

425
✓

AN OFFPRINT FROM VOLUME XIII

JULY 1971 - APRIL 1972

RACE

A JOURNAL OF RACE AND GROUP RELATIONS

THE INSTITUTE OF RACE RELATIONS LONDON

GAJENDRA K. VERMA AND BARRY MACDONALD

Teaching Race in Schools: Some Effects on the Attitudinal and Sociometric Patterns of Adolescents

Readers of this journal will be aware that the field of race teaching, while much discussed, has received scant attention from psychometric research. Were this not the case, there might be no place here for the results of a rather limited attempt to measure the effects on race attitudes and perceptions of a brief exposure to a particular teaching approach. Rather than summarize previous research, *ad hoc* reference will be made in the course of this article to particular studies which influenced it or produced comparable data. It is, however, necessary to outline the context in which the study was set up, and the rationale of the experimental treatment to which the subjects were exposed.

I. BACKGROUND TO THE STUDY

The Schools Council-Nuffield Humanities Curriculum Project has for the past three years been exploring the problems of teaching controversial social and ethical issues to adolescents aged 14 to 16. The work of the Project has been based on five major premises:

1. that controversial issues should be handled in the classroom with adolescents;
2. that the teacher accepts the need to submit his teaching in controversial areas to the criterion of neutrality at this stage of education, i.e. that he regards it as part of his responsibility not to promote his own view;
3. that the mode of enquiry in controversial areas should have discussion rather than instruction as its core;

GAJENDRA K. VERMA is Senior Research Associate, and BARRY MACDONALD Director of the Evaluation Unit, in the Centre for Applied Research in Education, University of East Anglia. The study was financed by the Schools Council.

Race, XIII, 2 (1971)

4. that the discussion should protect divergence of view among participants rather than attempt to achieve consensus;
5. that the teacher as chairman of the discussion should have responsibility for quality and standards in learning.

The problem, as the research team saw it, was how to allow adolescents to reach views responsibly without being restricted by the teacher's bias or subjected to undue pressures by their fellows. They approached this problem by attempting to stimulate and study a pattern of small group discussion in which collections of evidence, printed prose, verse, drama, photographs, paintings, and tapes, were used in a discussion situation which placed both teachers and pupils under the discipline of evidence. They produced collections of materials in such areas as war, education, relations between the sexes, the family, poverty, and people and work. During the years 1968-70, the first collections were used experimentally in thirty-six schools throughout England and Wales, while work proceeded on further collections.

It had been assumed that race qualified as a controversial theme of perennial human concern. But as the time to prepare the race pack drew near it became clear that the theme might present a number of distinct problems. Parliamentary legislation had to some extent put limits upon the expression of certain points of view, and although immigration had become an important issue with the increasing number of coloured immigrants in this country, some doubted whether the atmosphere surrounding the race controversy would provide an adequate safeguard of a manageable enquiry. In spite of the issues involved, a growing number of boys and girls of different colours and races attended the same schools, sat in the same classes, followed the same course. Obviously, one had to take careful precautions with regard to materials which represented a wide range of points of view on race. The next important consideration was whether teachers could press for diversity in discussion, expose issues in all their controversiality and at the same time meet their responsibilities to safeguard the welfare of their students. Some research findings suggested that the teaching of race might only serve to harden prejudice (Miller, 1967).

In view of all these factors, the Project decided that some kind of limited pilot study might provide them with guidance on how best to proceed. This paper therefore reports on a pre/post testing of the attitudinal and sociometric patterns of adolescents, and centres on any changes subsequent to the Humanities Project pilot teaching of race in six schools. There were no preconceived ideas as to what would emerge from the study, and it did not set out to test any explicitly formulated hypotheses. One of the main aims of this pilot study was to examine whether the teaching of race relations according to the principles and methods recommended by the Project has effects upon the attitudes, opinions, beliefs, and sociometric patterns of the students which teachers and others would

regard as undesirable. In the light of these effects the Project could decide whether to proceed in producing a pack of published materials on the race relations theme.

