 6 ninth; and 1 ten-twelfth grade) and 7 elementary (1 first; 1 second; 1 fourth; 2 fifth; and 2 sixth grade) teachers.

These teacners were from five secondary schools and five elementary schools designated as "project schools." All had volunteered to participate in the improvement project. Based on information supplied by the school records office, the project schools were representative of the diverse socio-economic characteristics in the school district and included urban, suburban, and small town communities and school populations ranging from nearly all white to nearly all black students. Project schools were considered representative of the system's approximately 35,000 students attending 60 elementary schools and 19 high schools.

Volunteer teachers from schools matched with project schools served as control teachers. The ten secondary and seven elementary control teachers were further matched by grade level, and in the secondary schools by subject taught.

Students. The students involved in the product evaluation were in classes taught by teacher subjects. For the twenty secondary teachers, one class in their schedules was randomly selected. Initially, there were 419 secondary (201 project; 218 control) students and 262 elementary (134 project; 128 control) students who completed the pretests. However, the final sample size used in the evaluation involved 345 secondary (165 project; 180 control) students and 214 elementary

(110 project; 104 control) students. The loss of subjects from pretest to posttest represented approximately 18% of all groups. The size of the loss reflected the fact that only students who took both pretests and posttests were included in the final sample, and flu accounted for a high percent of the absences. It was administratively impossible to retrieve the missing data.

In the final sample, a total of 157 males (63 project; and 94 control) and 188 females (102 project; 86 control) were in the secondary student sample; and a total of 111 males (55 project; 56 control) and 103 females (55 project; 48 control) were in the elementary student sample. The racial composition of the sample was 208 white (106 project; 102 control) and 137 black (58 project; 79 control) students in the secondary sample; and 117 white (54 project; 63 control) and 97 black (56 project; 41 control) in the elementary student sample.

Measurement Instruments

Process instruments. To measure the effect of the staff development component of the project process, a 48 item multiple choice instrument was constructed to measure specific objectives of the staff development. Eight areas were covered on the measure: Case Studies, Community Studies, Values Education, Moral Education, News media analysis, Simulation, Quantitative Analysis, and Decision-Making. The instrument was scored right/wrong. Scores could range from zero to 48 (with 6 items per area). Instrument validity was based on the direct relationship of the content of the items to the objectives of the staff development. Reliability of the instrument was estimated using Kuder-Richardson Formula 21. Reliability for pretest scores was .66, and .83 for posttest scores.

Additional data for the process evaluation was obtained from three questionnaires--Questionnaire on Management Support, Questionnaire on ICE Objectives, and Questionnaire on Materials including community resources (copies are found in Margolis, 1981). Each questionnaire had two parts. The first part asked teachers to rate the supportiveness or helpfulness of the specific process component on a scale of 1(low) to 4 (high). The second part asked teachers to review in more detail the effectiveness of each support area. This second part was designed for obtaining descriptive information for the project staff. No formal attempt was made to determine validity or reliability of these questionnaires. Informally, the instruments were verified by conversations with project teachers.

Product instruments. The Citizenship Knowledge Test (CKT) was used to collect data on the first product evaluation question. This test has both secondary and elementary editions. The secondary level test has two forms, Form A, and Form B;

and the elementary level test has two forms, Grades 1-2, and Grades 3-7.

The secondary level CKT has 73 items in eight subtests related to the specific objectives listed in Table 3. These are: National Government (11 items), State and Local Government (10 items), Politics and Political Ethics (11 items), Democratic Principles (7 items), Law and Individual Rights (10 items), Global Affairs (5 items), Participation Skills (10 items), and Analytic Skills (9 items). Variation in number of items per subtest (objective) reflected the emphasis each subtest area was given on the secondary level of the ICE project. The total test score for the secondary level CKT could range from zero to 73.

In the Project evaluation, Form B of the secondary level CKT was used (for details on the rationale for choosing Form B, see Hepburn and Strickland, 1979). The validity and reliability of Form B have been documented elsewhere (Hepburn and Strickland, 1979; Napier, Hepburn and Strickland, 1981). Cronbach alpha procedures were used to calculate reliability estimates of the total secondary student CKT scores which were .89 for the pretest scores and .90 for the posttest scores.

