

An Evaluation of

AMERICANS ALL

A National Multicultural Education Program

Results of an Evaluation of Program Effects
in Washington, D.C.

Mark F. Testa, Ph.D.

Associate Professor

School of Social Service Administration

The University of Chicago

February 1992

The changing demography of American classrooms is prompting a nationwide movement to restructure school syllabi to reflect society's growing ethnic diversity. The likely dimensions of this change are already discernible in the ethnic profiles of the five states with the largest populations under 18 years of age- California, Florida, Illinois, New York, and Texas. Together they account for 35 percent of the nation's child population. Presently, 43 percent of the children of these states do not have European ethnic origins. Rather, they are: Hispanic (24%), African (14%), and Asian (5%). By the year 2000, children of these diverse ethnic origins will become the majority of students in these five states. This pattern could well become characteristic of the entire nation toward the middle of the twenty-first century if current demographic trends continue.

To prepare students to function in a multicultural society, educators have begun to diversify school syllabi to acquaint students with the history of minority and diverse ethnic and racial groups whose experiences have hitherto been neglected or omitted from classroom curricula. Although many people welcome the diversification of curricula as enriching students' cultural identities, self-concepts, and multicultural awareness, others warn that the assumptions and goals of the current reforms risk undermining some essential foundations of American solidarity. Certain commentators caution that by schools' celebrating and strengthening children's sense of ethnic origins and identities, society runs the danger of encouraging social divisiveness at the expense of national unity (Ravitch 1990, Schelsinger 1991).

Progress toward settling such issues has been hampered by a lack of empirical studies of programs of multicultural education. Although many of the arguments in favor of a more diversified curriculum have been around for several decades, some of its key claims, such as improv-

from the Images of Ellis Island curriculum which commemorates the American immigrant experience and the historic importance of Ellis and Angel Islands. It supplements these materials with essays on the particular experiences of African Americans, Asian Americans, European Americans, Mexican Americans, Native Americans, and Puerto Ricans. Teachers are trained in a two-day workshop in the use of these products and in methods of appreciative learning to promote four general goals: (1) enriching children's cultural identities, (2) enlarging their multicultural awareness, (3) enhancing self-esteem, and (4) fostering critical thinking skills.

A core value of Americans All and other multicultural programs is cultural pluralism. The following two statements are indicative of this orientation. The first comes from the Introduction to the Americans All's teachers guide and second from the National Coalition for Cultural Pluralism:

The world in general and American society in particular are made up of many different groups of people with a variety of familial, ethnic and cultural backgrounds. Working cooperatively and effectively within communities is easier when people appreciate and value these multiple cultures and human experiences. Developing multicultural awareness increases one's sense of security or belonging, provides opportunities for expression and communication, and enhances one's sense of personal empowerment and freedom (Christopher and Sreb 1989, p.i-1).

Cultural pluralism is a state of equal co-existence in a mutually supportive relationship within the boundaries or framework of one nation of people of diverse cultures, with significantly different patterns of belief, behavior, color, and in many cases with different languages. To achieve cultural pluralism, there must be unity within diversity. Each person must be aware of and secure in his own identity, and be willing to extend to others the same respect and rights that he expects to enjoy himself (quoted in Suzuki 1979, p. 45).

Implicit in the above statements are a set of propositions that provide the rationale for Americans All and similar multicultural programs: (1) increasing students' awareness and appreciation of their own *cultural identity* and origins, (2) fashioning a basis for increasing their *multi-*

be seen as justifying or undermining ongoing reform efforts but rather as contributing to a deliberative process that informs policy makers, funders, and administrators about past accomplishments and future adjustments that need to be made in order to achieve desired goals.