II. DESIGN OF THE PILOT STUDY

The Sample. The experimental population consisted of 226 boys and girls drawn from six secondary schools. Three of these schools were Project schools chosen precisely because their staff and some of their students were familiar with the Project's methods. The three others were non-Project schools, which accepted an invitation to participate. The major consideration in drawing up the school sample was racial composition, the aim being to include schools with widely varying concentrations of immigrant pupils. To minimize cost, choices were limited geographically to the Midlands and South East of England. The following table briefly describes the six schools. Note the discrepancy between schools and experimental groups in terms of racial composition. Group composition was not pre-specified by the Project team.

Not all the pupils involved in the study were tested. Administrative problems in School A reduced the number of pupils available for testing in that school to about one third of its experimental sample. In addition, some pupils were absent from either pre- or post-testing, and a number of pupils failed to complete some of the tests. Neither category was included in the analysis. The average number of pupils whose scores were analysed was approximately 140 for each instrument.

The Teaching Programme. In line with the basic premises of the Project outlined earlier, the teaching approach adopted by all teachers in this study involved an acceptance of the following premises:

- a. that race relations was a theme of intrinsic importance, and, given adequate stimulus, support, and materials, one which should be incorporated into the curriculum of the senior part of the secondary school;
- b. that race relations qualified as a controversial theme;
- c. that the mode of enquiry in controversial areas like race should have discussion rather than instruction as its core;
- d. that for reason b. the teacher should submit his teaching to the criterion of neutrality, i.e. he should regard it as part of his responsibility not to promote his view on the issues under discussion;
- e. that the discussion should protect divergence of view rather than seek consensus.

Participant teachers new to the Project were familiarized with its aims and procedures before the experiment began. During October–November 1970 all six schools carried out a six- to eight-week programme of teaching race relations using a collection of materials

Table 1
Description of the Schools and Samples

School	Location	Type	Sex	Curriculum	Composition of Schools	No. of Pilot Students*		No. of Groups
						Native	Immigrant	
A	Urban	Sec. mod.	Mixed	Project	80% Native 20% Immigrant	45	24	4
B	Rural provincial	Sec. mod.	Mixed	Project	97% Native 3% Immigrant	13	1	1
C	Urban	Sec. mod.	Boys	Project	90% Native 10% Immigrant	36	9	3
D	Urban	Comp.	Mixed	Non-Project	40% Native 60% Immigrant	4	25	2
E	Urban	Sec. mod.	Boys	Non-Project	50% Native 50% Immigrant	15	17	2
F	Rural mining town	Comp.	Mixed	Non-Project	100% Native	37	—	3

The immigrant group consisted of Kenyan Asians, Indians, Pakistanis, and West Indians. All the native students in the

190

RACE

*Year
Groups*

5th and

6th

5th

4th

6th

4th

5th

study were

prepared and supplied by the Project. Exigencies of time and expense caused the trial pack to be an attenuated one. It was nevertheless balanced in that a representative selection of points and views and perspectives on race were included. Enoch Powell's Birmingham speech, editorials from right-wing political journals, photographs of a lynching, of segregated benches in South Africa, of an immigrant facing a 'No Coloureds' notice, and selections from official publications and reports were used extensively by teachers. Films too played a critical part, the most popular being *Flame In the Streets*, *Guess Who's Coming to Dinner*, and documentaries on apartheid. Tape recordings of Martin Luther King's 'I have a Dream' speech, and of an interview with an American Civil Rights worker also proved popular. Clearly the trial pack's deficiencies narrowed the scope of discussion and to the extent that an improved published pack will permit of more wide-ranging discussion and understanding, the effects outlined in what follows are likely to be modified.

The teachers were asked to hold a minimum of three discussion sessions per week throughout the experimental period. In fact most of the teachers chaired discussions for five or six periods per week.

Evaluation. The full evaluation of the pilot study included extensive recording and analysis of classroom discussion, interviews with teachers and pupils, and outside observer reports, as well as the measurement programme described hereafter. The decision to employ a battery of objective tests was taken with the understanding that, given a shortage of time and the limited availability of skilled personnel, a measurement programme was unlikely to be satisfactory as a conclusive gauge of effects. Some limitations of the tests used are discussed in Section IV.

The Tests.