On both forms of the elementary level CKT there are 41 items in seven subtests related to the specific objectives 1, 2, 3, 5, 6, 7, and 8 listed in Table 3. (The more abstract content of objective 4 was not emphasized in the elementary grades). The seven subtests are: National Government (6 items), State and local Government (8 items), Politics and Political Ethics (7 items), Law and Individual Rights (7 items), Global Affairs (5 items), Participation Skills (4 items), and Analytic Skills (4 items). Again, the variation in the number of items for each subtest reflected the emphasis each subtest area was given on the elementary level of the ICE project. The major difference between the Grade 1-2 and Grades 3-7 forms of the elementary CKT is in the complexity and difficulty of the wording used in the item stem or distractors. Basically the two forms are parallel. Total test scores could range from zero to 41.

Both forms of the elementary level CKT were used in the evaluation. The validity and reliability of the elementary level CKT is documented elsewhere (Napier, Hepburn, and Strickland, 1981). In the project evaluation, Cronbach alpha reliability estimates for pretest scores were .83, and for posttest scores .88.

The Opinionnaire on Political Institutions and Participation (OPIP) was used to collect data on the second product evaluation question. This test also has both secondary and elementary editions which are parallel. The OPIP is a 48-item instrument with six subtests each containing eight items.

The first and second subtests measure attitudes about political institutions (objectives 2.11 and 2.12). The first subtest, Political Institutions-Others, indicates the respondent's perception of how people in general should view American political institutions. The second subtest, Political Institutions-Self, reflects the respondent's personal attitude toward American political institutions.

The third and fourth subtests assess attitudes related to participation in public political processes (objectives 2.21 and 2.22). The third subtest, Public Political Participation-Others, reflects the respondent's perception of whether individuals should participate in public political processes. The fourth subtest, Public Political Participation-Self, measures the respondent's attitude toward personal political participation.

The fifth and sixth subtests measure attitudes towards participation in school political processes (objectives 2.31 and 2.32). The fifth subtest, School Political Participation-Others, reflects the respondent's perception of whether others should participate in school political processes. The sixth subtest, School Political Participation-Self, reflects the respondent's attitude toward personal participation in school political processes.

The secondary form is designed so that individuals respond to the 48 statements on a 5-point scale (Strongly Agree, Agree, Uncertain, Disagree, or Strongly Disagree). On the elementary form, individuals respond to the statements on a 3-point scale (Agree, Uncertain, or Disagree). In this evaluation, the secondary form was scored on a 3-point scale to allow for comparison across elementary and secondary grades. Also, the instrument was validated using that scale (Hepburn and Napier, 1980a; Napier and Hepburn, 1981). Consequently, the overall score on both the secondary and elementary forms could range from 48 (very negative) to 144 (very positive).

The validity and reliability of the secondary and elementary level OPIP are documented elsewhere (Hepburn and Napier, 1980a; Napier and Hepburn, 1981). Cronbach alpha reliability estimates for secondary subjects were .78 pretest and .89 posttest; for elementary subjects .84 pretest and .88 posttest.

Treatment

Teachers. Staff development for project teachers presented objectives of the ICE project and immersed them in teaching activities, strategies, and materials for achieving the objectives. The group participated in a 45-hour staff development program initiated in July, 1979, with a week of all-day sessions and followed up in the fall with five release-time days. An outline of the lessons and objectives for staff development training is presented in the project's Implementation Handbook (Margolis, 1981).

Administrators. An all-day meeting was held with principals and project schools in 1978 to inform them of objectives and methods of the project and enroll their support. In 1979 the project director met with principals individually to inform them of the progress and again encourage their support.

Students. Utilizing project objectives, materials, including draft versions of the secondary or elementary handbooks[2]; and with administrative support the teachers conducted six-weeks units aimed at improving student political knowledge and attitudes. Each project teacher was free to operationalize the units taught to fit the course and grade levels they were teaching. Therefore, some check was needed on whether they taught the content of the objectives, used the recommended activiites and strategies, and employed the suggested materials. To check on implementation, a daily log was kept by each teacher (a copy is in Margolis, 1981). For comparison, control teachers also kept the same log. Analysis of the log data later indicated that secondary project teachers spent significanly more time teaching the content related to the objectives and taught more citizenship topics. Also they used signficantly more recommended materials, and activities than the control teachers. However, the elementary project teachers did not spend significantly more time teaching the content, nor did they use significantly more topics, materials, or activities then the control teachers (see Table 5). Therefore, the effectiveness of the elementary implementation was questionned (see discussion in results section).

