FAIR TESTING IN EMPLOYMENT SELECTION

RAJVINDER SINGH KANDOLA

SUBMITTED FOR THE DEGREE OF Ph.D. TO
THE UNIVERSITY OF ASTON IN BIRMINGHAM

July 1985

UNIVERSITY OF ASTON IN BIRMINGHAM

SUMMARY

FAIR TESTING IN EMPLOYMENT SELECTION

RAJVINDER SINGH KANDOLA

Submitted for the degree of Ph.D., July 1985

This project represents the collaboration of Charta Mede Ltd and the Interdisciplinary Higher Degrees Scheme at the University of Aston.

The aim of the project was to monitor the effects of the Civil Service's Executive Officer Qualifying Test Battery on minority group applicants. Prior to monitoring the EO Test Battery, however, an ethnic classification had to be developed which was reliable, acceptable to respondents and appropriate for monitoring. Three pilot studies were conducted to examine these issues, during which different classifications and different ways of asking the question were trialled. The results indicated that by providing more precise instructions as to the meanings of categories, it was possible to obtain classifications which were acceptable and reliable. However, there were also certain terms and expressions which should be avoided such as those referring to colour and anthropological racial groups.

Two classifications were used in the Executive Officer Study - one derived from an Office of Population Censuses and Surveys classification and one developed for this project - the Multi-Cultural British Classification. The results indicated that some minority groups (Asians, West Indians and Africans in particular) pass the tests in significantly lower proportions than the majority group and also score significantly less well on the tests. Factors which were significantly related to pass/fail and test scores included educational qualifications and age on entering the UK (the latter being negatively correlated).

Using variables in this study, however, it was only possible to account for 5% of the variance in pass/fail rates and 11% of the variance in test scores. Analyses of covariance carried out indicated that the differences in test scores still remained even though the effects of significantly correlated variables were removed.

Although indirect discrimination could not be inferred from the data, further research into differential validity and fairer methods of selection is needed.

Key words

Race, Discrimination, Tests, Equal Opportunity

ACKNOWLEDGEMENTS

There are many people to whom I am indebted for their help and encouragement during the course of this project. In particular I should like to mention the past members of my supervisory team - Ivan Robertson, Michael Pearn and Martin Edis; John Barker and Sarah Paul at the Civil Service's Recruitment Research Unit for their hard work and cooperation in setting up the Executive Officer Study; the staff at Rubery Hill Hospital and at the Coventry and Birmingham Careers Offices for cooperating with the pilot studies; and to Alastair Cochran and Mike Hussey who have willingly provided assistance whenever it was required.

Finally, I wish to give heartfelt thanks to two people without whose support, encouragement and hard work this thesis would not have been completed. They are Heather Graham who has done a wonderful job of typing up this thesis, and my supervisor John Patrick whose interest in and commitment to the project has been immense. To John and Heather, thanks a lot and here's hoping you do not have to go through an experience like this again.

Please note that the materials contained in Appendix 1 ("Examples of Questions in the Executive Officer Qualifying Test Battery") are Crown copyright. I acknowledge gratefully the cooperation of the Civil Service Commission in allowing the materials to be reproduced here.

DISCLAIMER

It is important to note that the views expressed in this thesis are entirely those of the author, and do not necessarily represent the views of the Civil Service Commission.

CONTENTS

Page No.

Page No.

Page No.

Page No.

Page No.

LIST OF APPENDICES

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Participating Organisations and Outline of the Thesis

This project represents the collaboration of two organisations - Charta Mede Ltd, and the Interdisciplinary Higher Degrees Scheme of the University of Aston in Birmingham. Background information is provided on both of these organisations and their basic philosophies. In addition, summaries of each of the chapters in this thesis are provided.

1.1 The Participating Organisations

1.1.1 Charta Mede Ltd

Charta Mede Ltd grew out of the Runnymede Trust, and so in explaining its work, some mention is needed of its history.

The Runnymede Trust was established in 1968 as an educational charity whose deeds define its purposes as "the promotion of studies in political, economic, social and natural sciences, the humanities and other subjects or disciplines suitable for education; and the study and dissemination of knowledge of problems of and connected with ethnic, racial, national and other similar groups in England or otherwise and their effect upon the environment". (Runnymede Trust, 1968).

It attempts to be non-political and to provide reliable, valid and accurate information on all matters concerning race. It was felt, however, that there was a need for a specialist service to aid organisations in dealing with problems of discrimination in

employment. Consequently, in 1978, the Runnymede Industrial Unit was set up, and it had three basic functions:-

(i) research and development work

(ii) training and consultancy work

(iii) publication and dissemination of information and other advisory material.

It conducted work in all three areas and in addition published material on equal opportunities, ran seminars, etc.

Nevertheless, it was felt by the former director of the Unit, that some specific research on selection fairness issues should be conducted. To this end a postgraduate student was taken on from the Interdisciplinary Higher Degree Scheme at Aston University to investigate the issues of ethnic monitoring, and indirect discrimination of minority groups through the use of tests. The project formally began on 1 October 1980.

In 1981 plans were formulated for the Unit to formally dissociate itself from the Trust and to become a private consultancy. On 1 January 1982 Charta Mede Ltd formally came into being. Although a company limited by guarantee, Charta Mede is a charity. When the project was being conducted, it comprised three full-time members and one part-time member of staff.

1.1.2 The Interdisciplinary Higher Degree Scheme (IHD)

The IHD scheme was set up by Aston University in 1968, and was funded by the Joint SRC/SSRC Committee. The primary aim of the IHD scheme is:

"To equip postgraduate students for positions of responsibility in industry and publicly financed organisations by providing training in practical, real world problem-solving; more specifically in:

* application of existing knowledge
* generation of new knowledge needed
* appreciation of academic disciplines different from those of students' first degrees
* appreciation of practical constraints in real world problems
* implementation of solutions" (Cochran 1981)

The scheme operates by having a student attached, usually as a temporary employee, to an industrial or publicly-financed organisation.

The projects have to be large enough to warrant three years work, but cannot be so urgent that a solution is required in a few weeks. Projects should also involve some research, but not of the single discipline laboratory type - hence the word 'Interdisciplinary' in the scheme's name.

Each student is supervised by a small team. The main supervisor is chosen because of his/her expertise in the discipline considered to be of principal relevance. They have ultimate academic responsibility for the project. The Associate Supervisor - from a different Faculty - contributes in his/her area of expertise.

The Industrial Supervisor ensures that the project remains relevant to the collaborating company's needs, and also takes responsibility for the student within the company. The team is completed by the IHD tutor, who plays a co-ordinating role, and

tries to ensure that a balance is maintained between all the inputs to the project and generally monitors how well the basic aims of the IHD scheme are being realised.

1.2 Summaries of the Thesis Chapters

The following section provides summaries of each of the chapters in this thesis. This is intended to aid and to guide any readers through all of the material and to outline the general structure of the research.

Chapter 2 Anti-Discrimination Legislation and Equal Opportunities and Their Application in the Civil Service

The numbers and distribution of the ethnic minority groups are shown, together with research evidence on the nature and extent of racial discrimination within the United Kingdom.

To protect people from such discrimination the Race Relations Act of 1976 was passed. This contains two major definitions of discrimination - direct and indirect. It is the latter which is the particular concern of this project. In addition, the Commission for Racial Equality (CRE) has recently published its Code of Practice which is designed to help employers ensure that equality is achieved in their workplace. The Code has important implications for employers - particularly the sections relating to ethnic monitoring.

The organisation whose equal opportunity policy has come under the most scrutiny is that of the Civil Service. The Tavistock

Institute's report on the Civil Service (HMSO 1978) revealed that the policy is not working efficiently and one area where indirect discrimination could be occurring is in the use of the selection tests. Civil Service concern over this issue led to this project, and in particular the monitoring of the effects of the Executive Officer Qualifying Test Battery on minority group applicants. Details of the test battery and of the validation project on the tests are also provided. Finally, the objectives of the project are described.

Chapter 3 Psychological Testing and the Issues of Bias and Fairness

Many definitions of a psychological test exist, and by examining these an attempt is made to provide a definition for use in this project.

The use of tests is placed into a historical perspective and in doing so it can be seen that the principles of good test use were laid down many years ago. It is failure to adhere to these principles which has led to some of the problems of bias and unfairness.

The major sources of bias are identified together with solutions to overcome them. Two other areas, ie. differential validity and models of test fairness, are briefly discussed because, although they are only indirectly related to this project, they are nevetheless very important issues when bias in tests is being considered.

Chapter 4 The Overall Objectives and Methodology of the Pilot Studies Comparing Classifications For Use in Monitoring

There has been considerable debate about the need for ethnic monitoring. Some people feel it is discriminatory, others that it is impractical. Many methods of classifying people exist, but only organisations such as the Office of Population Censuses and Surveys have attempted to research the area. Many of the problems connected with monitoring are those dealing with terminology and suitable question format. The difficulties associated with finding such a question eventually led to the decision to leave out from the 1981 census a question on racial or ethnic origins.

Before any examination of the Executive Officer tests could begin, therefore, a systematic, reliable and acceptable way of classifying people had to be determined. This led to the first stage of pilot studies on classification systems.

Chapter 5 Comparison of Classifications For Use in Monitoring for Employment Selection: Stage I of Pilot Studies

A number of pilot studies were carried out with the overall objective to examine and compare different types of classification which could be used in the monitoring of applicants for employment selection. Five classifications were chosen, and the question was asked in a 'standard' way (ie. which group do you think you belong to?) and a 'reversed' way (ie. which group do you think someone else might think you belong to?)

Two pilot studies were conducted, each examining a different aspect of the problem. The first examined the suitability and objectionability of the classifications from the respondent's point of view. The second examined test-retest reliability.

Chapter 6 Comparison of Classifications For Use in Monitoring for Employment Selection: Stage II of Pilot Studies

As a result of the first set of pilot studies some classifications were thought to be of potential use in the Executive Officer study. Certain interpretative problems still had to be overcome, therefore refinements were made to these classifications and some newer ones were included in the second stage of pilotting. The major change was to include detailed instructions on many of the classifications as to what each category meant.

Two classifications were chosen for use in the Executive Officer study.

Chapter 7 The Executive Officer Study: Method and Sample Details

Although an investigation on the use of tests has not occurred in the United Kingdom Civil Service before, such an investigation was carried out in the American Civil Service. The nature of that study and its results have relevance to this project and so are summarised here.

Also, details of the development of the questionnaire containing the two classifications, the sampling technique and the sample itself are provided. In total, information was obtained on 3,449 candidates of whom 474 were from minority groups.

Finally, the objectives of the Executive Officer study are delineated.

Chapter 8 Results: Examination and Comparison of Pass Rates Between Groups on the Two Classifications

The first stage of the results was to examine and compare the pass rates of each of the groups on the two classifications.

Chapter 9 Results: Examination and Comparison of Test Scores Between Groups on the Two Classifications

The Executive Officer Qualifying test scores are examined in this chapter, and in particular the significant differences in the overall test score and individual test scores between the groups in the two classifications.

The results indicate that there are particular groups which appear to perform consistently less well on the tests.

Chapter 10 Results: Examination of Factors Related to Pass/Fail and Test Performance

In Chapters 8 and 9 no attempt was made to examine the factors related to the passing/failing of the tests or to test performance. In order to determine these, the multivariate techniques of discriminant analysis and multiple regression analysis were applied to the data.

An analysis of covariance was performed on the test scores to examine the effect partialling out significant variables had on them.

Chapter 11 Results: Examination and Comparison of Candidates' Reactions to the Classifications

An important part of the Executive Officer study was to determine candidates' reactions to each of the classifications. They were asked to categorise themselves on two classifications, and for each one they were asked how they would react if that classification were actually used to monitor candidates. Their reactions are examined in this chapter.

Chapter 12 Conclusions and Overview of the Project

The final chapter draws together the various findings of the project and presents some conclusions. In addition, the type and nature of future research are discussed.