1. *Opinion Questionnaire.* A version of the Wilson-Patterson Conservatism Scale (1968), which has an expanded sub-scale for the measurement of inter-ethnic attitudes, was included in the battery (Bagley, 1970). Cluster and component analyses of the pre-test schedules were carried out, and thirteen items formed a racial cluster which was used for before and after comparisons. The 'racialist' items were randomly distributed among other C-items.

2. *The Sociometric Test.* This was drawn up in a conventional manner. The purpose of this device was to obtain an objective picture of the relationships that existed in groups under study. Six areas of activity (ten items) representing significant aspects of the pupils' social environment were included. They were: friendship, work, play, social activities, leadership, and friendship after school. For each item a first, second, and third choice was called for. An attempt was made to examine the extent to which choices were

restricted to the ethnic in-group, and the number of choices which took place across ethnic group barriers.

3. *Likes and Interests Test* (NFER). In recent years much attention has been given to the study of personality characteristics associated with prejudice. This is based on a general assumption that ethnic prejudice both expresses and reflects the unique and fundamental dispositions of the person. Allport and Kramer (1946) have traced out the consistent way in which strong prejudice connects with the personality as a whole, which seems to be interesting and psychologically plausible. The test included in the battery was developed by Grygier (1961), and aimed at providing a general picture of personality organization in developmental terms.

Two sub-scales, out of thirty, formed the prejudice scale, and these were scored in the present study. Sixteen of the items of the two sub-scales were intended to measure 'Anal Sadism' or 'Fascist attitudes' (As), and 'Insularity' (Ai) i.e. reserve and mistrust, social, and racial prejudice. The rest of the 304 items were supposed to provide a fairly heavy disguise of the test's purpose.

4. *The Situations Test*. The test consisted of four perceptually ambiguous pictures, each representing white and coloured persons in different situations. The testees were asked to describe and to interpret the picture in the form of a story. The stories were then analysed independently by two judges according to a scheme of interpretation (see Appendix). It was hoped that the pupils would project their interpretations on the stimulus materials provided by pictures relating to real life situations, which would give an indication of their preceptions and attitudes in the field of race relations.

5. *The Semantic Differential*. Following the technique introduced by Osgood and his associates (Osgood, 1952; Osgood et al., 1957; Osgood, 1962), the final form of the semantic differential used in this study contained fourteen concepts and twelve scales. The test aimed at obtaining quantitative information about attitudes to twelve environmental and two self concepts. These concepts were selected subjectively from areas which appeared important in the context of race relations (see Table 6). The testees were asked to rate the concepts on a number of seven-point scales defined by opposing pairs of adjectives. The scales chosen were those which Osgood and others had found to have consistently high loadings on the evaluative factor across concepts. This factor may be described as general attitude. To obtain a measure of the testee's attitudes to the concepts, the scores on all the twelve scales were combined for each concept. The scales were:

wise	—	silly	unsuccessful	—	successful	pleasant	—	unpleasant
cruel	—	kind	interesting	—	dull	generous	—	mean
fair	—	unfair	selfish	—	unselfish	dishonest	—	honest
gay	—	sad	exciting	—	calm	tense	—	relaxed

The subjects were tested in October 1970, before the period of teaching, and immediately following its completion in December 1970. A search for shifts in their attitudinal and sociometric characteristics was made in terms of objective data. The purpose of the pre- and post-tests was to examine the degree and direction of shifts on the five measuring devices employed.

III. STATEMENT OF RESULTS

Three empirically defined ethnic groups (Native, West Indian, and Asian) were identified in the total sample. However, before and after comparisons were made between two broad groups, i.e. native and immigrants, on all instruments with the exception of the Sociometric test. In the analysis of test results, means and standard deviations of different variables under study were calculated, and apparent shifts were tested for significance. The broad results are presented in the following tables.

Table 1
Opinion Questionnaire

School	N	Mean R-Score before Teaching	Mean R-Score after Teaching	Significance of Change
A	24	11.80	11.30	NS
B	10	7.90	5.00	10%
C	24	12.35	8.42	5%
D	16	3.78	4.00	NS
E	29	8.20	6.75	10%
F	33	7.38	7.00	NS

The purpose of this scale was to measure inter-ethnic attitudes. The maximum score was 26, a high score (12 or more) indicating hostility to people of other races.