Data Collection

Process evaluation data. Project teachers took the 48-item multiple choice instrument as a pretest and again at the end of the 45 hour summer course. Project teachers responded to the questionnaires on objectives, management support, and materials at the end of the six-week implementation in Winter 1980. Unfortunately, two of the seven elementary teachers failed to return the three questionnaires despite repeated inquiries from the project staff for the return of these measures.

Product evaluation data. Project and control elementary and secondary students took the two pretests in early January 1980. The attitude test was given one day, and the knowledge test was given the next day. Both project and control subjects took the posttests in March with the tests given in the same order.

Evaluation Results

Process Evaluation

Secondary Teachers. Table 6 presents the descriptive data for secondary teachers for staff development pretests and posttests and responses to the first part

Table 5

Descriptive Statistics, F-Ratios, P-Values for Number of Materials, Activities, and Contents plus Time Covering Topics used by Secondary and Elementary Evaluation Project and Control Teachers

Source	Statistic	Secondary Project (10)	Secondary Control (10)	F	P	Elementary Project (7)	Elementary Control* (6)	F	P
Number of Materials (Scale 1-8)	Mean	7.6	5.0	17.69	.001	5.3	3.5	1.25	.288
	S.D	0.8	1.8			2.6	3.2		
Number of Activities (Scale 1-8)	Mean	7.0	4.6	12.23	.003	5.6	4.2	2.30	.158
	S.D.	1.1	1.9			1.3	2.0		
Number of Contents (Scale 1-8)	Mean	7.2	4.5	8.19	.010	6.4	4.3	3.32	.096
	S.D.	0.8	2.9			2.1	2.0		
Time on Topics	Mean	1332.0	726.0	10.95	.004	714.3	280.0	2.95	.114
	S.D.	234.1	529.8			563.3	271.5		

*One Control teacher never returned the daily log

Table 6

Descriptive Statistics for Staff Development Pretest and Posttest plus Responses to Questionnaire on ICE Objectives, Management Support and Materials for Evaluation Project Teachers

Source	Statistics	Secondary	Elementary
Pretest (Scale 0-48)	Mean	28.4	23.0
	S.D	5.8	8.7
	N	10	7
Posttest (Scale 0-48)	Mean	34.3	28.9
	S.D.	3.7	7.2
	N	10	7
Objectives (Scale 1-4)	Mean	3.4	3.6
	S.D.	0.7	0.6
	N	10	5*
Management Support (Scale 1-4)	Mean	2.6	2.8
	S.D.	1.1	1.1
	N	10	5*
Materials (Scale 1-4)	Mean	3.6	3.6
	S.D.	0.5	0.6
	N	10	5*

*Two Elementary Teachers never returned questionnaires

of the three questionnaires. A significant gain in knowledge was made from the staff development (t_d = 3.34; df = 9; $p < .01$). Questionnaire results indicated that the objectives were very helpful, principals were somewhat supportive, and the materials were very helpful.

Elementary teachers. Table 6 also presents the same descriptive data for elementary teachers. Like the secondary teachers, elementary teachers made a significant gain in knowledge from the staff development (t_d = 5.03; df = 6; $p < .01$). Results from the questionnaires indicated that the elementary teachers found the objectives very helpful, the principals were somewhat supportive, and the materials were very helpful.

Product Evaluation

For both the secondary and elementary data analysis of covariance procedures were used. The analysis was conducted blocking for effects by gender and race. According to Cook and Campbell (1979), analysis of covariance blocking on variables related to the dependent variable is a more powerful procedure than simple analysis of variance. The variables of gender and race were used as blocking variables because past studies have indicated there are relationships between these two variables and political knowledge and attitudes (see Hepburn and Napier, 1980b).

Secondary product results. Table 7 presents descriptive statistics for the secondary level students on the knowledge and attitude tests. Analysis of covariance on project and control results is presented in Table 8. A significant difference was found between project and control groups on both the CKT and OPIP.

The product evaluation questions asked about "practical" as well as statistical significance. Practical signfficance was determined by calculating "effect size" of statistically significant result. The effect size is a statistical "rule of thumb" suggested by Tallmadge (1977, p. 34). Effect size is calculated by subtracting the posttest mean scores of experimental and control groups and dividing the remainder by the standard deviation of the control group (see Glass, 1978). An effect size can range from 0% to 100% (or 0 to 1 if converted to fractions). For the CKT, an effect size of 25-50% was the criterion used to determine practical differences.[3] For the OPIP, an effect size of 10-25% was used. A lower effect size was expected from the OPIP because attitudes are difficult to change in any short-term treatment (Shaw and Wright, 1967; Hepburn and Napier, 1980b).