In the next section, I shall discuss the methods used in this study to evaluate the implementation of the Americans All program in the Washington, D.C. school system. The original design called for the random assignment of classrooms within schools to the Americans All programs or to a one-year waiting list (control group). Two-rounds of identical instruments were administered to students in the program and control groups. All students were asked to complete a brief, self-report measure designed to assess self esteem in children and adolescents. In addition, students in grades eight and eleven were asked to complete a questionnaire designed to measure scholastic interest, educational aspirations, perception of school climate, attitudes toward cultural pluralism, and their estimation of the contributions of different ethnic groups to American history. The primary hypothesis of this study is that children in the Americans All program will exhibit significant differences relative to a control group in self-esteem measures, ratings of minority ethnic groups' contributions to American history, and acceptance of cultural pluralism.

Method

Subjects

The subjects of this study are students attending grades 3, 5, 8, and 11 in Washington, D.C. public schools. In 1991, the principals of 23 school volunteered their institution's participation in the research. A total of 39 program classrooms and 39 control classroom out of an original 102 were deemed eligible for random assignment to the program or to a waiting list for participation in the next academic year. Classrooms on the waiting list constituted the control group.

ally regarded as psychometrically sound for assessing children's self esteem (Chiu 1988). In some quarters, however, it is considered better suited for children than adolescents (Blascovich and Tomaka 1991). Testing forms were purchased from Western Psychological Services.

Self-Esteem (Grades 8 and 11): Under licensing agreement with Western Psychological Services, the Piers-Harris scale was also included in a special questionnaire I developed for Americans All's use with eighth and eleventh graders. Because of the uncertainty over Piers-Harris's suitability for adolescents, the questionnaire also included a variant of the widely used Rosenberg (1965) Self-Esteem Scale. The Rosenberg scale was originally a 10-item inventory designed to measure adolescents' global feelings of self-worth. The particular version reproduced in the Americans All questionnaire is a modified version of the original scale which the U.S Department of Education used in its 1988 National Education Longitudinal Study (NELS:88) of eighth graders. The NELS:88 version was chosen over the original scale to facilitate drawing comparisons between the Washington D.C. sample and a 1988 national sample of eighth graders. Although the Rosenberg scale is typically scored using a Likert-like response format (strongly agree, agree, disagree, strongly disagree), the dichotomized format of the Piers-Harris scale was used in the Americans All questionnaire. This change was made partly for convenience sake, but also was motivated by research that pointed to the possibility of systematic biases in black student responses to Likert-type questionnaire items, who tend to favor the extreme response categories independent of item content. Bachman and O'Malley (1984) reported that black students scored significantly higher than whites in self-esteem scores when a full four- or five-point response range was used, but that the racial difference disappeared when a truncated scoring method was used. I obtained similar test results for the NELS:88 sample using a dichotomized scoring method to control for racial differences in the use of extreme response categories. A dichotomized scoring method will also be used in this study.

Figure 1. Feelings About Peoples Differences

How much do you agree with each of the following statements about people's differences?

(Circle One On Each Line)

Items	Strongly Agree	Agree	Disagree	Strongly Disagree
1. People of different color (whether they are black, white, brown or some other color) are basically the same.	1	2	3	4
2. People who live here but don't speak English aren't true Americans.	1	2	3	4
3. It would be better for everyone if people dated only people of their own race.	1	2	3	4
4.* People's differences in language, religion and customs are the real strengths of this country.	1	2	3	4
5. People who weren't born in the United States haven't contributed much to this country's history.	1	2	3	4
6. It would be better all around if people spoke the same language, practiced the same religion, and shared the same customs.	1	2	3	4

* Scoring reversed to indicate pluralistic appreciation.

measure students' feelings about people's differences. For the analysis, I constructed two additive scales from the items: degree of *pluralistic appreciation* (items 1, 4 and 6) and degree of *pluralistic acceptance* (items 2 and 5). The scores range from 0 to 15 and from 0 to 7, respectively, with higher scores indicating greater appreciation or acceptance. Item 3 on interracial dating was analyzed separately. As an additional indicator of multicultural awareness, I computed a measure of students' assessments of *other ethnic group's historical contributions* by adding together their responses on the contributions of different ethnic group's to the country's history. The scores range from 0 to 19 with higher scores indicating greater contributions.