An analysis of shifts by schools (Table 1) shows that in School C the attitudes of pupils changed significantly (at the 5 per cent level) in the direction of tolerance. Schools B and E changed moderately, while three schools showed no appreciable change. Further analysis showing change in the two broad ethnic groups, for boys and girls separately, is presented below.

Table 2
Opinion Questionnaire (Schools Combined)

Group	N	Pre-test		Post-test		Sig. Level
		Mean	S.D.	Mean	S.D.	
Native boys	60	13.21	6.56	11.82	6.30	10%
Native girls	28	5.68	3.68	3.96	3.26	5%
Immigrant boys	31	6.40	3.12	5.08	3.32	NS
Immigrant girls	17	3.94	3.16	4.38	2.44	NS

It can be seen from the above table that many native boys in the sample had hostile attitudes to other races on the pre-test situation, but on the post-test their intolerance was reduced moderately. The

Table 3
(Figures are shown in percentages)

School*		Native Students' Choice for			West Indians Choice for			Asians Choice for		
		Natives	West Indians	Asians	Natives	West Indians	Asians	Natives	West Indians	Asians
A	Pre	83.49	10.09	6.42	25.00	62.50	12.50	60.72	27.52	11.76
	Post	80.58	10.19	9.23	28.84	58.24	12.92	60.56	28.33	11.11
C	Pre	86.92	13.08	—	63.64	36.36	—	—	—	—
	Post	72.83	27.17	—	68.75	31.25	—	—	—	—
D	Pre	26.86	48.35	24.79	15.22	72.16	12.62	26.67	26.67	46.66
	Post	28.58	40.71	30.71	30.09	61.95	7.96	29.49	24.36	46.15
E	Pre	56.98	5.81	37.21	50.00	26.38	23.62	35.88	8.40	55.72
	Post	54.46	5.36	40.18	65.43	5.29	29.28	42.86	4.29	52.85

* Since our primary concern in using this instrument was to examine inter-ethnic choices, both School B and School F were omitted.

native girls, on the other hand, were markedly less hostile to other races than their male peers. Furthermore, after the race teaching programme their R-Scores fell from a mean of 5.68 to 3.96 and this change is significant at the 5 per cent level.

Interpretation of the test responses of non-white students is complicated by the fact that the Opinion Questionnaire, like the Semantic Differential, is not ethnically neutral and its validity as an instrument for measuring the attitudes of non-whites is questionable. Nevertheless, it is interesting to note that the post-test scores of native girls are lower than the scores of both immigrant boys and girls.

The Sociometric Test. Each pupil was asked to choose from among other pupils those with whom he or she would prefer to associate in specified situations. Choices were tabulated for the major ethnic groups. The results are presented in Table 3.

It should be kept in mind when inspecting the sociometric data that the proportions of each of the three groups in each school varied considerably (see description of samples). This is a variable which is likely to influence the extent of in-group preference. The present study was concerned with the extent and direction of preference changes.

In School A shifts were slight, but in the direction of out-group preference.

In School C both ethnic groups, native and West Indian, showed in the post-test a reduction of in-group preference.

In School D all three major groups increased their preference for native pupils, at the expense of the West Indian group. (This means of course a decrease in in-group preference on the part of the West Indian group.)

In School E there were reductions in in-group preference on the part of all three groups, again at the expense of the West Indian group. (There were only two West Indian pupils in this school sample.)

Taking the schools as a whole, the tendency was for all groups to increase their out-group preference. In the case of native pupils the shift was towards Asians where they had a choice, whereas both West Indians and Asians showed increased preference for native pupils. The general extent, however, of the shifts was slight.

Table 4
Likes and Interests

School	N	Pre-test		As Scale Post-test		Sig. Level	Pre-test		As Scale Post-test		Sig. Level
		Mean	S.D.	Mean	S.D.		Mean	S.D.	Mean	S.D.	
A	23	4.61	1.55	4.83	1.55	NS	7.09	2.28	7.13	2.17	NS
B	10	3.60	1.68	3.40	1.20	NS	5.40	1.96	5.40	1.96	NS
C	24	3.58	1.18	4.42	1.78	10%	6.13	2.37	6.29	1.59	NS
D	18	4.89	1.59	4.44	1.83	NS	5.78	1.36	5.83	1.61	NS
E	29	4.62	1.63	4.86	1.69	NS	5.55	1.85	5.41	1.99	NS
F	34	4.12	1.32	4.47	1.78	NS	5.85	1.89	5.91	1.56	NS

Inspection of the above table indicates very little variation from one testing situation to another. No significant differences were found between the means on either scale, except in School C (As Scale). This shows a moderate shift in the direction of intolerance.