Table 7

Descriptive Statistics for Main Effects of Group, Gender, and Race in Secondary and Elementary Evaluation Samples on Knowledge and Attitude Pretests and Posttests

Source	Main Effect	Class	N	Knowledge (CKT)					Attitude (OPIP)				
				Pretest		Posttest			Pretest		Posttest		
						Unadjusted		Adjusted			Unadjusted		Adjusted
				Mean	S.D.	Mean	S.D.	Mean	Mean	S.D.	Mean	S.D.	Mean
Secondary	Group	Project	165	31.9	12.5	36.8	12.5	35.8	122.0	9.3	123.2	9.9	122.0
		Control	180	29.7	10.3	30.3	10.9	11.2	119.3	9.9	119.2	10.5	120.4
	Gender	Male	157	30.3	11.4	32.7	12.5	33.4	118.9	10.1	118.5	11.2	120.0
		Female	188	31.2	11.5	34.0	11.7	33.4	122.1	9.1	123.3	9.2	122.1
	Race	White	208	33.7	11.5	36.6	11.9	34.1	122.3	9.2	122.7	9.9	121.5
		Black	137	26.3	9.8	28.5	10.8	32.3	118.1	9.9	118.7	10.8	120.6
Elementary	Group	Project	110	24.2	7.0	26.5	7.7	27.4	110.8	12.9	113.2	14.4	114.1
		Control	104	26.4	6.5	26.1	7.5	26.0	112.7	12.0	113.7	13.8	112.6
	Gender	Male	111	25.1	7.2	26.5	7.7	26.7	112.0	11.9	112.8	14.6	112.6
		Female	103	25.4	6.4	26.9	7.5	26.8	111.4	13.1	114.1	13.6	114.5
	Race	White	117	27.6	5.9	28.2	6.9	26.3	114.0	12.2	116.6	12.4	114.5
		Black	97	22.5	6.9	24.9	8.0	27.2	109.0	12.3	109.6	15.0	112.0

Table 8

Analysis of Covariance Tests on Knowledge and Attitude Posttest for Secondary and Elementary Evaluation Subjects

Source	Secondary					Elementary				
		Knowledge (CKT)		Attitude (OPIP)			Knowledge (CKT)		Attitude (OPIP)	
	df	MS	F	MS	F	df	MS	F	MS	F
Covariate	1	33233.3	756.8*	20663.2	439.4*	1	6529.1	248.2*	28517.4	448.9*
Group	1	1776.2	40.5*	205.7	4.4*	1	93.9	3.6	114.4	1.8
Gender	1	0.1	0.0	376.8	8.0*	1	0.6	0.0	131.7	2.1
Race	1	219.78	5.0*	55.2	1.2	1	31.1	1.2	319.3	5.0*
Group x Gender	1	24.46	0.6	10.6	0.2	1	15.4	0.6	0.7	0.0
Group x Race	1	30.6	0.7	25.4	0.5	1	159.2	6.1*	0.3	0.0
Gender x Race	1	151.0	3.4	0.0	0.0	1	12.2	0.5	36.0	0.6
Group x Gender x Race	1	127.6	2.9	42.3	0.9	1	26.3	0.7	3.1	0.1
Residual	336	43.9		47.0		205	26.3		63.5	
Total	344	146.5		108.3		213	57.6		198.0	

*Significant at the $p < .05$ level

The effect size found for the CKT was 42%, and for the OPIP 11%. Thus, the answers to product evaluation questions were positive for the secondary evaluation. The ICE Project did significantly and practically affect the political knowledge and attitudes of secondary students.

Elementary product results. Table 7 also presents the descriptive statistics for the elementary students on the CKT and OPIP tests. Analysis of covariance results on the CKT and OPIP posttest comparisons between project and control students (Table 8) indicated no significant difference between project and control groups on either test. Therefore, answers to the two product evaluation questions were negative for the elementary project evaluation. The political knowledge and attitudes of elementary students were not significantly affected.