Statistical Analysis

With a pretest-posttest design with two nonequivalent groups, the analysis of program effects is most straightforwardly approached using a conditional or regression model of change (Plewis 1985). The simplest way of doing this is, first, to find the best-fitting linear relationship between the posttest and pretest scores of the variable for which one is seeking to observe change. Second, one computes the residuals, which is done by subtracting the actual posttest score from the score predicted on the basis of the linear relationship to the pretest. Next, one repeats the same two steps this time substituting an indicator variable that indexes membership in the program or control groups for the posttest variable. Finally, one regresses the posttest residuals against the program residuals. Since the resulting regression coefficient can be thought of as measuring the adjusted relationship between the posttest score and program indicator after removing their respective linear relationships with the pretest score, the size of the coefficient can be interpreted as an indicator of programmatic change. The larger the coefficient, the more change can be attributed to the program. Of course, this simple adjustment procedure assumes that no other important differences remain between the program and control groups which could affect the outcome. Otherwise, one would need to introduce these influences beforehand into the statistical adjustment of the posttest and program variables.

The above steps can be summed up in the following multiple regression equation:

$$Y_2 = \alpha + B_1Y_1 + B_2M + B_3Z + e_2, \quad (1)$$

where Y_2 is the posttest score, Y_1 is the pretest score, M is a indicator variable that indexes membership in the program or control group, and Z is all other important influences of Y_2 which are correlated with M . If the error term e_2 satisfies the necessary statistical assumptions, the ordinary least squares estimate of B_2 will provide the best linear estimate of the program's effect.

Table 1-- Characteristics of Eighth Graders

	NELS:88	Pretest Sample	Final Sample
Sample N	200	203	88
% Female	52.4	58.4	51.1
% African American	65.8	85.6	89.3
% Living with mother only	38.1	43.6	40.1
% College graduate (parents)	19.3	33.8	34.5
% Repeat a grade	28.9	29.3	33.0
Age		13.8	13.8
Self-esteem (Rosenberg)	2.6	2.6	2.7

in the study. The NELS:88 subsample was restricted to urban, public schools in the northeastern United States with minority enrollments of 90 percent or more. This is as close as one can come to matching the NELS:88 sample to the average profile of the Washington, D.C. public school system.

The data show that while the NELS:88 subsample includes a lower concentration of African-American students than the Americans All's sample of eighth graders, they are otherwise quite comparable in several important respects. Almost equivalent proportions of students lived with their mother only (around 40%). A lower proportion of the NELS:88 subsample had college-educated parents (19% vs 34%), but similar proportions of students had ever repeated a

Data on racial and ethnic self-identification are available only for the upper school sample. The data show that classrooms with larger concentrations of African-American students were more likely to complete the study. The slight racial imbalance in the composition of the program and control groups at pretest became even more uneven at posttest. As a consequence, control group data are available only for African-American students.

A convenient way to summarize the compositional differences between the program and control groups is to compute a correlation matrix. The smaller the group differences are, the closer the correlation coefficients will be to zero. Large differences ($r > 0.150$) indicate areas of non-comparability. Table 2 shows that the lower school groups differ significantly by gender and the upper school groups differ significantly by grade and race. The signs of the correlation coefficients indicate that in the 3rd and 5th grades males are under-represented in program group. In the 8th and 11th grades, juniors are under-represented in the program group and blacks over-represented. For the other variables, the correlation coefficients are sufficiently close to zero so that the differences are ignorable.

In the following sections, I report the results of the regression analyses of the effects of Americans All on changes in student's self-esteem, cultural identity, and multicultural awareness. My first pass at the data assumes that the program and control groups are equivalent. No other variables aside from the pretest score and program indicator are included in the regression model. The program's coefficient in this regression model can be interpreted as an estimate of the program's effect under ideal experimental conditions. Since these ideal conditions were not achieved in practice, my second pass at the data introduces the appropriate control variables of gender, race and grade to adjust for pre-existing group differences. Because none of the control

the upper grades.