It should be recalled that the test was made up of thirty sub-scales, dealing with thirty different personality traits. The mean score of the test standardization was about 5, a low score (5 or less) always meaning liberal and tolerant personality characteristics (As) and sensitivity to social and racial discrimination (Ai). The average mean score of our samples was in the region of 5. According to the general psychological meaning given to the two personality variables, pupils' initial attitudes (at the pre-test stage) were moderate except in school A (Ai scale), and the teaching of race relations had not been followed by any significant change. The results of this study do not support Miller (1967) who, using the same test, found that the level of prejudice increased significantly after the teaching of race relations. His sample's score average was 7.5. Students' exposure to the teaching programme in the present study was between twenty and thirty hours compared with a three-hour (maximum) exposure in Miller's study.

Table 5
The Situations Test

School*	N	Pre-test		Post-test		Sig. Level
		Mean	S.D.	Mean	S.D.	
A	23	28.17	3.39	29.25	3.09	NS
B	11	26.45	3.65	26.36	4.68	NS
D	16	26.31	4.63	28.81	3.88	10%
E	30	25.53	4.48	26.87	4.09	NS
F	33	26.21	4.006	26.61	4.59	NS

* Data from school C were not available.

The results of the Situation test follow a pattern of the Likes and Interests test. This instrument did not reveal significant shifts on the mean score in four schools. The pupils in School D, however, obtained significantly higher mean value at the post-test; the critical ratio was significant at the 10 per cent level of confidence.

The purpose of this test was to gain an understanding of the pupils' perception in the field of race relations. The test utilized four pictures of coloured and white people in ambiguous situations and relationships to which the testee had to respond in the form of stories. The stories were marked on a 5-point scale. For each pupil the final score was obtained by totalling the scores that were assigned to the four stories by two judges; hence the maximum possible score was 40 and the minimum 8, the average being 24. The basic assumption was that the way the testee perceives and structures some stimulus is a function of his attitudinal and personality characteristics.

The codification in terms of ethnocentrism was on a continuum, a high score (24 or more) meaning tolerance of perceptual ambiguity. The results of this test have to be weighed in the light of inter-scorer and inter-story reliabilities which were regarded as reasonably satisfactory (see Appendix).

There were no significant differences between the mean scores attained on the pre- and post-tests although the scores on the second test were slightly changed in the direction of tolerance. It is possible that this finding is specific to the sample used, which was composed of pupils with already high initial scores and hence the programme had no effect on their perceptual dispositions. However, the data obtained from this instrument should be interpreted in the light of the reservations outlined in the Appendix.

Table 6
The Semantic Differential (Schools Combined)

Concepts	Native N = 88		Immigrants N = 54	
	Pre-test Mean	Post-test Mean	Pre-test Mean	Post-test Mean
Coloured immigrants	45.32	47.05	61.12	61.95
Myself now	58.03	60.55	65.87	65.05
Mixed race marriage	45.08	48.89	47.97	53.17
Right of Negroes	43.47	48.65	55.85	59.47
Myself in 5 years' time	63.92	64.70	67.40	69.17
Coloured equality	46.36	51.00	60.72	61.35
West Indians	45.01	48.74	59.75	59.85
Pakistanis	38.02	45.27	48.20	50.02
Apartheid	36.04	34.72	33.72	27.02
Jews	39.21	40.92	43.40	47.67
Rule by minority	39.54	37.89	35.87	35.72
Chinese	43.51	44.31	52.22	47.40
Coloured bus conductors	51.65	52.60	60.80	57.70
Jamaican neighbours	50.54	51.06	57.40	58.45

Table 6 shows the results for the complete sample tested on the Semantic Differential. The range of possible scores on each concept is from 12 to 84. The average score for each concept is 48, with a high score always meaning a more positive concept in terms of evaluative attitude. It should be remembered that the 'favourable' extreme of the scales was sometimes on the right hand and sometimes on the left, to try to counteract response set and 'halo' effect.