There were circumstances related to the implementation of the ICE Project which could have accounted for the negative findings in the elementary evaluation. Originally, a ten week implementation had been designed for both secondary and elementary programs. For elementary classes the average daily time devoted to content was 23.8 minutes, while the secondary average was 44.4 minutes. Obviously four additional weeks of treatment would be more essential for elementary students. However, report deadlines set by the funding agency had forced the shorter six-week treatment. Cook, Cook, and Mark (1977) warn that failure to give sufficient exposure to process results in loss of statistical power. Also, because two project teachers did not return the three questionnaires documenting process, the evaluators wondered whether they had actually implemented the project properly. Staff and reviewers decided that re-evaluation in elementary grades was necessary before final conclusions about the effectiveness of the project at this level.

Supplemental information on gender and race. Although product evaluation questions did not call for examination of gender and race, data on these two variables were collected and used in the statistical analysis to send some light on the effects on student sub groups. There was a significant difference between races on the knowledge test (CKT) and between genders on the attitude test (OPIP). However, no significant interactions were found which means that gender did not make a significant difference in performance by project or control and race did not make a difference by project and control. Nevertheless, within project and control groups whites scored better than blacks on knowledge, and females scored better than males on attitudes (see Table 7)

Procedures for Re-evaluation

Subjects

Teachers. In the re-evaluation study 20 elementary teachers were originally involved. (These were different teachers than in the first evaluation.) These 20 project teachers were matched with 20 control teachers in the same manner used in the initial evaluation. Unfortunately, one control teacher left the school system and the matched project teacher was dropped from the study. Also, two project teachers never returned daily logs and they with their control matches were dropped. The final sample of teachers consisted of 17 project and 17 control teachers all matched by grade (1 first; 2 second; 4 third; 2 fourth; 2 fifth; 3 sixth; and 3 seventh).

Students. The students involved in the re-evaluation were in classes taught by project and control teachers. Initially, 611 students (290 project; 321 control) took the pretests. However, the final sample size was 546 (260 project; 286 control) -- approximately a 10% loss for each group. The loss reflected the fact that only students who took both pretests and posttests were included in the final sample. The final sample had 272 males (132 project; 140 control) and 274 females (128 project; 146 control). Racial composition was 264 white (129 project; 135 control) and 282 black (131 project; 151 control).

Measurement Instruments

Process instruments. Only the three questionnaires-Questionnaire on ICE objectives, Questionnaire on Management Support, and Questionnaire on Materials were used in the re-evaluation. These instruments were identical to those described earlier in the Procedures for Evaluation section.

Product instruments. Again the elementary forms of the CKT and OPIP were used to collect data from the student subjects. Cronbach alpha reliability estimates of the total CKT in the re-evaluation was .85 for the pretest, and .84 for the post-test. Cronbach alpha reliability estimates for the total OPIP score were .83 for the pretest, and .83 for the posttest.

Treatment

Teachers. The project teachers, none of whom were previously involved with the project, received staff development training similar to that in the initial evaluation. The major difference was in the length of the treatment. In the re-evaluation, five days of staff development (approximately 40 hours) was used, but the major elements of the original staff development treatment were maintained, and the staff

felt the treatment was equivalent. Since the effectiveness of the staff development process was documented in this evaluation and in an earlier study (Hepburn and Napier, 1979), it was not monitored again.

Administrators. Again meetings were held with principals in project schools to make them aware of objectives and methods of the project. They were encouraged to be supportive to teachers involved in the project.

Students. Project teachers followed the objectives, and utilized the materials, including a finalized version of the elementary handbook, and with administrative support taught their citizenship units over a ten-week period. Again, each project teacher was free to operationalize the units taught. Therefore, daily logs were again kept by both project and control teachers to document daily treatment to students. Analysis of the log data indicated that the project teachers spent significantly more time covering contents related to the objectives; and taught using significantly more of the topics, materials, and activities than control teachers (see Table 9).

Data Collection

Process evaluation data. All 17 of the elementary project teachers responded to the three questionnaires on project objectives, management support, and materials at the end of the ten week implementation in the Fall, 1980.

Product evaluation data. Project and control elementary students took the two pretests at the beginning of Fall quarter 1980. The attitude test (OPIP) was given one day, and the knowledge test (CKT) was given the next day. Posttests were given in November at the end of the ten weeks of treatment, and the tests were given in the same order. Since it was impossible to arrange to retrieve missing data, only those students who took all four tests were included in the evaluation.