CHAPTER 2

Anti-Discrimination Legislation and Equal Opportunities and Their Application in the Civil Service

The research evidence on discrimination has been steadily accumulating since the mid '60s and it reveals, consistently and overwhelmingly, that minority groups, particularly black groups, face discrimination at all levels of employment and it does not matter whether the people are British and British-born or immigrants. This evidence led to the Race Relations Act of 1976 which incorporates two notions of discrimination - direct and indirect - and it is the latter which is the interest of this project. To help employers ensure equality of opportunity the Commission for Racial Equality has published its Code of Practice which has important implications for organisations.

The organisation which has come under most pressure to have an effective equal opportunity policy is the Civil Service. Monitoring studies carried out, however, particularly by the Tavistock Institute (HMSO 1978), reveal that the policy is not working efficiently. One area of special concern is that of the use of selection tests and the possibility that they might be indirectly discriminating. This concern in the Civil Service expressed itself in their willingness to allow a project monitoring the effects of the Executive Officer Qualifying Test Battery on minority group applicants. Details of the test battery and of a validation project are provided.

2.1 Evidence of Racial Discrimination

2.1.1 Employment and Unemployment of Racial Minorities

In 1971, the population of the United Kingdom included 3 million people born in foreign countries, of whom 1.2 million were born in the New Commonwealth and Pakistan.* Of these 1.2 million, 0.3 million were white. Of people born in Britain 0.5 million were black, bringing the total black population of Britain to 1.4 million people. By 1977 this figure was estimated to be 1.85 million or 3.4% of the total population.

The black population is not evenly distributed throughout the United Kingdom - it is numerically larger in some areas than others.

The areas in which black people have chosen to live is a direct consequence of where the immigrants settled in the 1950s and 1960s. The most favoured areas were the larger connurbations where employment was to be found, in particular London and the South East, the West Midlands and the North West. A more detailed breakdown between districts is provided in Table 2.1.

* The definition used here is that contained in "Britain's Black Population" (Runnymede Trust and Radical Statistics Race Group 1980). New Commonwealth (NC) refers to ex-British colonies, most of which now belong to the Commonwealth and whose citizens can be described as black. This includes all Commonwealth countries except Australia, New Zealand and Canada (Old Commonwealth). Pakistan left the Commonwealth in 1973 and so is referred to separately.

Table 2.1 Population Born in New Commonwealth in 1971 (areas with more than 20,000 NC born)

	Number of NC Born (thousands)	Percentage of total population NC born
GREATER LONDON	476	6.4
BIRMINGHAM	64	6.3
LEICESTER	23	8.2
BRADFORD	23	4.9
WOLVERHAMPTON	20	7.4

(Source: Runnymede Trust and Radical Statistics Race Group, Britain's Black Population 1980)

In comparison with the total population, the black population is young. In 1971 40% of the black population were born in the United Kingdom; 41% were under 15 years of age and 49% aged between 15 and 44 years which compares with 24% and 39% for the total population. One effect of the age differences is that a relatively high proportion of black people are working - 91% in comparison with 77% of the total population.

Despite working in a wide range of industries black people tend, nevertheless, to be concentrated in the unskilled and semi-skilled jobs (see Table 2.2). Some of the reasons why the black population work in such low status jobs are that:

(i) this was usually where the gaps in the labour market occurred.

(ii) these were the occupations which offered employment in the 1950s and 1960s.

(iii) some people, especially those who have not been resident in the UK for too long, may not yet have acquired the skills and language for work in higher grades.

(iv) movement out of these industries and occupations is hindered by racial discrimination.

Table 2.2 Percentage Socio-Economic distribution: Economically Active Males and Females, 1977

Socio-economic group	West Indies	India*	Pakistan & Bangladesh	'Whites'
Professional	0.4	7.5	4.2	4.2
Employers & Managers	2.1	7.0	4.5	11.9
Other non-manual	25.8	21.7	9.1	31.0
(Total non-manual)	(28.3)	(36.2)	(17.8)	(47.1)
Skilled manual	28.2	27.1	29.1	27.6
Semi-skilled manual	32.2	27.0	37.6	18.7
Unskilled	11.1	9.6	15.6	6.7
(Total manual)	(71.5)	(63.7)	(82.3)	(52.9)

* The 'Indian' group contains relatively large numbers of white collar employees who, though born in the Indian sub-continent were in reality of British descent.

(Source: Field, et al. Ethnic Minorities in Britain 1981)

Furthermore, according to both Department of Employment figures and the P.E.P. survey (Smith, 1974), unemployment amongst black people increases disproportionately when the general level of unemployment is rising. This is borne out by recent evidence. Between November 1973 and February 1980, unemployment as a whole doubled - but the number of black unemployed quadrupled. (Field et al 1981).

Several factors have been suggested to explain this, which include:

(i) increasing numbers of young black people leaving full-time education and entering the employment field at a time when employment prospects for the young generally are reduced.

(ii) the shorter average duration of minority group workers makes them more vulnerable to the "last in - first out" rules in times of redundancy.

(iii) a higher degree of unemployment in those industries and areas in which minority groups work and live, eg. the South East and West Midlands.

(iv) a lack of appropriate skills and qualifications of recent immigrants.

(v) discrimination by individual employers which means that in times of plentiful labour supply minority groups will suffer more in terms of job acceptances.

Although the other factors should not be discounted, the evidence to date suggests that discrimination against minority groups is in fact a major problem.

2.1.2 Research on Racial Discrimination

Research studies in the last 15 to 20 years have consistently shown the existence of discrimination against minority groups in selection, placement, promotions, etc.

One of the most important and wide-ranging studies was conducted by Smith (1976). Using census data, Department of Employment figures on the number of people unemployed (in particular people from minority groups) and data from his own surveys he found

that unemployment amongst young West Indians was much higher than in the population generally; when total unemployment rises the increase is much greater amongst minority groups; black workers with academic qualifications hold lower level jobs than whites with comparable qualifications, eg. 55% of black people with 'A' levels managed to obtain non-manual jobs compared with 83% of white people.

Daniel (1968) looked in more detail at the extent of discrimination in certain situations. In carefully controlled studies he had black and white actors present themselves at factory gates enquiring after unskilled and semi-skilled jobs. The actors had been matched for age, sex, qualifications, motivation and presentability. An act of discrimination was taken to be any occasion when one candidate was treated differently to the other eg. where the white candidate was offered an interview but the black candidate was refused - or vice versa. In such cases black people were discriminated against in 48% of cases.

Smith (1976) used correspondence tests to examine the extent of discrimination for white collar jobs. Letters were sent out for a sample of advertised jobs. Two applications were made for each job - one in the name of a white British applicant and the other in the name of an immigrant. An act of discrimination, again, was any occasion when one candidate was invited to interview and the other was not. The levels of discrimination against the three minority groups in the study were very similar - West Indians 33%, Indians 27% and Pakistanis 30% (these are net figures, which allow for cases where minority candidates were

favoured). Similar results have been found by other researchers, eg. Hubbock and Carter (1980).

The advantage of the methodology adopted by these two studies is that all aspects of the presentation and application were controlled.

The drawback is that they only test for discrimination at the first screening stage of recruitment - which means that only minimum estimates of the level of discrimination are produced.

It could also be argued, however, that black immigrants occupy low status jobs because of their lack of knowledge of the English language, British conventions, etc. Ballard and Holden (1975), therefore, looked at the employment of a group they termed 'Black Britons' ie. black people who were born and brought up in Britain. They attempted to discover how a group of black university students fared in applying for their first job when compared with a matched group of white students.

Their survey revealed that:

(i) length of residence appeared to make little difference to the chance of getting employment

(ii) black students were far less successful than their white counterparts at all stages of job finding.

They concluded that in such situations it is the suitably qualified white candidates who are considered first while black candidates wait at the back of the queue; where there is an excess of suitable applicants for the job, coloured graduates may

never reach the front. "The career prospects of coloured job seekers, however well-qualified, look bleak, and will remain so without a radical change of attitudes on the part of employers." (p. 175). It should be borne in mind that their research was carried out in the early 1970s when unemployment was much lower and graduate opportunities were rather greater.

Overall, a substantial amount of research has been accumulated which indicates that black people are concentrated in low status jobs; that their unemployment rate increases proportionately faster than the majority group during times of recession; that they face discrimination at all levels of employment; that the extent of discrimination is considerable even amongst black people born and educated in Britain; and that black graduates find it more difficult to obtain employment than their white counterparts.

2.2 Government Responses - The Race Relations Acts of 1968 and 1976

Research studies, such as those quoted above, indicated with some force that racial discrimination was not a problem which would disappear merely through the passage of time and the changing of attitudes. In America it was recognised that legislation, forceful legislation, was needed to overcome disadvantage and discrimination. In the UK this realisation led to the 1968 Race Relations Act, but its scope was very narrow, dealing only with overt or direct discrimination. With the 1976 Race Relations Act (which can be seen as an admission that the previous Act was too weak), it was recognised that direct discrimination is not

the only form of discrimination which needs to be legislated against, and so it contained the concept of indirect discrimination - a topic central to this project.

2.2.1 Direct Discrimination

The evidence from research studies had shown that discrimination was widespread, and it became clear that in order to overcome it, some form of legislation was required to provide the impetus for change. The case for legislation is laid out clearly in the "Government White Paper on Racial Discrimination" (1975). "Legislation is the essential precondition for an effective policy to combat the problems experienced by the coloured minority groups and to promote equality of opportunity and treatment. It is a necessary precondition for dealing with explicitly discriminatory actions or accumulated disadvantages. Where unfair discrimination is involved the necessity of a legal remedy is now generally accepted. To fail to provide a remedy against an injustice strikes at the rule of the law. To abandon a whole group of people in society without legal redress against unfair discrimination is to leave them with no option but to find their own redress. It is no longer necessary to recite the immense damage, material as well as moral, which ensues when a minority loses faith in the capacity of social institutions to be impartial and fair" (p.5/6).

The 1968 Race Relations Act was an attempt to provide minority groups with some legal redress against cases of overt or direct discrimination. Discrimination was defined as follows:

"A person discriminates against another if, on the grounds of colour, race, sex or ethnic or national origins, he treats that other...less favourably than he treats or should treat other persons" (s.1(1)).

The term "on the grounds of" indicates that the "less favourable treatment" must be intentional, and is based on the 'discriminator's' belief about the race or colour of the person before him/her. As Grosz and Bindman (1979) explain "discrimination was defined in terms of the 'attitude' and intention of the discriminator and not in terms of its effect" (p.21).

The Act, therefore, saw discrimination as a direct wrong done by one individual to another.

2.2.2 Indirect Discrimination

In the United States of America, civil cases brought by individuals against companies had shown that direct discrimination was not the only form discrimination could take. People could be discriminated against unwittingly due to the institutional practices and procedures which an organisation may follow, eg. a condition may have to be fulfilled before a person can be considered for a job but the nature of this condition or requirement is such that, although applied equally to all applicants, certain groups will be less able to fulfill it. This is known as indirect discrimination, and in the USA this was legally recognised in the Supreme Court decision in the case of Grigg v Duke Power Company. In this case it was pronounced that the Civil Rights Act of 1964 "prescribes not only overt discrimination, but practices that are fair in form but discriminatory in operation.

The touchstone is necessity. If an employment practice which operates to exclude negroes cannot be shown to be related to job performance the practice is prohibited" (401 U.S. 424 (1971)).

Continued discrimination and the American experience highlighted the weaknesses of the 1968 Act, which was accepted by the Government in its White Paper in 1975. Whilst acknowledging that the 1968 Act had made cases of overt and crude discrimination less common, it was also asserted that "...while it is right that motive should be relevant in determining whether an illegal discriminator should compensate his victim, it is insufficient for the law to deal only with overt discrimination. It should also prohibit practices which are fair in a formal sense, but discriminatory in their operation and effect." (p. 8).

To overcome these inadequacies, the 1976 Race Relations Act widened its definition to include indirect as well as direct discrimination.

In the 1976 Act, direct discrimination is dealt with in Section 1 subsection (1) paragraph (a), and it is defined as a person who when dealing with somebody else "on racial grounds treats that other person less favourably than he would treat other persons". This definition is comparable to that given in the 1968 Act.