The results are highly consistent across grades and instruments. There are virtually no statistically significant differences between the program and control groups. Once comparisons are adjusted for pretest scores, students assigned to the Americans All program evinced no greater improvement in self-esteem relative to students in the control group. In fact, the American All's standardized (beta) coefficients are slightly negative for the Piers-Harris scale. While the standardized coefficient is slightly positive for the Rosenberg scale, its size would have to be twice as large to be considered practically and statistically meaningful. The introduction of additional controls for gender, race and grade does little to alter this conclusion.

Table 3. Self-Esteem Coefficients

Variables	Grades 3 and 5		Grades 8 and 11			
	Beta	p-value	Piers-Harris Beta	Piers-Harris p-value	Rosenberg Beta.	Rosenberg p-value
Americans All	-.106	0.184	-.017	.413	.122	0.13
Pretest Score	.504	0.005	.708	.000	.197	0.04
N	58		91		90	
R^2	28.9%		50.2%		6.9%	

Multicultural Awareness

Table 5 presents the results from an analysis of the program effects on children's multicultural awareness. In this instance, the program effects are strong and statistically significant. Students in the Americans All program exhibited sizable differences in their appreciation and acceptance of cultural pluralism as compared to students in the control group. On average, there was an approximately 30 percent standard unit difference in Americans All participants' appreciation and acceptance of cultural pluralism. This corresponds to a 1.04 unit difference on the appreciation scale (posttest $\bar{x} = 7.42$, s.d. = 1.69) and 0.85 unit difference on the acceptance scale (posttest $\bar{x} = 4.54$, s.d. = 1.41).

Figure 2 gives a visual display of the findings for pluralistic appreciation. It is a scatterplot of the posttest scores against the pretest scores. The two lines show the pooled within-group regressions. The lines differ by a constant of 1.04 units which equals the unstandardized program coefficient. It can be interpreted as the estimated effectiveness of Americans All in promoting students' appreciation of pluralistic values. For fixed values of the pretest, program participants scored, on average, one unit higher on the posttest than students in the control group. This corresponds to a 30 percent standard unit gain.

The results in Figure 2 were obtained by constraining the regression analysis to fit parallel lines. Closer inspection of the data points suggests, however, that the within-group regressions might not be parallel. This possibility can easily be accommodated in the regression analysis by allowing for a group-pretest interaction term. Given the small size of the sample, the interaction term is unlikely to achieve statistical significance. But if it is included, then the group regressions cross at the lower end of the scale. Figure 3 illustrates the effect: the average gain declines as the pretest score becomes lower and rises as the score becomes larger.