Re-evaluation Results

Process Evaluation

Table 10 presents descriptive data for responses to the first part of each questionnaire. Results again indicated that objectives and materials were perceived as very helpful. Unlike the initial evaluation, the project teachers perceived management support in the re-evaluation as very supportive.

Product Evaluation

Again, analysis of covariance procedures were used blocking on gender and race. Descriptive statistics for student tests are in Table 11. The analysis of covariance

Table 9

Descriptive Statistics, F-Ratios, P-Values for Number of Materials, Activities, and Contents plus Time Covering Topics used by Elementary Re-evaluation Project and Control Teachers

Source	Statistic	Project (1)	Control (16*)	F	P
Number of Materials (Scale 1-8)	Mean	6.0	3.9	17.55	.000
	S.D.	1.4	1.5		
Number of Activities (Scale 1-8)	Mean	5.8	3.4	37.41	.000
	S.D.	0.8	1.3		
Number of Contents (Scale 1-8)	Mean	7.4	3.8	29.57	.000
	S.D.	1.4	2.3		
Time on Topics	Mean	1114.4	285.0	22.49	.000
	S.D.	632.5	307.2		

*One Control Teacher never returned the daily log

Table 10

Descriptive Statistics for Responses to Questionnaire on ICE Objectives, Management Support and Materials for Re-evaluation Project Teachers

Source	Statistic	Project
Objectives (Scale 1-4)	Mean	3.5
	S.D.	0.5
	N	17
Management Support (Scale 1-4)	Mean	3.8
	S.D.	0.4
	N	17
Materials (Scale 1-4)	Mean	3.8
	S.D.	0.6
	N	17

Table 11

Descriptive Statistics for Main Effects of Group, Gender, and Race in Elementary Re-evaluation Sample on Knowledge and Attitude Pretests and Posttests

Source	Class	N	Knowledge (CKT)					Attitude (OPIP)				
			Pretest		Posttest			Pretest		Posttest		
					Unadjusted		Adjusted			Unadjusted		Adjusted
			Mean	S.D.	Mean	S.D.	Mean	Mean	S.D.	Mean	S.D.	Mean
Group	Project	260	23.8	6.8	26.7	6.3	26.4	108.6	11.7	113.2	12.2	113.6
	Control	286	23.2	7.1	24.3	7.3	24.5	109.8	12.1	112.2	12.8	111.7
Gender	Male	272	23.5	7.2	25.3	7.1	25.3	109.0	12.3	112.0	12.9	112.0
	Female	274	23.5	6.8	25.6	6.7	25.6	109.4	11.6	113.3	12.0	113.2
Race	White	264	25.3	6.6	27.0	6.7	25.7	110.7	11.0	114.7	12.5	113.6
	Black	282	21.8	7.0	24.0	6.8	25.2	107.8	11.7	110.7	12.2	111.7

(Table 12) indicated a significant difference between project and control groups on both the attitude and knowledge test. The effect size for the CKT was 26%, and for the OPIP 12% indicating positive answers to both product evaluation questions. Based on re-evaluation results, the ICE Project model does significantly and practically affect the political knowledge and attitudes of elementary students.

Supplemental Information on Gender and Race

Examination of Table 12 indicated that there were no significant differences between gender on either the knowledge (CKT) or attitude (OPIP) tests. Also, there was no significant difference between races on knowledge; but there was a significant difference between races on attitude with whites scoring higher (Table 11). However, again there were no significant interactions which means race did not make a significant difference in performance by project and control groups. Within project and control groups, whites did score better than blacks on attitude.

Conclusions

The Improving Citizenship Education Project was evaluated in the real world of the school system where many factors human and mechanistic prevent the implementation of an ideal evaluation design. Holiday vacations, student absences, change in class schedules, teacher job changes, and the failure to return a few questionnaires -- these are some of the many problems that quite naturally compromise any attempt to conduct an ideal evaluation. In another paper, the authors have discussed criteria of useful project evaluation design, which though modeled on evaluation ideals, must realistically be modified to meet limitations of the field setting in which it is applied (Napier and Hepburn, 1981). In this section we will focus on specific limitations of the actual design used and review the results for their implications regarding social studies curriculum change.

Limitations

It must first be emphasized that results of the process evaluation in this study apply only to "process" as operationalized by four process instruments (staff development test and three questionnaires) and the daily teacher log. Likewise, results of the product evaluation apply only to products which are operationalized by the specific objectives of the project and the knowledge and attitude instruments which assessed the student outcomes.