S1(1)(b), on the other hand, concerns indirect discrimination, and it says it is when a person "applies to that other a condition or requirement which he applies equally to persons not of the same racial group as that other but:

(i) which is such that the proportions of persons of the same racial group as that other who can comply with it is considerably smaller than the proportion of persons not of that racial group who can comply with it, and

(ii) which he cannot show to be justifiable irrespective of the colour, race, nationality or ethnic or national origins of the person to whom it applied, and

(iii) which is to the detriment of that other because he cannot comply with it."

The Act also includes as discrimination any form of victimisation or segregation, and applies to discrimination in education, housing, the provision of goods, services and facilities as well as in employment.

Although three conditions have to be satisfied before an employer can be said to be guilty of indirectly discriminating, virtually any case will be resolved around the meaning of the word 'justifiable'. Unfortunately, the Act gives no guidance as to what is meant by 'justifiable' which has led to the concept being interpreted by the Industrial Tribunals. Following the American example, justifiability has become associated with necessity, ie. a condition is justifiable if it is considered to be necessary for a business to operate and is not merely convenient or desirable.

No test case of indirect discrimination under the 1976 Race Relations Act has yet passed through the legal system, and so an example will be given of a case brought under the Sex Discrimination Act (1975), which also legislates against indirect discrimi-

nation, and which made reference to the notion of necessity (Steel v Post Office 1977 IRLR 288 (EAT)).

In this case, the Post Office had a rule that where postal workers were competing for favoured routes, it would be given to the most senior person in terms of years served - but only full-time years of service were counted. However, until 1975 women had only been allowed to work part-time; and so when Mrs Steel came to compete for a particular route, it was given to a man who had been employed at the Post Office for a shorter period because her part-time years of employment, prior to 1975, were not taken into account when calculating seniority. It was this condition that was seen as being indirectly discriminatory, since a considerably smaller proportion of women than men could comply with this rule. The Industrial Tribunal supported Steel's claim and she succeeded in her action.

Discrimination, as Pearn (1979) points out, is nearly always the result of a decision, and the basic premise of indirect discrimination is that a lot of decisions being made cannot be justified because there is not enough information about the decision-making process to see whether the decision was justifiable or not. The important point of this, though, is that whenever the numbers of ethnic minority members or women in any occupation, at any level of seniority, are substantially lower than for the majority group, then the question of indirect discrimination can and should be raised. If the disproportion in numbers, or adverse impact, is due to an act, policy or requirement which cannot be shown to be necessary, then that employer has broken the law. In the United States, this could lead to the imposition of heavy fines and

payment of damages to those who suffered 'detriment' from those actions.

2.3 Implications of the Legislation for Selectors: The Commission For Racial Equality's Code of Practice

The Act, and in particular those sections relating to indirect discrimination, should have had a considerable impact on the work of professional selectors, including occupational psychologists. The Act places the onus on employers and/or selectors to demonstrate that no psychological test or assessment procedure has an adverse impact on minority group applicants. If adverse impact cannot be denied then they must show that the particular selection requirement is predictive of a necessary feature of job performance.

In the light of this, Wallis (1979) recommends that selectors must ask themselves two questions. Firstly, has the selection requirement been set because of some demonstrably necessary feature of job performance, and secondly does the selection instrument itself have a demonstrably relevant relationship with job performance?

More importantly, the Act set up the Commission for Racial Equality (CRE) with the duties of working towards the elimination of discrimination and promoting equality of opportunity and good relations between persons of different racial groups. In April 1984, the CRE 'Code of Practice' came into effect and it states the responsibilities of employers, employees, trades unions and employment agencies in the promotion of equality. Since the rules are in a Code they are not legally binding and so a breach of

code will not be a breach of the civil or criminal law. However, as Jenner (1983) points out in a paper addressed to personnel managers "a code is no mere guidance note. Failure to comply may, and undoubtedly will be used in evidence against you in any proceedings civil or criminal, to which it is relevant" (p.25).

The Code states that "Responsibility for providing equal opportunity for all job applicants and employees rests primarily with employers. To this end it is recommended that they should adopt, implement and monitor an equal opportunity policy to ensure that there is no unlawful discrimination and that equal opportunity is genuinely available" (p 8).

The aim of an equal opportunity policy is to ensure that no job applicant receives less favourable treatment on racial grounds, that no-one is discriminated against; and that, in suitable cases, employees of under-represented racial groups, through positive action programmes, should be given training and encouragement to achieve equality within the organisation. To ensure its effectiveness the CRE recommends: "making an initial analysis of the workforce and regularly monitoring the application of the policy with the aid of analyses of the ethnic origins of the workforce and of job applicants" (p 9).

This recommendation of all the Code's recommendations is the one generally considered most controversial (IRLIB February 1984) and by the CRE as the most misunderstood.

Significantly, the Code calls not for 'ethnic monitoring' but for 'the monitoring of equal opportunity' ie. the effects of selection

decisions and personnel practices and procedures. The best way to obtain information for effective monitoring is through "records showing the ethnic origins of existing employees and job applicants". (p 18).

No method of monitoring is recommended (although the CRE has its own classification of minority groups). Instead the Code states that "it is open to employers to adopt the method of monitoring which is best suited to their need and circumstances, but whichever method is adopted, they should be able to show that it is effective" (p 18). For example, in a small firm it might be adequate to assess the distribution of employees from personal knowledge and visual identification. A more comprehensive analysis would cover

"(a) the ethnic composition of the workforce of each plant, department, section, shift and job category, and changes in distribution over periods of time;

(b) selection decisions for recruitment, promotion, transfer and training, according to the racial group of candidates, and reasons for these decisions"

(p 18/19).

The information should be carefully analysed to identify areas which may need attention and a number of key questions should be asked which are : "Is there evidence that individuals from any particular racial group:

(a) do not apply for employment or promotion or fewer might apply than might be expected?

(b) are not recruited or promoted at all, or are appointed in a significantly lower proportion than their rate of application?

(c) are under-represented in training or in jobs carrying higher pay, status or authority?

(d) are concentrated in certain shifts, sections or departments?" (p.19)

If the answer to any of these questions is 'yes' then "the reasons for this should be investigated. If direct or indirect discrimination is found, action must be taken to end it immediately" (p 19).

The Code also provides guidelines for developing selection criteria and the use of tests. "In order to avoid direct or indirect discrimination it is recommended that selection criteria and tests are examined to ensure that they are related to job requirements and are not unlawfully discriminatory" (p 11). In particular "selection tests which contain irrelevant questions or exercises on matters which may be unfair to racial minority applicants should not be used (for example, general knowledge questions on matters more likely to be familiar to indigenous applicants)" and also that "selection tests should be checked to ensure that they are related to the job's requirements, ie. an individual's test markings should measure ability to do or train for the job in question" (p 12).

The Code contains the CRE's strategy for implementing an equal opportunity programme. Other general intervention strategies have been devised, eg. Pearn (1978), Wainwright (1980), Hitner et al (1981). Obviously, these approaches differ from one another but the authors all agree that if equal opportunity is to be taken

seriously then organisations must conduct some form of audit or self-analysis of the origins of job applicants and employees.

Neither the Code, the CRE nor any of the writers in this area provide any specific advice to organisations on how to construct or find the most suitable ethnic classification for their particular circumstances. There are many problems which have to be resolved before an 'effective' method of monitoring can be found and these are explored in Chapters 5 and 6.

2.4 Equal Opportunities within the Civil Service

This project is particularly concerned with equality of opportunity within the Civil Service and especially in determining whether the Executive Officer (EO) qualifying test battery has any adverse impact against minority groups. It is appropriate, therefore, to delineate the results of monitoring studies carried out within the Civil Service (in particular a study by the Tavistock Institute), to outline the work of and selection procedure for EOs and to report the results of a validation study on the EO qualifying test battery.

2.4.1 Monitoring Studies within the Civil Service

Perhaps one of the most important studies into equal opportunities and monitoring to be carried out in the United Kingdom to date, was the Tavistock Institute's examination of the effectiveness of the Civil Service's equal opportunity policy.

In 1975, the Civil Service produced a statement which seemingly outlined their attitude toward equal opportunities. It stated that "It is a statutory requirement that a person must satisfy the nationality qualifications to be eligible for appointment to a post in the Civil Service. It is the policy of the Civil Service that all persons so eligible shall have equal opportunity for employment and advancement in the Civil Service on their ability, qualifications and fitness for the work. There shall be no discrimination against any person eligible under the nationality rule whether in recruitment, promotion or in any other way, on the grounds of colour, race or ethnic or national origin. This applies to all departments and to all grades and positions in both the industrial and the non-industrial Civil Service".

The Civil Service employs about 750,000 people divided into 167,000 industrial staff (manual grades working in H.M. dockyards, munitions factories, etc) and 566,000 non-industrial staff (clerical, professional and managerial grades).

Since the government, through the Civil Service, is the largest single employer in the country, it was felt, both by Government ministers and those working in the Civil Service Department (the department formerly responsible for personnel matters) that there was a need for a detailed and objective study into the recruitment, training and promotion procedures in order to discover whether they were in any way operating in a discriminatory manner, and to the disadvantage of ethnic minority groups.

On this basis, the Tavistock Institute of Human Relations were asked to undertake a study "to explore the possibility of developing a system of monitoring equal opportunities in the Civil Service, acceptable to all interested parties which would clearly show whether there are equal opportunities for various racial groups" (HMSO 1978). Between 1976 and 1977 seven separate studies were conducted on recruitment and employment procedures at various Civil Service sites.

Before starting their fieldwork, however, a method for classifying ethnic minority groups had to be devised. On the Civil Service application forms no direct references are made relating to ethnic or racial origins. The researchers, however, decided to infer such information by looking at other details actually asked for on the application forms, eg. name, birthplace, nationality of the applicant and his/her parents, and countries where they had been educated or resided. According to this information the person would be placed into one of two categories - white or coloured.

The differences found between the white and coloured groups were quite marked on a number of criteria. For example, in the Examiners in Insolvency competition, 19% of the eligible white group were offered posts compared with 5% of the eligible minority group. For an Executive Officer competition the success rates for the white and coloured groups were 20% and 8% respectively. At the DHSS Headquarters the white group's chances of success were twice as good as the coloured group's - success rates being 79% and 40% respectively. In addition it was found that of candidates who were suitably qualified to hold a partic-

ular post, the coloured candidates were more likely to be offered a post at a lower level than were white candidates. Similarly of a sample of 96 (46 coloured and 50 white) people already working within the Civil Service, it was found that 30 possessed qualifications which were additional to those required for entry to the job. Of these 22 were coloured.

These results emerged even though the method of classification was relatively crude. The method which was adopted may be unreliable because within certain minority groups, especially West Indians, who were born in the UK, some have English-sounding names and British nationality. Consequently it would be easy to place them in the wrong category. The effect of this would be to increase the rejection rate of the white and reduce it for the coloured groups, thus making the differences between them less than they may be in reality.

What this study does reveal though is that even a crude classification of this sort can in fact produce interesting results. A more sophisticated method would hopefully provide a more detailed breakdown of the way individual minority groups are treated.

It should be borne in mind, however, that interpretation of the figures is perhaps the most important and the most difficult part of any sort of monitoring. The figures which result from such analysis will have to be seen in context since a lot will depend upon local circumstances, eg. the size of the local population, the distribution of minority groups in that area, etc.

The study attracted considerable interest, but despite that the Government decided not to implement a system of ethnic monitoring for the Civil Service. This decision was strongly attacked, since it was, and still is, widely felt that the Civil Service, because of its size and unique position within the United Kingdom, has an obligation to set an example as a model equal opportunity employer. In June 1981 the issues were debated both in the House of Commons and the House of Lords which, taken together with the pressure from other organisations (including the Runnymede Trust and the Runnymede Industrial Unit), indicates the concern which was being felt.

Also in 1981 the Home Affairs Committee on Race Relations' produced their report which recommended that the Government reconsider its decision on ethnic monitoring, and Lord Scarman published his report on the riots of the summer of that year. In a House of Commons debate on Scarman, Home Secretary William Whitelaw said "To help in the attack on racial disadvantage we must place greater emphasis on what is generally known as ethnic monitoring. We accept the need for the Government to give a lead in this area." It was decided to conduct an experimental monitoring exercise in Leeds.