It can be seen from Table 6 that appreciable differences were found between the two testing occasions for both native and immigrant students, showing rises at the second testing in the direction of 'favourable' evaluation. It is interesting to note that the immigrants' self-assessment (present self and future expectations) was higher than the natives'. As might be expected, the immigrant group regarded all the interracial concepts as highly significant and there was a tendency to attach more positive attributes to these.

concepts after the teaching of race relations. The native group, on the other hand, evaluated some concepts (mixed race marriage, rights of Negroes, coloured equality, West Indians) negatively at the pre-test session, but after the teaching programme these concepts received higher positive attributes from them.

This instrument revealed many similarities on the attitudinal patterns of the two groups. Both seemed to have low regard for 'apartheid' and 'rule by minority'. They showed increased disapproval of these concepts on the post-test. It is rather surprising that the native group rated 'Pakistanis' and 'Jews' more negatively than other ethnic group concepts although the post-test mean values increased nearer to the average. This pattern suggests that the groups were not rejecting people of other races on the basis of colour, which is in line with the findings of Hill (1970). An analysis in terms of shifts clearly shows that following the teaching programme both native and immigrant pupils were more favourably disposed towards inter-ethnic relationships.

IV. DISCUSSION AND CONCLUSIONS

The present pilot study was concerned with examining the effects, if any, of race teaching according to principles recommended by the Project, upon the attitudes, opinions, personality, and friendship patterns of students. The limitations of this study should perhaps be considered before interpreting the results.

Limitations of the Design. Firstly, six to eight weeks is a very short period of time between the use and re-use of test instruments of this kind. Ideally, parallel forms of tests should be employed in such circumstances to avoid distortion due to recall and recognition, but these were not available. Secondly, it was not feasible for one person to administer the test programme in all six schools. Three people shared the task, and therefore an administrator variable was introduced. Thirdly, the sample was not representative of the adolescent population as a whole, since criteria derived from the interest of the Project's main research were employed.

Limitations of the Tests. The weaknesses of the Situation test are discussed in the Appendix. Of both the Semantic Differential and the Opinion Questionnaire it must be stressed that whereas their usefulness as measures of the attitudes of the white population is not questioned, their ethnic bias raises problems of interpretation regarding the responses of non-white subjects. Initial trials of the Semantic Differential included a more ethnically balanced range of concepts but statistical analysis of pilot work led to the exclusion of seven concepts and imbalanced the contents in its final form. Further trials are necessary to elicit viable concepts which would redress the balance but these could not be carried out before the present study took place.

use of the Likes and Interests test which consists of only items forming two sub-scales, abstracted from a test of ability development involving over 300 items and thirty sub-items may also be open to question. It was used in this experiment for comparison with the widely publicized findings of Miller

conclusions. We have thought it important to stress the limitations of the programme, to ensure that the data obtained are treated with caution, but it may still be useful to summarize the trend of the findings.

Given the prudential nature of the concern which motivated the setting up of this pilot study, the major finding is that no marked change was manifested in the attitudinal or personality characteristics of the students was manifested in their test responses after exposure to the teaching programme. The effects of the experiment, although not statistically significant, tended to suggest a shift in the direction of interracial tolerance.

The results of the Opinion Questionnaire (Table 2) lend support to the general assumption that greater understanding leads to greater tolerance and many studies designed to reduce intolerance are based on this hypothesis (Miller and Briggs, 1958). The data obtained by the geometric method indicate a clear pattern with regard to the magnitude and direction of shifts. It is not evident that the minority group showed a greater tendency towards in-group selection and rejection of the out-group than the majority group. The present findings (Table 3) are in considerable agreement with earlier work on friendship choice (Silberman and Spice, 1950). It would seem reasonable to say that in this research the teaching programme was evaluated by shifts in the pattern of friendship selection in the direction of interracial tolerance. In contrast to Miller's (1967) findings, there were no signs to suggest that the teaching of race relations had led to change in an adverse direction as assessed by the Likes and Interests test (Table 4). With regard to the Semantic Differential there were shifts in respect of most concepts evaluated, and with the exception, all shifts were in the direction of interracial tolerance (Table 6).