Second, the internal validity of the study was not absolutely assured, although teachers and students were matched and quasi-experimental design was used in the product evaluation. One source of rival causes was not controlled with the evaluation design used--the interaction of selection and maturation or history. Also,

Table 12

Analysis of Covariance Tests on Knowledge and Attitude Posttest for Elementary Re-evaluation Subjects

Source	df	Knowledge (CKT		Attitude (OPIP)	
Covariance	1	14947.9	778.5*	46392.6	667.0*
Group	1	482.3	25.1*	503.1	7.2*
Gender	1	9.6	0.5	108.0	2.7
Race	1	26.3	1.4	439.5	6.3*
Group x Gender	1	29.6	1.5	11.1	0.2
Group x Race	1	7.9	0.4	42.1	0.6
Gender x Race	1	5.6	0.3	2.3	0.0
Group x Gender x Race	1	70.8	3.7	54.1	0.8
Residual	537	19.2		69.6	
Total	545	47.5		155.9	

*Significant at the $p \leq .05$ level

since pre-experimental design was used to examine the process, the effectiveness of the process could be the result of such rival causes as past history, selection, Hawthorne effect, and so forth. Thus, results of the study could be attributed to extraneous variables.

A third limitation is in the statistical validity of the evaluation and re-evaluation. Analysis of covariance with blocking was the statistical procedure used to analyze product data. Possibly different analysis procedures would produce different results.

Finally, the external validity of the study may be questioned since volunteer teachers were used and only students who completed all pretest and posttests were analyzed. It seems likely that non-volunteer teachers and non-attending students would produce different results. Also, only one school system was used in the evaluation. Again, it seems likely that teachers and students in different school systems could produce different results.

On review of the above-meantioned limitations, the evaluators regard the issue of generalizability of the instrumentation as a minor one. The instruments used in the evaluation of student knowledge and attitude outcomes were general, norm-referenced measure which had extensive piloting and analysis in validity studies. Moreover, most school systems find it feasible to evaluate educational programs with such paper and pencil measures.

Regarding internal validity, in defense of the design, consideration should be given to the fact that teachers and student groups were matched carefully. Also, the use of analysis of covariance procedures with blocking lessened the likelihood of initial differences as an internal validity problem.

The use of analysis of covariance with blocking was not viewed as a major limitation either. Simple inspection of the descriptive statistics indicated that differences between groups, genders, or races were logical.

Only external validity was considered to be a serious limitation. The citizenship education model may not work with other teachers, students, or school systems. Consequently, replication with other teachers and students in other school settings is the appropriate way to verify the conclusions of this evaluation.

Implications

Results of the test of the ICE model suggest implications for social studies curriculum change which are significant. The model centered on teacher involvement and teacher behavior as a key factor in generating improved student political knowledge and attitudes. If the results of this evaluation are valid, then it is

logical that any attempt to modify social studies curriculum and instruction must involve teachers. Likewise, the involvement of principals, leaders with the local school system, and community leaders appears important to the effectiveness of improvement projects. Without the combined support of these actors, change in the curriculum might not occur.

The findings indicate that clear explication of objectives and an explanation of how the objectives are linked to content and teaching methods contribute to the success of a curriculum change project. Likewise, teacher education which provides content information and instructional ideas and materials related to the objectives also plays a role in generating change in the classroom. These components provide teachers with the intellectual framework and the classroom tools to integrate and adapt the modified curriculum at their grade level. Probably a major impetus for change came from the attitudinal effects of the cameraderie of a group of teachers working together in a staff development program where their involvement is highlighted. All support components were aimed in their direction. They developed their own study units for students, and they were considered the important actors in the project. The result was that they changed what they were doing in their classrooms in regard to citizenship education, and the ultimate outcome was that student performance on knowledge and attitude tests, regardless of gender or race, showed significant positive changes.

Again, it should be made clear that the Improving Citizenship Education Project was designed to enhance citizenship education as defined by the specific objectives of this project--which are essentially political knowledge, political skills, and political attitude objectives. Given a different set of objectives this multiple support model centered on the teachers might work as well. However, we can only claim that the model is valid and effective for the ICE project.