The survey was carried out in two sections, first a voluntary self-classification survey of existing staff, and secondly voluntary self-classification by job applicants. The classifications in both cases were White, Black (West Indian, African, Other), Asian (Indian, Pakistan, Bangladeshi, East African, Other) and Other. Job applicants were also asked to complete a second

questionnaire which contained 15 sub-categories such as Nigerian, Polish, Chinese.

Results of the survey of staff showed that 2.2% were to be of ethnic minority origin. The largest number of ethnic minorities were in typing grades (3.1%). Of 372 staff in middle management two came from ethnic minority groups, and in the grades requiring no qualifications there was only one of 233.

The second survey revealed that of 965 applicants at the short-listing stage, 48 (5%) were of ethnic minority origin. Of those, 11 were ineligible (ie. not meeting educational or other requirements). 26 of the remaining 37 applicants were rejected (4% of all rejections). Of the 10 interviewed, three were declared successful (3% of all successes) and two were offered appointments (2.5% of all offers).

Following the Leeds experiment the Civil Service has started to carry out ethnic monitoring in other regions.

The two projects quoted above appear to highlight the need for effective monitoring of equal opportunity, but there has in fact been very little research examining reliable, accurate and acceptable ways of classifying people.

Secondly, it appeared from the Tavistock report that the Civil Service selection examinations were proving to be a difficult hurdle for many minority groups, and so could possibly be indirectly discriminatory.

These two issues in fact form the major focus of the project, and the Civil Service tests chosen to be examined were those for entry into the Executive Officer grade.

2.5 Executive Officers' Work and Qualifying Tests

2.5.1 The Work of Executive Officers

The Executive Officer (EO) grade is an administrative one, but there are a wide variety of activities which they might find themselves doing, including the supervision of clerical staff, assessing senior staff, dealing with public within the office or outside, etc. There are also more specialised areas such as Examiners in the Insolvency Service of the Department of Trade; Immigration Officer in the Home Office; Collector of Taxes and Tax Officers in the Inland Revenue.

After four years' service EOs can be considered for promotion to Higher Executive Officer (HEO) and to Senior Executive Officer (SEO) after three years' service as an HEO. For the grades above SEO there are no laid down minimum periods of service. The EO selection procedure is a hurdle which must be overcome, therefore, for anyone who wishes to progress to the highest levels of the Civil Service.

Eligibility

There are a number of conditions which candidates must satisfy if they are to be considered eligible for EO posts. These are:

(i) Age - candidates must be at least 17½ and under 45 years of age.

(ii) Nationality - candidates must be, and must have been from birth, a British subject and also one of their parents must be, or was at death, a British subject. For certain posts, eg. in the Ministry of Defence, both parents must have been British subjects at all times since birth.

(iii) Educational requirements - all candidates must have one of the following: 5 GCE 'O' levels (or equivalent) which must include English, and two 'A' levels; an Ordinary National Certificate or Diploma (or equivalent); Higher National Certificate or Diploma; Business Education Council (BEC) National Certificate or Diploma with either four acceptable passes at 'O' level, or a Business Education Council General Diploma at credit level; BEC Higher National Certificate or Diploma; a university or CNAA; a pass in an examination held by a professional body which is accepted by the commissioners as being of a satisfactory standard.

In addition, candidates must achieve a qualifying mark in the EO written examinations.

2.5.2 The EO Selection Procedure

On applying to the Civil Service each candidate is sent details of the job, brochure, etc, together with an information sheet. Every candidate who returns the information sheet is invited to sit the qualifying test. (This procedure was introduced after the Tavistock report and is intended to eliminate the discrimination which occurred during the screening of application forms.) No candidate can be exempted from taking the tests or be allowed to carry forward a performance on an earlier occasion. Each candi-

date is also sent a pamphlet which contains examples of the types of questions they will be faced within the tests.

The tests are computer scored, and a check made of the information sheets to determine whether those people who have passed the tests also meet the other entry requirements. If they are not eligible they are not invited to be interviewed, even though they may have passed the test.

The interview forms the final part of the selection procedure.

2.5.3 The Qualifying Test Battery and Its Validity

The EO qualifying test battery consists of three tests:

(i) Enquiries from management - a multiple choice test of numeracy which lasts 25 minutes in which candidates must interpret statistical information presented in the form of graphs and tables, etc.

(ii) Intelligence - this test consists of three separately timed tests each lasting 15 minutes and intended to measure "intelligence, accuracy and commonsense". The first part contains 17 problems which vary in nature, and candidates must work out solutions from the information given. The second part contains 30 sentences in which two of the words need to be interchanged in order to make the sentence read sensibly. The third part comprises 27 questions, each of which consists of a design or part of a design covering a number of squares. In one of the squares the design is incorrect and must be identified.

(iii) Executive Problems Test - this lasts for one and a half hours and consists of 20 short answer problems "of the kind that can arise in administration".

Examples of each of the tests are provided in Appendix 1.

The purpose of any sort of pre-selection test is to produce information about the probability of an applicant's success ie. it should have some validity. In order to determine the effectiveness of the EO qualifying test battery the Civil Service conducted a predictive validation study where EOs taken in the 12 month period between June 1971 and July 1972 were followed up. 90% of the EOs go into ten departments with many of them going into general administrative posts and a few entering specialist areas. The validation study concentrated on the general administrative positions.

Annual Confidential Reports (ACRs) were used as the job performance measure. At the time of the study new ACRs were being introduced so data was obtained for both the old and new forms for the first two years of a person's service. It was felt that "although the first Annual Report was unlikely to reflect the EOs abilities very accurately, it was expected to reflect his ability to adapt to the new work situation. The Annual Report completed at the end of the second year's service was felt to be a more reliable indicator of levels of job performance". (Paul and Kuplens 1977 p.3.)

The ratings between some of the ACR criteria and the 'Intelligence' test are shown in Table 2.3. For the first year it correlated significantly with three of the criteria - foresight, judgement and ability to produce constructive ideas - all of which assess cognitive aspects of work performance. Correlations with criteria dealing with the quality or type of work produced were not given. Positive correlations with overall performance were very low $-^{+}.06$ for the first year and $^{+}.09$ for the second. The second year correlation was in fact significant ($p < .05$) but this is probably due to the large sample.

For those people who were rated on the same form for the first and second years there are no significant correlations between the tests and second year performance. Indeed, the correlation with overall performance is negative.

The second test, called the 'Graphs and Tables' test, appears similar to the 'Enquiries from Management' test used in the 1982 selection testing. The correlation between overall performance in the second year and test score was stated as significant but neither the correlation coefficient nor the level of significance were provided. The 'Graphs and Tables' test produced significant correlations with numerical ability $-^{+}.16$ ($p < .05$) for the first year, and $^{+}.09$ ($p < .05$) for the second. Despite their significance, the correlations were nevertheless low. In posts where numerical ability was of high importance the correlation between the test score and numerical ability was $^{+}.21$ ($p < .05$). For posts where numerical ability was of low importance, however, the correlation was only $^{+}.05$ which is not significant.

Table 2.3 Correlation Coefficients Between the Intelligence Test and the Annual Reports

INTELLIGENCE TEST SCORE	ACRs			OLD FORMS	
	1st Year	2nd Year	2nd Year[1]	1st Year	2nd Year
IT x Overall performance	.06 (n=277)	.09* (n=592)	-.01 (n=162)	.05 (n=377)	not enough cases
IT x Foresight	.18** (n=227)	.05 (n=598)	.02 (n=163)	not enough cases	.04 (n=88)
IT x Penetration	.13 (n=227)	.10* (n=598)	.02 (n=165)	.13** (n=474)	.16* (n=219)
IT x Judgement	.16** (n=227)	.04 (n=598)	.01 (n=165)	.06 (n=843)	.04 (n=254)
IT x Ability to Produce Constructive Ideas	.13* (n=227)	.05 (n=595)	.02 (n=163)	.10 (n=221)	.09 (n=68)

** $p < .01$
* $p < .05$

1 This sample consists of those individuals for whom both sets of reports were in the new-style format and refers to coefficients produced on their second year reports. (Source: Paul and Kuplens 1977)

The study then shows that the 'Intelligence' test is useful for predicting some criteria - all related to cognitive aspects of work performance - and that the 'Graphs and Tables' test is particularly useful for predicting performance in jobs where the numerical content is of high importance, but is less useful for jobs where it is of low importance. In all cases, however, the significant correlations with job performance are low (sometimes below .1). Overall, the results of this study can be said to be mixed, and certainly do not provide incontrovertible evidence as to the tests' validity.

Furthermore, the relevance of this validation study to the present situation is uncertain. The study was initiated in 1973, and in the intervening 12 years the posts of Executive Officer may have changed drastically. Also, it would appear that the selection testing was altered, since the 1971/72 candidates did not take the Executive Problems test and consequently there is no validation evidence for it.

2.6 Objectives of This Project

This project is particularly concerned with ethnic monitoring within the Civil Service, determining whether the Executive Officer qualifying test battery has any adverse impact against minority groups and, if so, attempting to discover the reasons for it.

After a long period of discussions with the Civil Service, in particular with the psychologists of the Recruitment Research

Unit, a project was eventually agreed upon. Figure 2.1 presents the terms of reference which were finally agreed at the start of the project.

Figure 2.1 Outline of Feasibility Study

I. Data Collection

Stage A	Pilot on non-civil servants.
Goal	To test several self-classification schemes (including that suggested by the Commission for Racial Equality).
Method	A sample of secondary school students who fulfil the eligibility criteria for the Executive Officer Competition would be asked to complete one or more classifications while an observer (or the teacher) provided ratings of racial identification. To guarantee confidentiality, the data would be rendered anonymous in the students' presence.
Stage B	Pilot on live candidates.
Goal	To test the better versions of monitoring systems from Stage A.
Method	The Executive Officer Competition allows for data-gathering on larger samples (than in Stage A) at a point remote in time from the selection decision. Groups of candidates are

invited to sit a 3-hour qualifying test on the basis of which those who pass are subsequently invited to a Selection Board. The data-gathering with self-classification and observer ratings* would be followed by the return to candidates of any information that might identify them as individuals at a later stage in the selection process. Analysis would be confined to the monitoring system and the test scores, in ignorance of the sample's subsequent experience in those later stages.

Subject to the findings of the above stages, further research (requiring re-negotiation of the contract) might be envisaged. Subsequent stages would consider the indirect discrimination at the selection decision point in the Executive Officer Competition and that of the employment process of serving Executive Officers.

II. Projected results of such a study

The results would include:

(a) The evaluation of several monitoring systems including that advocated by the Commission for Racial Equality.

(b) Evidence on samples of ethnic/racial minorities in one of the CS Commission's largest candidate fields.

(c) Information on the response rates and "spoilt" data with the different monitoring systems tested.

* It was originally envisaged that not only would EO candidates classify themselves, but also that test administrators should attempt to classify them. This proved impractical to set up, however, and so was not carried out.

CHAPTER 3

Psychological Testing and Issues of Bias and Fairness

As part of this project will be to examine the effects in terms of pass/fail rates that the Executive Officer Qualifying Test Battery has on ethnic minority groups it would be useful here to outline some of the current areas of concern with regard to test use. There are many different definitions of psychological tests. By examining them a comprehensive definition is provided for use in this project. There is considerable literature on psychological tests and problems in their use. By adopting a historical perspective it can be seen that the major principles of testing have been in existence for many years. It is partially a failure in adhering to these principles that the problems concerning bias and unfairness arise. However, it is also possible to identify the major sources of bias and provide solutions to overcome some of them. Two other areas considered here are differential validity and models of selection fairness. Despite the attention these subjects have commanded they appear to have revealed more about the biases of the researchers than of the tests being investigated. As a result many issues still remain unresolved.

3.1 Definitions of Psychological Tests

There are many definitions of a psychological test, each of which takes a slightly different view. Taking these together it is possible to arrive at a definition which incorporates the salient points.