The combined picture of the results seems to indicate that there is no general tendency towards intolerance after a seven- to ten-week teaching programme. There is no evidence to suggest that the students generally became less sensitive to or tolerant of members of other racial groups. These results cannot be considered as constituting proof. Analysis of the pilot study along other lines is incomplete, but a decision has already been made to proceed with the setting up of a full collection of materials on race, on the grounds that some of the problems encountered in the course of the study justify the abandonment of further research. No teacher involved in the programme abandoned the course, or found it necessary to reject any of the premises described earlier. In February

1971, each of the schools which participated sent team members to an evaluation conference at which they expressed willingness to undertake the teaching of race with other students in the future.

APPENDIX: THE SITUATIONS TEST^a

This test was devised in an effort to diversify the range of instruments employed in the measurement battery. As the study was exploratory, it was thought that an indirect technique of a projective nature might yield useful data. Additional support for this approach was derived from consideration of the fact that the subjects were likely to be highly conscious of the nature and broad purposes of the study. A device utilizing disguised and unstructured stimuli might reduce the degree of guardedness which normally influences subject response in more 'hard-nosed' approaches.

The technique used has its origin in the well-known Thematic Apperception Test (TAT). Considerable criticism has been directed against projective techniques of this kind on the grounds that they are unstandardized in administration and subjective in scoring. Furthermore the validity of the techniques has been, and continues to be, a major issue in the measurement field. While they are widely used by clinicians, their adaption to psychometric procedures is still fraught with problems. Even their opponents, however, would agree that such techniques are least open to objection when they are specially designed to focus on a particular issue or problem, and when the data they elicit is used 'incrementally' as a supplement to information gained by other means. Both conditions were met in this particular case.

At the design stage of the pilot study, preliminary discussions about the possible effects of race teaching on adolescent pupils raised the question of whether it was possible to distinguish between shifts of perception and shifts of attitude, it being held by those who favoured the distinction that whereas a short exposure to teaching in this field was unlikely to shift race attitudes it was bound to strengthen pupils' dispositions to perceive relevant social situations in racial terms. Following up this line of reasoning, an attempt was made to devise an instrument which would isolate perception and enable the hypotheses to be tested. In the event, those involved in the development and use of the instrument failed to agree about the nature and validity of the distinction and, as a result, the instrument which emerged, the Situations test, was interpreted as a measure of ethnocentrism in which perception and attitude were undifferentiated.

The trial material consisted of a set of seven pictures, each showing coloured and white persons in ambiguous situations and relationships. The pictures were subjectively chosen because they appeared to be minimally structured. These stimuli were subjected to a trial testing on a group of 228 students, 14 to 15 years old, drawn from six schools. The trial schools were randomly selected from a wide area of England and Wales, representing both urban and rural settings.

In administering and scoring the test, the instructions of McClelland and others (1953) were strictly followed. Each picture was shown for thirty seconds; following each presentation the subjects were given four minutes to write a story, invented on the spur of the moment.

The stories were rated independently by two judges (psychologists), using a five-point scale, in terms of the pupils' perception of and attitude towards race relations. Judges were asked to take into account the intensity, duration, and frequency of race-oriented expressions in the language and content of the stories.

^a Copies of the pictures, for information only, may be obtained on loan from Dr. Verma at the University of East Anglia.

The results of the inter-rater agreement are shown below:

Table 7
Inter-rater Agreement: *r* Values

Judge	Picture						
	I	II	III	IV	V	VI	VII
A-B	.54	.76	.73	.69	.71	.61	.52

Considering the projective nature of the test, scorer agreements on four pictures (II to V) were reasonably high, despite the fact that the judges reported some difficulty in maintaining an adequate frame of reference in the interpretation of the stories. In an attempt to find additional reliability for the test inter-story correlations were calculated. The reliability coefficients ranged from .47 to .81. This was regarded as moderately reliable and consistent.

In the pilot study of race four pictures (II, III, IV, and V), which yielded *r* values above .65, were used. The procedures followed in the administration and scoring of the test remained the same. The analysis of inter-rater agreement on the four pictures revealed satisfactory correlations of .68, .69, .69, and .72.