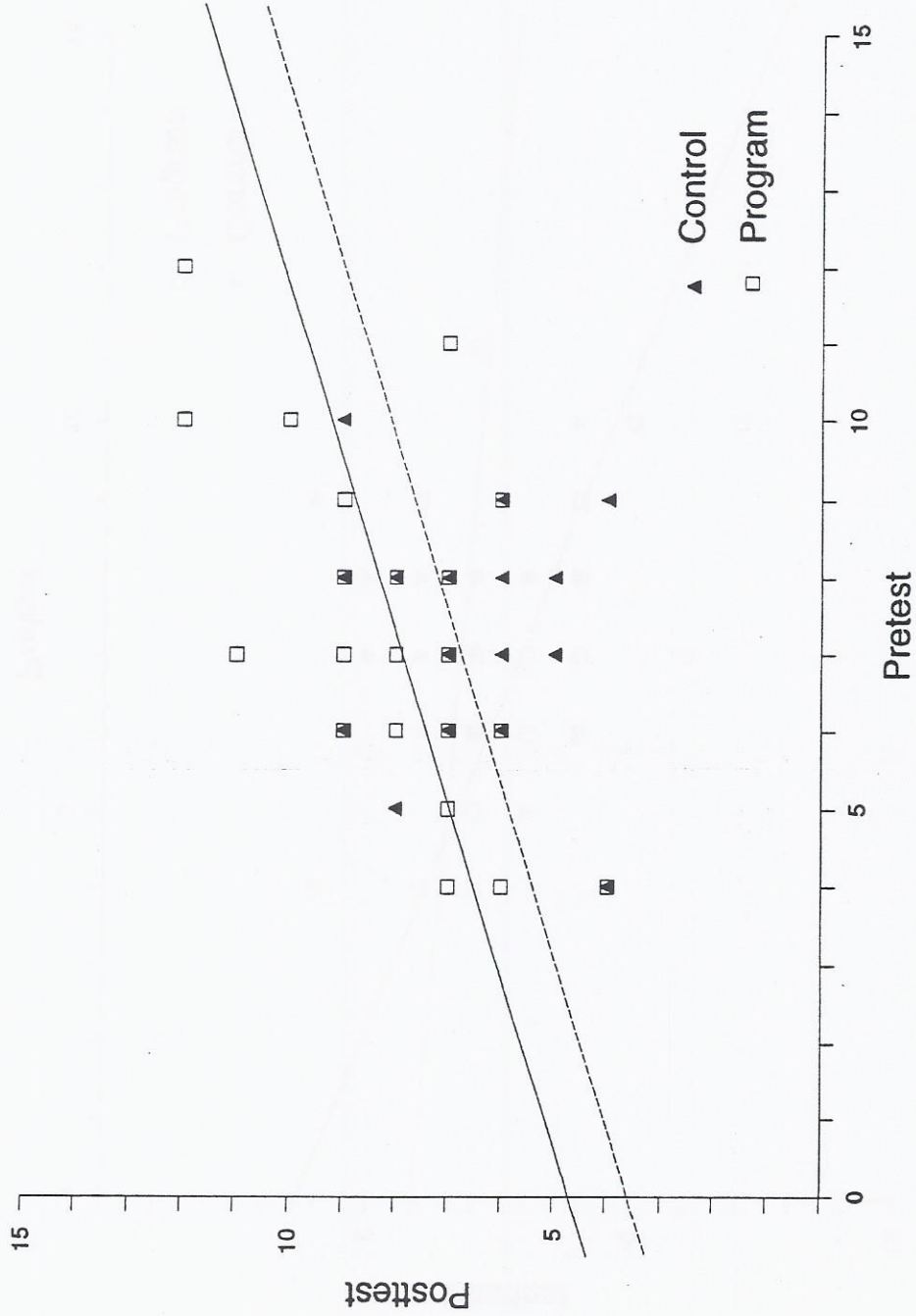


Figure 2. Plot of pluralistic appreciation posttest against pretest. The lines give the pooled-group regressions. Program: — Control: - - - -

Table 6. Ethnic Contributions to American History

Variable	African American		European American		Mexican American	
	Beta	p-value	Beta	p-value	Beta	p-value
Americans All	.227	.032	-.197	.062	-.088	0.465
Pretest Score	.346	.003	.266	.020	.341	0.005
N	61		60		60	
R ²	20.8%		9.9%		11.5%	

Variable	Puerto Rican American		Asian American		Native American	
	Beta	p-value	Beta	p-value	Beta	p-value
Americans All	-.021	.433	-.029	.406	.067	.287
Pretest Score	.389	.001	.478	.000	.478	.000
N	58		57		60	
R ²	15.1%		23.4		24.6	

does point to the need for replicating the study with a larger and more diverse sample of students.

Discussion

Americans All is a national multicultural education program that strives to promote appreciation for self and sensitivity to cultural diversity by drawing on the history of voluntary immigration to the United States and the involuntary incorporation of various ethnic and racial groups through colonization, enslavement, or territorial annexation. The results of this study show that Americans All can be effective in promoting students' sensitivity to cultural diversity. Students who participated in the program exhibited significant gains relative to the control group in their approval of pluralistic values and in their acceptance of foreign-born persons.