The Institute of Personnel Management (IPM) define psychological tests as 'standardised sets of questions or problems which allow a candidate's performance to be compared with that of other people of a similar background' (Sneath et al 1976 p.1). The emphasis is on the usefulness of tests in providing comparisons of people through the use of appropriate norm tables.

Anastasi (1975) sees the basic function of tests as being to 'measure differences between individuals or between the reactions of the same individual on different occasions' (p.3). Again, the use of tests for comparison is emphasised, but also the reliability of tests is implied, ie. if a person's score changes on the second time of taking a test it is due to some change within the person rather than within the test.

Following the Civil Rights legislation in America, the Office of Federal Contract Compliance (OFCC) extended the definition to include many other aspects not formerly called tests. "All other formal, scored, quantified or standardised techniques of assessing job suitability including, for example, personal history and background requirements which are specifically used as a basis for qualifying or disqualifying applicants or employers' specific educational or work history requirements, scored interviews, biographical information blanks, interviewers' rating forms and scored application forms" (OFCC 1971). This wide-ranging definition of a psychological test has gone too far. Tests can be differentiated from procedure such as interviewers' assessment forms, scored application forms, etc. Tests ask the candidate for a response to a number of problems or questions and furthermore these questions remain the same, so that if a person were to take

a test again they would be faced with precisely the same items. Nor is the information biographical - test scores should provide information which biographical information alone could not provide.

Bearing these points in mind the definition of a test used for selection purposes in this project is a standardised set of questions or problems, completed by the person being assessed, which demonstrably measures specific skills and abilities, and which allows comparison between individuals from a group of similar background or between the same individual on two or more separate occasions.

3.2 History of the Development of Selection Tests

The beginnings of selection testing are often seen as being the turn of this century. Anastasi (1975), however, quotes research which reveals that the Chinese Empire had a system of Civil Service examinations 3,000 years ago and, furthermore, that the ancient Greeks considered testing an established part of the selection process.

The history of modern testing, however, really starts in the mid-19th century with developments occurring almost simultaneously in a number of countries. In France, Esquirol and Seguin developed tests for identifying and classifying the mentally retarded. In Britain, Galton and Spearman were pioneering the use of questionnaires, rating scales and applying mathematical principles to tests.

In 1905 the Binet-Simon scale appeared, again primarily intended for identifying the mentally retarded, and it was possibly the first test to gain worldwide attention. The American adaptation of this test - the Stanford-Binet test - introduced the term Intelligence Quotient (IQ) which describes the relationship between mental age and chronological age.

The greatest strides in test development took place during the First World War. There was an urgent need to devise an instrument which could quickly and effectively place people into particular posts and grades in the United States Army. The work eventually led to the Alpha and Beta tests which were used successfully and represented considerable advances in test construction.

The success of these tests led to a dramatic growth not only in test research and development but also in their use in personnel selection. So much so, in fact, that by the 1950s certain tried and tested principles had entered into American personnel practice. These were summarised by Guion (1976) into what he termed "The Tenets of Orthodoxy", which state that:

(i) the purpose of personnel selection testing is to predict job performance

(ii) predictors and criteria should be selected on the basis of job analysis

(iii) measuring instruments must be standardised

(iv) tests should be empirically validated

(v) validation is situation specific

(vi) more than one test should be used

(vii) only one criterion should be used

(viii) tests are preferred over 'non-test' predictors

(ix) individual differences should be recognised in evaluating tests

(x) tests are supplements to existing employment procedures

These tenets, which remain sound principles of selection even today, became accepted and, with the successes which tests were achieving, attitudes towards testing changed. The tenets were there for everyone to follow and as a result the interests of industrial psychologists moved away from test construction and validation towards other areas such as attitudes, leadership, motivation, etc. Non-psychologists were now using tests and the tenets were not so rigidly applied allied with a lack of awareness of the consequences of what was being done.

Interest revived though, with the passing of the Civil Rights Act in 1964. The case of Myart vs Motorola led to a clause being included in the act stating that "... it shall not be an unlawful practice for an employer... to give and act upon the results of any professionally developed ability test, provided that such test, its administration or action upon the results is not designed, intended or used to discriminate because of race, colour, religion, sex or national origin" (Section 703(h)). The key phrase is "or used" because it makes it clear that a test that inadvertently discriminates is just as illegal as one that does so deliberately. This is the concept of indirect discrimination which is recognised in the Sex Discrimination and Race Relations Acts of the United Kingdom.

3.3 Testing and Minority Groups

A very controversial area in test use remains that of bias, and whether minority groups perform worse on particular types of test. Despite the raging debates some trends in the data seem to be emerging. Minority groups appear to do less well on tests of verbal ability and verbal reasoning, and that the results are worse for candidates of working class background and for those who are bilingual.

In a study concentrated on primary schools in inner London, Little et al (1968) administered the Wechster Intelligence Scale for Children (WISC) to groups of indigenous and immigrant pupils. They found that on the verbal component of WISC 82% of immigrant pupils scored well below average.

Eels et al (1951) found similar results when comparing 5,000 children from low and high social status groups. The low status groups performed particularly worse on 'intelligence' test items, particularly the vocabulary and verbal ones.

In one of the most well-known studies in this area, Jensen (1971) found differences in overall levels of ability and patterns of abilities between three ethnic groups - negro, white and Mexican. As a result he hypothesised two broad hierarchical levels of ability: Level 1 - Associative Ability depends largely on digit memory, rote learning and paired associate learning, and Level II - Conceptual Ability involving abstract reasoning processes such as concept learning and problem-solving. He found no differences in Level I ability, but children of low socio-economic status, especially ethnic minority children, were approxi-

mately one standard deviation below the general population mean for tests of Level II ability. Jensen claims these differences are due to genetic rather than environmental factors, and this is a debate which remains active today.

Vernon (1965 (a) and (b)), however, compared 50 West Indian boys with 100 English children on a series of tests and assessments and concluded that although West Indians did poorly overall, the deficit varied for different tests and was most notable for practical and non-verbal tests of intelligence rather than verbal ones.

This is supported by the results of a large scale study by Coleman et al (1966). This was the largest study of the United States' school testing programme ever carried out - 650,000 children were tested on a number of aptitude and achievement tests. Numerous academic criteria were used and it was found that the verbal IQ measures showed less predictive bias than the non-verbal ones. These results are supported by Hegarty and Lucas (1978) who argue that verbal and number tests are, in general, better predictors of educational achievement than non-verbal tests - although they strongly criticise the view that verbal reasoning tests are also equally good predictors.

Differences have been found, however, between the results of immigrant and non-immigrant minority group children. McFie and Thomson (1970) for example found differences between matched samples of West Indian and English children on the verbal component of WISC. The West Indian children performed poorly compared with the English group. However, those West Indian

children who arrived in Britain before the age of 5 years performed better than those who arrived after that age. In a study which supplements the above findings Watson (1973) found that performance of immigrant children on tests increases relatively with the length of schooling.

Watson's study was small-scale but his findings were supported by later researchers eg. Yule et al (1975) conducted two studies - one involving a sample of 2,281 children of whom 354 were West Indian and found that West Indian children born in the United Kingdom scored significantly higher on educational tests than those born in the West Indies. However, the British born West Indian group's scores were still well below the national average.

In reviewing the literature on the performance of minority groups on verbal and non-verbal tests Jensen (1980) concludes that "... in large representative samples of American blacks there is probably little, if any, difference in level of performance on verbal and non-verbal tests" (p.529) but goes on to state that "... some verbal tests show possible bias for Mexican-Americans that is not demonstrable in the black samples" (p.587). It appears, therefore, that verbal tests could well be biased against at least some minority groups and that verbal reasoning tests might possibly show more bias.

It seems, however, that bilingual groups perform worse on these types of tests than monoglot groups. Alleyne (1962) gave a variety of verbal and non-verbal tests to groups of children and found that the monoglot groups were constantly superior to bilinguals and that the degree of bilingualism had a progres-

sively adverse effect, but that this occurred to a lesser extent on the non-verbal tests. Those children who were moderately or highly bilingual and who had been born outside the UK performed particularly badly.

These findings are supported by Jensen (1980) who found that all non-English speaking or bilingual groups examined in America "... obtained higher scores on every kind of non-verbal and performance test than on any kind of verbal test presented in English, oral or written. It seems most unlikely that the general findings would be entirely the result of an inherent deficiency in these groups' verbal ability relative to their non-verbal ability" (p.605) (own emphasis). Furthermore, he found that "On standard (English language) tests of scholastic achievement, pupils from foreign language backgrounds perform relatively better on arithmetic items than on language items" (p.605). Even though these conclusions were reached on reviewing American studies he feels that they are generalisable to groups in Britain such as Asians. He concludes that "Test users should especially note significant discrepencies between verbal and non-verbal scores in bilingual subjects, for whom failure to use non-verbal tests is clearly a form of psychometric malpractice" (p.619).

It is important to note, however, that all of these studies were conducted in educational settings, which are different from industrial or occupational settings and from occupational testing. The skills needed for success at school, as Kirkpatrick et al (1969) point out, could be quite different and would need different tests from those needed for success on a given job. Furthermore, the criteria for success will vary from job to job

which means that each individual test would have to be examined for possible unfairness regardless of what these educational studies show. Therefore, whilst they provide some helpful lines of enquiry, studies are needed in occupational or industrial settings to provide answers to the problems of unfair discrimination.

3.4 Major Sources of Bias in Testing

Despite the fact that testing procedures are standardised in terms of administration, scoring, etc, tests can still be biased against particular groups. Some of the factors which cause bias are identified here and they include atmosphere and conditions in which the testing is conducted; the small range of abilities which are accurately measured; the use of poorly validated tests; the interpretation of scores.

3.4.1 Bias Due to Atmosphere

The ways in which a test is administered can be a cause of bias due to the anxiety it creates within candidates.

One area of possible anxiety is that of time limits. The results are not conclusive, however, some researchers (eg. Flanagan 1971) feel that many minority groups have difficulty adjusting to time limits, whereas others (eg. Evans and Reilly 1972) have found that extending or shortening time limits does not significantly help minority groups.

There is considerable agreement, however, that the degree of test sophistication of a candidate is an important determinant of how well they will perform on the test. Tests are couched in an artificial format and largely use multiple choice item forms, have stop and start signals, etc. The test sophisticated candidate has become used to this artificiality and adapts to the situation quickly, does not waste time on items to which he/she sees no immediate solution, etc. It has been shown that the Western middle class school-child has greater opportunity to gain this sort of experience compared with some minority groups (eg. Biesheuvel 1949, Irvine 1966, Vernon 1969).

These biases can be countered, however, and Pearn (1975) suggests the following:

(i) giving longer and more detailed instructions which should be written down as well as read aloud

(ii) giving practice on similar tests or similar test items. The more practice candidates have with the type of test they will be taking, the more equal they become in terms of their familiarity with the test

(iii) establishing a rapport with candidates to reduce their anxiety and help them feel at ease

(iv) providing some form of coaching, advice and instructions on how to tackle different types of problems and tests. All candidates would receive this information, but it has greater effect for handicapped candidates and produces test results which are more indicative of their true abilities

3.4.2 Bias Due to Abilities Measured

Although there is a variety of tests commercially available they in fact only measure a narrow range of skills and abilities and these become over-emphasised when assessing people for jobs.

In a survey of test use in the UK Brotherton et al (1978) concluded that "the present array of selection measures within companies could be somewhat reduced as there is much redundancy in the information collected. Performance on one test frequently predicts performance on another" (p.28).

The areas which are measurable, though, are measured extremely well, but the lack of validity for tests generally is due, says Flaugher (1976), to those areas which cannot be measured accurately. This has led to psychologists over-emphasising those skills and abilities which can be measured, and this could lead to bias "because people could be screened out of jobs on the basis of tests which are not relevant to the position in question. When these tests are represented as assessments of highly valued parts of the spectrum, the issue of test bias is legitimately raised, since it is a leap from being unable to work out a few problems on a paper and pencil test to being declared lacking in practical judgement." (Flaugher 1976 p.6/7).