The limitations of this technique should be kept in mind. Firstly, the development of this instrument has been continuously bedevilled by arguments about the significance of perception as a reflection of attitude. Consequently the scheme of interpretation communicated to the scorers lacked precision. The fact that inter-scorer and inter-story correlations were relatively high probably indicates the operation of crude categories and only broadly comparable frames of reference. Secondly, the problem was compounded for the judges by the brevity of many of the stories which they were required to score. To put it simply, although both scores and scorers were reasonably consistent, it is difficult to say precisely what the scores stand for. What we have is an ill-defined assessment of ethnocentrism which does not seem to distinguish between perception and attitude.

Additional weaknesses, such as its shortness (only four pictures), and the use in re-testing of the same stimuli, further undermine the usefulness of this instrument to the point where it became necessary for us to suggest that the obtained results should be treated with extreme caution. Further research with a refined and clear-cut scheme of interpretation is needed.

Our experience of working on this instrument reinforces many of the widely held reservations about the problems inherent in such devices. At the same time it should not be forgotten that our time was limited, and the development of sound projective instruments requires very extensive trial and analysis, which we were unable to devote to it. We think that such techniques are worth developing, principally because they are less susceptible than others to distorting motivational influences in the subjects. This may be particularly important in the assessment of school programmes and if, as we suspect, the teaching of race becomes a standard element in secondary school curricula, more and better efforts will need to be made in this field. This Situations test will undergo further work to eliminate some of the deficiencies. In the meantime it is here reported in some detail as a cautionary tale.

References

- Allport, G. W. and Kramer, B. M., 'Some Roots of Prejudice', *J. Psychol.* (22, 1946), pp. 9-39.
 Bagley, C., 'Racial Prejudice and the Conservative Personality: A British Sample', *Political Studies* (18, 1970), pp. 134-41.
 Grygier, T. G., 'Likes and Interests Test', *N.F.E.R.* (1961, Copyright).
 Hill, D., 'The Attitudes of West Indian and English Adolescents in Britain', *Race* (XI, 1970).

Miller, H. J., 'A Study of the Effectiveness of a Variety of Teaching Techniques for Reducing Colour Prejudice in a Male Sample (aged 15-21)', M.A. Thesis, London University (1967).

Miller, K. M. and Briggs, J. B., 'Attitude Change through Indirected Group Discussion', *J. Educ. Psychol.* (49, 1958), pp. 224-8.

McClelland, D. C., Atkinson, J. W., Clark, R. A., and Lowell, E. L., *The Achievement Motive* (New York, Appleton-Century-Crofts, 1953).

Osgood, C. E., 'The Nature and Measurement of Meaning', *Psychol. Bulletin* (49, 1952), pp. 192-237.

Osgood, C. E., 'Studies on the Generality of Affective Meaning System', *American Psychologist* (17, 1962), pp. 10-28.

Osgood, C. E., Suci, G. J., and Tennenbaum, P. H., *The Measurement of Meaning* (Urbana, University of Illinois Press, 1957).

Wilson, G. D. and Patterson, J. R., 'A New Measurement of Conservatism', *Brit. J. Social and Clinical Psychol.* (22, 1968), pp. 9-39.

Minerva

A REVIEW OF SCIENCE, LEARNING AND POLICY

VOLUME IX NUMBER 3

JULY 1971

EDITORIAL

No Salvation outside Higher Education

ARTICLES

Science Policy in American State Government *Harvey M. Sapolsky*

The Support of Science without Science Policy in Australia *S. Encel*

The Urban University and its Urban Environment *Kermit C. Parsons and Georgia K. Davis*

'World Leadership', the 'Technological Gap' and National Science Policy *Richard R. Nelson*

DISCUSSION

Consultation or Voting Power *Eric Ashby and Mary Anderson*

Reflections on a Debate *James Drever*

Letters from *Harold Orlans, F. E. Balderston and Thomas Walsh, G. D. Parikh, Michael Moravcsik*

Reviews by *Jürgen Habermas, Fritz K. Ringer, Norman Kaplan, John Ziman*

Single copy £1.00/\$3.00 Annual subscription £3.50/\$10.50

4 LITTLE ESSEX STREET, LONDON WC2R 3LF