Currently, the ICE Project model is being disseminated throughout the State of Georgia with approximately 20 school systems and 67 schools in the process of adopting the change model and utilizing the guiding materials. Students in these adopting schools are being tested, and from the dissemination information, the replicability of this evaluation will be discovered.

Notes

[1]The authors thank Helen W. Richardson, Project Director, and Sheila L. Margolis, Project Coordinator, for invaluable assistance. We also thank Gerald Klein of the Georgia State Department of Education; and the parents, administrators, teachers, and students who participated in the evaluation.

[2]Information about the final versions of the Secondary and Elementary Handbooks can be obtained from:

Sheila L. Margolis
Improving Citizenship Education Project
Liberty-Gunn Center
4820 Long Island Drive
Atlanta, Georgia 30342

[3]While there are no established criteria indicating that a given effect size is of practical significance, the work of Cohen (1977) represents the most useful source. Basically, Cohen defends using the idea of small, medium, and large effects. A small effect (approximately .10) is a change which is unlikely to be observable in outward behavior, but which has theoretical value. A medium effect (approximately .25) is a change which has the possibility of perception in outward behavior and the effect most often found in psychological research. Finally, a large effect (approximately .50) is a change which is seen in outward behavior, but is a rare finding in single psychological research studies. In this study a small effect was expected for attitude changes while a medium effect was expected for knowledge change. Attitudes are difficult to alter, yet any positive change would be of value to a curriculum project. Knowledge is more easily changed, but since the knowledge test was a general norm-referenced instrument, a large effect was unlikely.

References

Berry, M. The teacher and the Rand Study. Today's Education, 1978, 68, 38-40.

Cook, T.D., Cook, F.L., and Mark, M.M. Randomized and quasi-experimental designs in evaluation research; An introduction. In, Rutman, L. (ed.) Evaluation research methods: A basic guide. Beverly Hills, CA: Sage, 1977.

Cook, T.D. and Campbell, D.T. Quasi-experimentation: Design and analysis issues for field settings. Chicago: Rand McNally, 1979.

Glass, G.V. Integrating findings: The meta-analysis of research. Review of research in education, Vol. 5. Itasca, IL: Peacock, 1978.

Hepburn, M.A. and Napier, J.D. Evaluation of the Improving Citizenship Education Project staff development and pilot study. Unpublished manuscript, 1979.

Hepburn, M.A. and Napier, J.D. Development and initial validation of an instrument to measure political attitudes. Educational and Psychological Measurement, 1980, 40 (4), 1131-1139. (a)

Hepburn, M.A. and Napier, J.D. Patterns of development of student attitudes toward political institutions and political participation. Paper presented at the Annual Meeting of the National Council for the Social Studies, New Orleans; 1980. (b)

Hepburn, M.A. and Strickland, J.B. The development of instruments to measure students citizenship knowledge and skills. Journal of Social Studies Research, 1979, 3 (2), 50-57.

Margolis, S.L. (ed.). Improving citizenship education: Implementation handbook. Atlanta, GA: Fulton County Board of Education, 1981.

Napier, J.D. and Hepburn, M.A. Manual for the Opinionnaire on Political Institutions and Participation. Athens, GA: Dept. of Social Science Education, University of Georgia, 1981.

Napier, J.D. and Hepburn, M.A. Social studies project evaluation: A case study. Paper presented at the Annual Meeting of the National Council for the Social Studies, Detroit, 1981.

Napier, J.D., Hepburn, M.A., and Strickland, J.B. Manual for the Citizenship Knowledge Test. Athens, GA: Dept. of Social Science Education, University of Georgia, 1981.

Rand Corporation. Implementing and sustaining innovations (Vol. 8 of the Federal programs supporting education change). U.S. Office of Education, Dept. of Health, Education, and Welfare, 1978.

Rutman, L. Planning and evaluation study. In, Rutman, L. (ed.), Evaluation research methods: A basic guide. Beverly Hills, CA: Sage, 1977.

Shaver, J., Davis, O.L., and Helburn, S. An interpretive report on the status of precollege social studies education based on three NSF funded studies. Social Education, 1979, 43, 150-153.

Tallmadge, G.K. Ideabook. Washington, D.C.: U.S. Department of Health, Education, and Welfare, 1977.

Wiley, K. The status of pre-college science, mathematics, and social science education 1955-75, Volume 3: Social science education. Social Science Education Consortium, 1977 (Report to the National Science Foundation, Contract C7620667).