3.4.3 Bias Due to Irrelevant Test Standards

In times of high unemployment, with a large capable pool of people from whom to select, the temptation for many organisations is to ask for a higher level of skills or qualification than is actually required to perform the jobs adequately, and so could lead to indirect discrimination against some groups. An industrial

tribunal case involving the British Steel Corporation illustrates this point (Runnymede Trust 1979). Here, nine Bangladeshi workers overstayed their holidays in Bangladesh and effectively terminated their contracts. On re-applying for their labourers posts they found a test of 'literacy' had been introduced which they sat and failed. However, they had all worked successfully as labourers previously, so the test did not appear to be measuring abilities relevant to the job. A subsequent job analysis revealed that the test was setting unnecessarily high standards of English for a post where written English was not relevant. This not only is an example of indirect discrimination but also illustrates the need for conducting job analysis prior to introducing tests.

3.4.4 Bias Due to Lack of Validity

Organisations appear often to use tests without any regard for their efficiency, or validity in selecting people.

Sneath et al (1976) found that of organisations using tests 51% had not attempted any form of validation, and only 5% applied any statistical techniques to their data. They felt that personnel specialists relied upon face validity and that not enough time or resources were allocated to the problems of validation.

In a very thorough survey of the effectiveness of using tests Ghiselli (1966) found that validity coefficients were in the order of 0.5 for predicting training criteria but 0.3 or less for job performance criteria. In an intensive survey of nine organisations in Britain, Brotherton et al (1981) found validity coefficients of between 0.27 and 0.59 for training criteria and 0.16 to 0.46 for

proficiency criteria. They surmise that one of the reasons for the low validities is because "too few instruments are being devised to meet the needs of industry. Too many instruments in use in Britain have originated in the United States and have not been appropriately modified to meet new cultural, linguistic and normative demands" (p.66).

3.4.5 Bias and Society

Much time has been devoted to the development of 'culture-fair' tests. As Arvey (1972) points out, these efforts have largely been in vain since they often tend to be measures of general ability and it is not clear what they are measuring. In addition, they are not always 'fairer' to minority groups as they sometimes widen rather than close the gap in average scores. Anastasi (1976) feels that tests should reveal differences between groups. "Differences in the experiential backgrounds of groups or individuals are inevitably manifested in test performance. Every psychological test measures a behaviour sample. In so far as culture affects behaviour, its influence will and should be detected by tests. If we rule out cultural differentials from a test we may thereby lower its validity as a measure of the domain it is designed to assess" (p.58). It is important to remember that differences between people of different backgrounds should only be considered relevant if they affect performance on the job. Bias, therefore, can occur either through the way the test is used and/or in the way the scores are interpreted. This implies tests can be fair or biased according to the circumstances in which they are applied - a test may be appropriate in certain situations but not in others. As Flaugher (1976) says, "a test

cannot be better than the person using it; if the test is used improperly the fault quite possibly lies with the user and not with the test" (p.5). He goes on to argue that in interpreting test scores a wider perspective should be adopted than currently prevails. So if applicants from a particular group receive low scores on a test it is essential not to reject that group out of hand, but to discover whether low scores mean low performance at work for that group and to attempt to discover why those low scores came about. This is the issue of differential validity which will be discussed in the following section.

3.5 Research Evidence on Differential Validity

In the areas of differential and single-group validity, discrimination is not just taken as a difference in mean test performance for two groups, but also as a question concerning the nature of the relationship between test and criterion - groups could differ in criterion as well as test performance and it is important to establish the relationship between these two variables.

The definitions of single group and differential validity which are most widely accepted are those provided by Boehm (1972). Differential validity, she says, exists where "(a) there is a significant difference between the correlation coefficient of a selection device and a criterion obtained for one ethnic group and the correlation of the same device with the same criterion obtained for the other group. And (b) the validity coefficients are significantly different from zero for one or both groups" (p.33).

Single group validity is where the validity coefficient is significantly different from zero for one group only, but where there is no significant differences between the two groups.

Boehm (1972) went on to examine 13 studies looking at Negro-white differences. In all there were 160 validity coefficients - of which 100 were not significant for either group. From the remaining 60 cases she found 22 cases of single group validity and 7 of differential validity - approximately one half of cases studied.

Despite this Boehm concluded in a widely quoted passage that "The concept of differential validity has received far more attention than the data would seem to justify." (p.37). In a widely ignored passage, however, she goes on to say that "more research is needed in this area as even the occasional instance of differential validity can have consequences both socially and practically undesirable" (p.39).

Single group validity has been shown for three major reasons to be a theoretical rather than real problem. Firstly Zedeck (1971) feels that the concept actually does not mean much. "The validity coefficients for two groups might be 0.40 and 0.42 - one being significantly different from zero but the other not. It would not be reasonable to conclude that the predictor is useful only for one group. Nor does a higher correlation coefficient necessarily mean better prediction because the Standard Error of Estimate has also to be considered."

Secondly, Schmidt et al (1973) found that the number of times single group validity occurred was not more than would be accounted for by chance.

Finally, Bartlett et al (1977) showed that single group validity is only a sample problem and "... is due to insufficient information about the underlying population. When the phenomenon is encountered in research it should be a warning that the sample information is insufficient and that further research may be needed to see if differential validity is operating in the population" (p.157).

There is far less agreement, however, on the existence of differential validity. Indeed it is an area which has revealed more about biases in researchers rather than in the tests they were studying. A typical example of the kind of work which characterises this topic is Fisher (1975) who makes very bold claims on limited evidence. After reviewing the evidence for differential validity he claimed that "Neither differential validity nor statistical adjustments for test bias is supported by the empirical data available" (p.481). He claims that the results of studies where differential validity was found were 'tenuous'. Two studies attracted his praise, however, neither of them finding any evidence of differential validity. The studies of Grant and Bray (1970) and Gael and Grant (1972) are described as "well-conducted, involved substantial numbers of white and black employees and used well- developed proficiency criteria" (p.481). For both studies the proficiency criteria were performance on other, specially developed tests. Although the majority group scored higher on the tests the slopes and intercepts of the regression lines were not significantly different. The major flaw with these studies is self-evident - they both correlate performance on one test, which they choose to call the predictor, with

that of another, which they choose to call the criterion. Even though the criterion tests may be job-related, if it is factors in the testing situation which are causing the differences in test scores then use of such methodology will lead to minority groups performing less well on both occasions.

Schmidt and his colleagues are the strongest opponents of the differential validity hypothesis, and after reviewing the subject they stated that it was in fact only a "pseudo-problem" (Schmidt et al 1972 p.8).

Katzell and Dyer (1977) criticised this finding on two points. Firstly, the tests used had not previously been shown to be significantly correlated with a relevant unbiased criterion of performance on the job in question. Secondly, many of the tests reviewed by Schmidt et al were not significant for either group. This compounds the above problem because "if tests are truly invalid for the job in question, it is obviously nonsensical to use them to investigate the hypothesis that they are invalid in both groups with equal frequency" (p.138). Furthermore, Schmidt et al (1973) had broken very basic assumptions in the statistical techniques they had used.

Katzell and Dyer (1977) therefore took those correlation coefficients which were 'appreciable' for at least one group. They found that the number of significant differences between validity coefficients was itself significant. They concluded that "this set of studies casts serious doubts on the hypothesis that tests are equally valid in the two ethnic groups" (p.44). Hunter and Schmidt (1978) criticised the method of preselecting validity

coefficients and they in turn were criticised by Boehm (1978) who claimed that their (Hunter and Schmidt's) research "appear to reflect an idealised world far removed from the one in which industrial-organisational psychologists practice - a world in which statistics is regarded as a purely mathematical (rather than probabilistic) enterprise" (p.182).

Similar debate followed a paper by O'Connor, Wexley and Alexander (1975) who claimed that "... there is no evidence that selection tests are more often valid for whites than blacks" (p.354).

Bartlett et al (1977) again found flaws with the techniques used - and again basic assumptions were broken regarding the statistical tests employed.

It can be seen, then, that the major reviews of studies of differential validity which loudly proclaimed that it is not a problem have very serious methodological faults. To overcome this Katzell and Dyer (1977) have laid down a set of conditions which would have to be satisfied in order to test the null hypothesis that no differences exist between groups. The only real conclusion which can be drawn is that made by Katzell and Dyer (1978) when they observe that "because of their theoretical and methodological deficiences the available set of studies cannot serve as a basis for firm conclusions on the issue of differential validity, much less the conclusion that the null hypothesis has been proven" (p.21).

The issue is far from being resolved. The research to date has had flaws, methodological weaknesses and, perhaps worst of all, has appeared biased. What is needed is more research more objectively carried out (possibly following Katzell and Dyer's 1977 principles). In short, as Katzell and Dyer (1977) state "... what is needed is not more research, but better research" (p.144).

3.6 Theoretical Models of Selection Fairness

Over the past few years there has been a debate in the United States concerning different ways of formally defining fairness in selection. The debate has attracted attention and often heated arguments, because some approaches recommend or require the use of double standards to increase the numbers of women and ethnic minorities in certain types of jobs, in education, etc. Since this debate is not of direct relevance to this project, but is nevertheless of some interest, the models of fairness will be discussed briefly.

Jensen (1980) provides, in his "Essential Definition of Predictive Bias", a model for the perfectly fair test.

This model is psychometrically and logically sound and appears to be an indisputable definition of a fair test. How, then, can there be other models of fairness? Firstly, Jensen's model is only a theoretical one since he makes assumptions which do not exist in the real world such as that of perfect reliability. Secondly, other models do not attempt to define fair tests per se, but their fair use.

Hunter and Schmidt (1976) categorise the various models into three philosophical stances which are:-

(i) Unqualified Individualism - where the person appointed should be the one best qualified to do the job. Adoption of this philosophy would minimise selection error, but it could also lead to race or sex being explicitly used as a predictor of future performance.

(ii) Qualified Individualism - this philosophy is identical to Unqualified Individualism when a fair test is used. With less fair tests, however, the test user would take societal values into account and would refuse to use variables such as race, sex, etc as predictors. Adoption of this philosophy, however, could lead to a reduction in the tests' effectiveness.

(iii) Quotas - a philosophy which does not give psychometric considerations the main priority, focussing instead on societal values and attempts to compensate for the past exclusion of minority groups.

The Regression Model (Cleary 1968) perhaps best embodies the Unqualified Individualist view, but it also includes the Equal Risk Model (Einhorn and Bass 1971) and the Multiple Regression Model (McNemar 1975).

The Constant Ratio Model (Thorndike 1971), the Conditional Probability Model (Cole 1973) and the Equal Probability Model (Linn 1971) are group parity models and fit into the qualified individualist philosophy.

There is, then, an abundance of selection models and it is difficult at times to detect the differences between some of them. This confusing state of affairs was recognised by the American Psychological Association (1974) who stated that "it is important to recognise that there are different definitions of fairness and whether a given procedure is or is not fair may depend upon the definition accepted" (p.44).

All of these models purport to provide the definition of the fair use of tests but the subjective judgements of the model builders are evident in their design. In fact, Breland and Ironson (1976) compared the results of applying a number of models to a real-life selection situation and found they each produced very different results - each of them unfair to the others.

The model builders are attempting to relieve a conflict between maximising test validity whilst at the same time reducing test discrimination. This conflict is resolved says Darlington (1971) by a subjective decision taken by model builders on behalf of organisations and society. Furthermore, Petersen and Novick (1976) showed that these models were in fact internally inconsistent and thus logically unsound. "What is required" say Novick and Ellis (1977) "is a means of awarding compensatory treatment based on individual disadvantage rather than the possession of racial or ethnic characteristics" (p.18). There needs to be a move away from the use of prescribed models and towards allowing organisations to make this subjective decision themselves and to state it explicitly.

This has led to the call for the use of statistical decision theory such as that developed by Gross and Su (1978). This system looks at people as individuals and not as members of groups and takes into account their unique combination of advantages and disadvantages. The problem though is in attempting to determine what factors constitute disadvantage and how much weight should be given to each of them (Pearn 1979). This indicates the vast amount of data that will have to be collected and the enormous amount of research that will have to be conducted before such a model can be fully used.