The results on enhancing children's self-esteem and reinforcing their sense of cultural identity are less promising. There is little evidence of a significant program impact. One possible reason for the lack of change in students' self-esteem is that the children's scores were already high at the outset. Hence, there was little room for promoting gains. Another is consistent with the idea that self-concept is a relatively stable trait (Blascovich and Tomaka 1991). Like other person traits (e.g. intelligence, Type A behavior), self-esteem cannot readily be manipulated experimentally. Even when dealing with young children, it may be difficult to induce measureable changes in a limited time frame when children are assessing themselves against 8 to 13 years of self-evaluative experiences. This inability to manipulate self-concept experimentally obviously poses difficulties for researchers interested in evaluating interventions designed to raise self-esteem. One possibility is to extend the period of intervention in the hopes that change becomes more observable with time. Another, suggested by Blascovich and Tomaka (1991), is to focus on self-evaluations of very specific or novel attributes. For example, specific

counter this erosion of self-concept by neutralizing negative stereotypes. But as mentioned previously, this intervention strategy is called into question by research showing that minority children have no lower or perhaps higher self-esteem than majority white students (Crocker and Major 1989, Hoelter 1983, Rosenberg 1979).

Another explanation is the theory of self-fulfilling prophecies: minority children do poorly in school because teachers do not expect them to succeed. According to this view, minority students are assigned to remedial classes and inferior schools by means of misclassification and biased testing. Eventually, they may also come to behave in ways that are consistent with these evaluations (Miller and Turnbull 1986). Multicultural education can help to break this self-fulfilling cycle by sensitizing teachers to the cultural biases and prejudices they may bring to the classroom. Although this explanation is generally accepted, Ogbu (1991) criticizes it as failing to explain why immigrant students who attend the same inferior schools do relatively better than native-born blacks and other minorities. For the same reasons, he criticizes the theory of cultural discontinuities which posits that school failures are related to cultural conflicts between teachers and students in styles of instruction and learning. This theory, he notes, also falls to account for why immigrant students with significant language and cultural differences often perform better than native-born students of the same ethnic origins.

Ogbu (1990) argues for a distinction to be drawn between minority students whose presence in the United States is the historical result of voluntary immigration and minority students whose presence is the historical result of involuntary incorporation through colonization, territorial annexation or enslavement. Most descendents of European immigrants and many from Asia fall into the former grouping, while most Native Americans, African Americans, Puerto Ricans, and southwestern Mexican Americans fall into the latter. Ogbu says that immigrant students tend to differ from involuntary minority students in their orientations toward the cultural and language

high academic achievement as "acting white." Black students who wanted to achieve felt pressured to downplay their academic interests and to engage in peer-approved activities that jeopardized their scholastic standing. As Ogbu notes, they face a conflict that does not seem to confront immigrant students between "striving for academic success and maintaining their minority identity and cultural frame of reference" (Ogbu 1991, p. 536).

The Americans All questionnaire for the upper grades included a question: "What are the things about your ethnic or cultural group that make you feel most proud?" Black students less often selected scientific, historical, religious, and academic achievements ($\bar{X} = 5.8$) than white and Hispanic students ($\bar{X} = 6.6$). One way that multicultural education can assist students in crossing cultural boundaries without fearing the loss of cultural identity is to acquaint them with the scholarly and scientific contributions of their ethnic forbears. For example, the Americans All program highlights the scientific, literary, and historical contributions of members of various ethnic groups to American culture. Although this study revealed no significant change in students' attitudes toward such accomplishments, further investment in exploring student sentiments and feelings about cultural identity might help to provide more secure transitional bridges for crossing perceived cultural boundaries.

Challenges of Multicultural Education

As social studies become more diversified, the history of racial and ethnic subordination in this country will become a more prominent feature of the school curriculum. Some people will welcome the change as finally balancing out the scales of historical interpretation. Others will no doubt worry about tipping the scales too far in the direction of an overly ethnic interpretation of American history. Whether one views the coming change as equilibrating or not, it will nonetheless serve as a painful reminder of the terrible compromises the nation has made in extending