3.7 Conclusions

The work reviewed here has highlighted many of the problems involved in test use, and, in the United States, they have led to calls for the banning of tests altogether, or, not withstanding that, at least a moratorium on their use. These calls have been dismissed on two points. Firstly, it is said that it is society which is biased and that tests merely reflect that. Banning tests would only be killing the messenger of bad news - nothing would have been done to improve the situation on a larger scale. Secondly, it would lead to the use of even more discriminatory techniques. The Equal Employment Opportunity Commission in the United States recognise the benefits of good selection. In the "Guidelines on Employment Procedures" (1970) the EEOC stated that "properly validated and standardised employee selection procedures can significantly contribute to the implementation of non-discriminatory personnel procedures as required by Title VII (of the Civil Rights Act). It is also recognised that professionally developed tests when used in

conjunction with other tools of personnel assessment and complimented by sound programs of job design may significantly aid in the development and maintenance of an efficient workforce and, indeed, aid in the utilisation and conservation of human resources generally".

In conclusion, then, it should be stated that the advocates of equality of opportunity are generally not against the use of tests - just the use of tests in an improper or inefficient manner. By using tests which have been properly constructed and implemented following professional and established procedures the discriminatory effects of tests should be reduced and validity maintained.

CHAPTER 4

The Overall Objectives and Methodology of the Pilot Studies Comparing Classifications for Use in Monitoring

4.1 Overall Objectives of the Pilot Studies

The overall objectives of the pilot studies were to

(i) compare and evaluate a number of classifications which are based on different criteria eg. colour, racial group, etc.

(ii) evaluate the effectiveness of asking the monitoring question in different ways eg. the standard method used in classifications is to ask respondents to categorise themselves based on how they perceive themselves. If, however, the purpose of a classification is to monitor equality of opportunity it might be more useful to discover how the respondents think they are perceived by others.

The criteria for evaluating the classifications and the questions were determined after examining some of the problems associated with ethnic monitoring. These included examining:

(i) the way in which respondents reacted to particular classifications and questions

(ii) the reliability of the classifications

(iii) the ways in which people classify themselves

The discussion below outlines how these criteria were formulated.

4.2 An Examination of the Criteria Used in Classifications

Useful, meaningful classifications can only be developed, once their purposes have been defined. Classifications used in employment situations are intended to monitor equal opportunities and to identify the disadvantaged groups. The problem thereafter becomes one of obtaining a reliable criterion for grouping people together. Different methods of classifying people are discussed along with the problems associated with them.

Bulmer (1980) in defending the use of 'ethnic' questions in censuses, says that such questions attempt to distinguish people "in terms of their situation in a particular society which makes them socially distinct groups in their own eyes and in the eyes of others." (p.5). He goes on "what the ethnic question is trying to do is to find out in an as objective a way as possible how members of British society identify themselves in terms of the racial or ethnic categories prevalent in that society." (p.5).

If this is accepted then it means that there will be little consistency between classifications used in different societies because the groups which it is appropriate to distinguish in a particular country will depend in every case on its social composition and historical development.

So, firstly you need to identify those labels most widely used within a given society to distinguish between people. This information, however, is not collected purely to provide information on numbers or even of popularity of labels used. Such information is collected by particular societies because "these societies' racial or ethnic divisions are closely associated with social and

economic inequalities, wherever such inequalities exist there is a case on policy grounds for collecting information about them." (Bulmer 1980, p.9).

So classifications whether used for monitoring or in censuses attempt to provide information on people who have suffered from unequal opportunities and deprivation of one sort or another, and in doing this people are grouped together according to the labels which are most prevalent, widely used and accepted by members of that society.

The problem then becomes one of deciding on the 'best' criterion (or criteria) for grouping people together.

In the past, classifications have attempted to group people into separate races on the basis of physical criteria, eg. skin colour, cephalic index, or on a combination of factors. These attempts revealed that mankind could be split, almost arbitrarily, into any number of races - some researchers have found only 3, some 17 and some 34. Benedict (1942) said that, "In all modern science there is no field where authorities differ more than in the classification of human races." (p.19). There are too many similarities between people to be able to split them up into separate physical categories. The systematic study of categorising people is what Montagu (1974) has called the scientific image of race. It is an image of race based on facts and the results are "overwhelmingly clear that biological traits of both individuals and populations isolates overlap to such an extent that similarities overshadow any distinctive differences" (p.352).

Neither does it help to look not at 'race' or 'racial group' but at 'ethnicity' or 'ethnic group' as a way of differentiating between people. In fact, the distinction between race and ethnicity is blurred and the terms, in this country at least, are often used interchangeably, eg. an OPCS question pilotted prior to the 1981 census and used in the subsequent studies asked respondents "to show the racial or ethnic group" to which they belonged (OPCS 1980).

Some 'ethnic' classifications, like some 'racial' classifications, are in fact based on identifying people's skin colour. Other ones appear to be linked more closely with nationality or the origins of the person's forebears. The essential ingredient of all the classifications, on whatever criterion they are based, is that they are attempting to identify disadvantaged groups reliably. However, there have been few, if any, studies comparing the effectiveness of these different classifications for monitoring equal opportunity. One of the main objectives of the pilot studies, and of the project overall, was to compare different types of classification and examine their suitability for monitoring purposes.

4.3 An Examination of the Problems Associated with Questions and Terminology

Some of the most important issues connected with monitoring are those dealing with terminology, ie. what sorts of questions to ask, what classifications to use, and the analysis of results. The Committee on Race Relations Research (1975) focussed on the issues of terminology and said that any data collected must be

suited to the purpose for which it was to be used and "what this requires is the use of terms that are unambiguous and well understood, and that also lend themselves to the more precise definitions necessary to sort out marginal cases." (p.12).

The Committee also felt that any terminology developed should firstly be able to identify people's race or ethnic origins in terms of colour and country of origin. Secondly that exploratory testing was needed, thirdly that no term which was felt to be derogatory should be used, and finally that guidelines should be developed showing which kind of terminology is likely to be most appropriate under different circumstances. They also make the point that the terminology used may well be subject to change as their usage changes in society. To date, there is little systematic work which has been conducted into those issues in the specific area of monitoring.

In terms of questions, White (1980) says that there are two basic questions which could be asked, with three possible hybrids. Firstly, you could have a list of all the possible labels which are applied to people, and the respondent ticks the one which he/she sees as being the most relevant. Alternatively, the respondent could be asked to state their national origins but instead of having a list to refer to, a blank space is provided and the respondents fill in what they think their origins are.

The hybrids possible from these two include:

(i) listing all the possible responses and adding a space for any description not provided

(ii) ending the list with the word - "other"

(iii) listing all groups of interest and assigning all other respondents to an undifferentiated miscellaneous category.

There is, however, another type of question which could be considered for monitoring certain situations, eg. interviews. In such situations how the candidate perceives him or herself might be less important than how the interviewer perceives him/her. For example, if the candidate perceives himself to be British, but the interviewer considers him to be West Indian and is prepared to discriminate unlawfully on the basis of race, then it is the interviewer's perception of the candidate which is most important. In such instances, therefore, it might be useful to ask the candidate what other people might think they are.

4.4 An Examination of the Suitability and Reliability of Classifications

Although there are very few research studies which have been conducted on ethnic monitoring, some lessons can be learnt from the experiences of the Office of Population Censuses and Surveys (OPCS) in attempting to develop an 'ethnic' question for the 1981 census.

The most difficult problem which they encountered was that of obtaining a suitable format for the question and categories. Many different forms of question were tried out. Some, for example, listed many categories, including European. It was found that many West Indian parents were placing their children in the European category rather than the West Indian. People from the Indian sub-continent said that they would like to be differentiated according to their religions as well as their national origins, but when this was attempted respondents sometimes only completed the religious category and not the national origins one. Further issues concerned people of mixed ancestry who were unsure of which category to use and recording the origins of children born in this country of immigrant parents. Many of the latter see themselves, quite legitimately, as being British.

However, if such information is indeed being collected to examine social and economic inequalities, and it is believed from research evidence (eg. PEP, 1974) that minority groups who have foreign origins are the people suffering such privations, then the classification of themselves as British will not enable a full examination into the extent of the problem to take place.

Therefore some method of asking the question is required which will identify and distinguish between the national origins of minority groups.

Another problem encountered during the field trials (OPCS 1980) was the amount of opposition to the concept of asking questions on people's origins at all.

Eventually, it was decided to drop the question entirely from the 1981 census. The reason given was that it was "impracticable at this stage to devise a question on ethnic origin which would give reliable results in the 1981 census and would not put at risk the success of the census as a whole." (p.1 OPCS, 1980). The implications behind this were that:

(i) the form of the question was inappropriate

(ii) it would not be reliable

(iii) it would not be acceptable.

These are the main problems associated with ethnic monitoring, and ones which had to be addressed in the pilot studies.

4.5 General Methodology of the Pilot Studies

In this section the methodology will be explained, in particular how the classifications, questions and samples were chosen.

4.5.1 The choice of classifications

In order to meet overall objective (i) (see section 4.1) care had to be taken in choosing the classifications. It was decided that the classifications to be used and compared in the pilot studies should be based upon:

(i) a different criterion (eg. colour, racial groups as used in anthropology, specific identifications of national origins (ie. English, Indian, West Indian, etc), and general indicators of national origins (eg. Asian, African, European, etc.)) in order that comparisons could be made of a variety of classifications

(ii) ones already used for monitoring in order to increase and enhance the practical relevance of the project.

Five classifications were adopted (see Appendix 2)

Classification 1 : OPCS - this is a classification which was tested by the OPCS during its pilot studies prior to the 1981 census. Although this was not the classification finally chosen to be used in the census it is the one which comes most closely to having a single criterion of classification. This criterion is national origins - so it wants people who are immigrants or are descended from immigrants, to state from where their forebears came. It is the most specific of the five classifications chosen containing as it does thirteen separate classifications plus one category of 'other' for those not covered by any of the individual categories.

Classification 2 : CRE - This is the classification which the CRE recommends employers to use when monitoring. It is far less specific than the OPCS classification and contains five fairly general categories, ie. African, Asian, European, Caribbean, UK European and Irish, and a category of 'other' for those not covered by any of the above.

Classification 3 : Tavistock - This is the classification which was adopted by the Tavistock Institute in their Civil Service study and contains only two categories, those of "coloured" and "white" and they refer to skin colour. This is the shortest classification and possibly the crudest.

Classification 4 : Multi-Cultural British (MCB) classification - This is the only classification which was specifically designed for the pilot studies. It has eleven separate categories plus one of 'other' but it has more in common with the generalised categories of the CRE than the specific ones of the OPCS. The factor which differentiates this classification from the others is that it acknowledges that there are many people in this country who are seen as belonging to minority groups, but who are also British. Therefore not only is there a British category, but there are also categories of British Asian, British West Indian, British African, British Chinese. The categories are then repeated but without the British prefix. It was hoped that this would enable people to state their origins while at the same time acknowledging that some minority group members are British (an objection encountered during the OPCS trials).

Classification 5 : Anthropological - This is based upon anthropological work which has attempted to classify people according to specific racial groups. The three categories in this classification are thought to be the most mutually exclusive which have been achieved.

4.5.2 The choice of questions

For each classification the question introducing it was asked in two different ways representing, in effect, different instructions as to how the classification should be completed.

The first method of questioning (called the standard method) basically asked "Please state what category you belong in on the classification below." The second method (called the reversed method), asked "Please state what category you think someone else might think you belong in." Thus it would be quite possible to get two different responses on the same classification.

The reversed question was developed because it was felt that if a person is being discriminated against it is not his/her self-perceptions which are relevant but instead the perceptions of the other person taking the decision. This could be particularly useful in any face-to-face situation. Furthermore, since so many people objected to being directly asked to describe their ethnic origins in the pilot studies prior to the 1981 census, it was thought that if the question was asked in this indirect fashion they might not object as strongly.

4.5.3 Obtaining the samples

Given the time scale it was felt to be essential that organisations which we approached for access to conduct our questionnaires:

(i) should respond quickly and positively to our request

(ii) should be able to provide a multi-racial sample of people from their organisations

(iii) should provide a sample partly comprised of school-leavers, graduates and others who might be eligible to apply for the EO grade.

Access was eventually gained to a hospital in Birmingham, and also to the Birmingham and Coventry Careers Services. In addition permission was obtained to question university students at Aston, and also to test and re-test a particular lecture group.

CHAPTER 5

Comparison of Classifications For Use In Monitoring For Employment Selection: Stage 1 Of Pilot Studies

This chapter describes Stage 1 of the pilot studies on the classifications for use in monitoring. In Stage 1 there were two pilot studies: (i) Pilot Study 1 - comparing classifications and questions in terms of their suitability and the extent to which respondents objected to them; (ii) Pilot Study 2 - examining test-retest reliability.

The two pilot studies are described separately in terms of their objectives, method, results and conclusions. The overall conclusions are presented at the end of the chapter.

5.1 Pilot Study 1 : Comparison of Classification/Questions on Levels of Suitability/Objectionability

5.1.1 Objectives

There were four main objectives of this pilot study which were to examine:

(i) the extent to which respondents objected to each of the classifications, and which ones they felt to be most suitable for monitoring purposes.

(ii) which type of question was thought as most suitable, ie. the 'standard' or 'reversed' question. (It should be noted that the primary focus of attention here was the actual question and not the classification.)

(iii) how respondents classify themselves on each of the classifications. The purpose of using ethnic or racial classifications is to monitor equality of opportunity. Any classification used, therefore, must be able to identify the disadvantaged groups in society. This in turn requires the provision of categories in the classifications which people from different groups can identify with and which are unambiguous. If members of minority or disadvantaged groups do not categorise themselves as such, then the purpose of having a system of monitoring will have been negated. In this pilot study, therefore, it was attempted to determine how people of similar origins classified themselves. If the classifications were being interpreted in the same way, then people of similar origins should, to a large extent, categorise themselves in the same way. If the classifications and the categories within them are ambiguous or unclear then there should be a lack of consistency in people's responses.

5.1.2 Method

The development of the questionnaire is described first, followed by a description of the sample.

5.1.2.1 The questionnaire

A questionnaire was developed (Appendix 2) which included the five classifications, with standard and reversed questions, which respondents were asked to complete. Below each classification respondents were asked to state on an interval scale from 1 (suitable/not object) to 5 (unsuitable/object very much):

(i) how suitable they thought the classification was for monitoring purposes

(ii) how suitable they thought the question was

(iii) the extent to which they would object to being given such a classification by an employer.

At the bottom of each page a space was provided for any additional comments respondents wanted to make.

In addition respondents were asked to state:

(i) their own country of birth

(ii) the country of birth of their parents

(iii) their sex and age

(iv) whether their educational qualifications were greater than, equal to or lower than 5 'O' levels and 2 'A' levels (the minimum level of qualifications required for an Executive Officer).

The responses to the country of birth questions (in particular the parents' country of birth), were accepted as an indicator of respondents' origins. This provided the method for grouping together people of similar origins.

5.1.2.2 Description of the sample

The total sample consisted of two sub-samples:

(i) school-leavers using the Careers Services in Coventry and Birmingham, plus nurses from a Birmingham hospital

(ii) undergraduate students at Aston University (see Table 5.1).

(i) The Careers Service/Nurses sample

Sixty questionnaires with covering letters of explanation were given to a hospital nursing administrator with the instructions

Table 5.1 Origins of sample as determined from parents' place of birth

ORIGINS OF PARENTS OF RESPONDENTS	UNIVERSITY STUDENTS	SCHOOL LEAVERS/ NURSES	TOTAL
United Kingdom & Eire	7	59	66
West Indies	5	36	41
India	22	12	34
Pakistan	5	5	10
Sri Lanka	3	-	3
Nigeria	13	-	13
Sudan	6	-	6
Zambia	5	-	5
Zimbabwe	3	-	3
Ghana	2	-	2
East Africa	2	-	2
Malaysia	11	-	11
Singapore	3	-	3
China/Hong Kong	7	-	7
Indonesia	2	-	2
Middle-East	8	1	9
Greece	3	-	3
Italy	-	2	2
Poland	-	2	2
Other European	-	3	3
Other Non-European	2	3	5
	109	123	232

that of these half should be given to ethnic minority group members and half to the majority group members.

25 completed questionnaires were returned - 19 from people identifiable to the majority group and 6 from ethnic minorities; producing an overall response rate of 41.75%.

In order to increase the response rates, particularly from minorities, it was decided that a direct approach should be made to respondents so that a personal explanation could be provided about the purpose of the studies.

This method was used for the school leavers using the Coventry and Birmingham Careers Services. School leavers were approached because they comprise a large proportion of the applicants sitting the Executive Officer tests. At the end of a normal careers interview the school leavers were asked if they would complete the questionnaire and the response rate was 100% using this approach.

(ii) The University sample

An individual approach was adopted with the university sample. On this occasion students studying in the library were approached. The library was chosen for the sampling because there is a high turnover of people so that on different occasions during the day it will be populated by different people. Since people were often too busy to complete them at that moment, they were asked to leave the completed questionnaires at a particular point in the library. The response rate for completed questionnaires in this case was 64%.

5.1.3 Results

The results are presented under three major headings. First the results of the opinion questions are considered, with each question being treated separately. Second there is the analysis of the respondents' comments; and finally the degree of correspondence between self-ratings and the external criterion of parents' birthplace is examined and particular issues of self-classification discussed.

5.1.3.1 Examination of the three opinion questions

An analysis of variance with repeated measures was applied to each of the separate groups of means (see Table 5.2). The suitability of the classifications is considered first, followed by the analysis of the suitability of the questions and finally an analysis of the objections to the classifications.

Suitability of classification

An analysis of variance was applied to the means for the total sample for the question on suitability of the classification for monitoring purposes and the result was significant ($F=25.11$, $df=9,231$, $p<.01$). (For analysis of variance table see Appendix 3a.)

Significant values were also found when the analysis of variance was applied to the school leaver/nurses sample ($F=17.11$, $df=9,122$, $p<.01$) (for analysis of variance table see Appendix 3b) and the university sample ($F=13.37$, $df=9,108$ $p<.01$) (see Appendix 3c for analysis of variance table).

Table 5.2 Mean Scores for Three Opinion Measures on all Classifications for School Leavers/Nurses, University Students and Total Samples

	STANDARD QUESTION			REVERSED QUESTION		
	SCHOOL LEAVER/ NURSES SAMPLE	UNIVERSITY SAMPLE	TOTAL SAMPLE	SCHOOL LEAVER/ NURSES SAMPLE	UNIVERSITY SAMPLE	TOTAL SAMPLE
OPCS						
Suitability of Classifications	2.54	2.71	2.62	3.07	2.99	3.03
Suitability of Question	2.65	2.94	2.79	3.19	2.96	3.08
Degree of Objectionability	2.49	2.91	2.69	2.92	2.98	2.99
CRE						
Suitability of Classifications	2.93	2.48	2.71	3.20	2.82	3.03
Suitability of Question	2.97	2.82	2.76	3.4	2.96	3.17
Degree of Objectionability	2.71	2.59	2.65	3.14	2.99	3.07
TAVI						
Suitability of Classifications	3.14	3.60	3.35	3.08	3.37	3.21
Suitability of Question	3.28	3.72	3.48	3.27	3.44	3.35
Degree of Objectionability	3.20	3.77	3.46	3.03	3.54	3.28
MCB						
Suitability of Classifications	2.65	2.62	2.63	3.07	3.06	3.07
Suitability of Question	2.90	2.67	2.75	3.14	3.14	3.14
Degree of Objectionability	2.65	2.62	2.63	3.02	3.11	3.06
ANTH						
Suitability of Classifications	3.93	3.65	3.81	3.79	3.35	3.6
Suitability of Question	3.94	3.70	3.83	3.78	3.49	3.65
Degree of Objectionability	3.64	3.50	3.62	3.55	3.42	3.49

CODE: OPCS - OFFICE OF POPULATION CENSUSES AND SURVEY'S CLASSIFICATION
CRE - COMMISSION FOR RACIAL EQUALITY'S CLASSIFICATION
TAVI - TAVISTOCK INSTITUTE'S CLASSIFICATION
MCB - MULTI-CULTURAL BRITISH CLASSIFICATION
ANTH - ANTHROPOLOGICAL CLASSIFICATION

In order to discover the significance between individual means, Tukey's (1949) multiple comparison of means was applied. Table 5.3 shows the significant differences between means for the total sample.

Table 5.3 Significant differences between means for total sample on suitability of classification

		STANDARD QUESTION					REVERSED QUESTION				
		OPCS	CRE	TAVI	MCB	ANTH	OPCS	CRE	TAVI	MCB	ANTH
OPCS	SQ		NS	**	NS	**	**	**	**	**	**
CRE	SQ			**	NS	**	NS	NS	**	NS	**
TAVI	SQ				**	*	NS	NS	NS	NS	*
MCB	SQ					**	**	**	**	**	**
ANTH	SQ						**	**	**	**	NS
OPCS	RQ							NS	NS	NS	**
CRE	RQ								NS	NS	**
TAVI	RQ									NS	NS
MCB	RQ										**
ANTH	RQ										

Key

NS - Insignificant

* - Significant to .05 level

** - Significant to .01 level

OPCS Office of Population Censuses & Surveys classification

CRE Commission for Racial Equality classification

TAVI Tavistock classification

MCB Multi-Cultural British classification

ANTH Anthropological classification

The means between OPCS, CRE and MCB were not significantly different from one another, and the MCB and OPCS were also significantly more suitable than all other classifications except for the CRE. The CRE classification on the other hand was only seen as being significantly more suitable than the Tavistock and anthropological classifications (standard and reversed). The anthropological classification was significantly less suitable than virtually all other classifications.

The significant differences between classifications for each of the subsamples was also considered.

The university students considered the CRE classification to be most suitable but there were no significant differences between their three most favoured classifications, ie. CRE, MCB and OPCS. In this respect the responses of the university sample most closely resembled those of the sample taken as a whole. (For table showing significant differences see Appendix 4a.)

The most striking feature of the school leavers/nurses sample was that the OPCS and MCB classifications were considered to be significantly more suitable than the CRE classification ($p<.01$, and $p<.05$ respectively). There were no significant differences between the OPCS and MCB classifications. (For table of significant differences see Appendix 4b.)

The responses to the opinion questions of the three largest groups in the sample were examined. These groups were United Kingdom and Eire (n=66), West Indian (n=41) and Indian (n=34). (For full breakdown of mean scores for the opinion questions by ethnic

group see Appendix 5.) The UK group considered the OPCS classification to be most suitable followed by the CRE and MCB. The West Indian and Indian groups both considered the MCB classification to be the most suitable.

All groups considered the Tavistock and Anthropological classifications to be the least suitable.

Suitability of the questions

The analysis of variance on the total sample for the second opinion question on the suitability of the question was significant ($F=23.68$, $df=9,231$ $p<.01$) (For analysis of variance summary table see Appendix 6a).

Table 5.4 shows the significance of difference between individual means.

Table 5.4 Significant Differences Between Means for Total Sample on Suitability of Question

		STANDARD QUESTION					REVERSED QUESTION				
		OPCS	CRE	TAVI	MCB	ANTH	OPCS	CRE	TAVI	MCB	ANTH
OPCS	SQ		NS	**	NS	**	*	**	**	*	**
CRE	SQ			**	NS	**	*	**	**	**	**
TAVI	SQ				**	NS	**	*	NS	*	NS
MCB	SQ					**	*	**	**	**	**
ANTH	SQ						**	**	**	**	NS
OPCS	RQ							NS	NS	NS	**
CRE	RQ								NS	NS	**
TAVI	RQ									NS	**
MCB	RQ										**
ANTH	RQ										

For OPCS, CRE and MCB the means for the standard question were significantly better than the means for their reversed question. There were no significant differences between the standard and reversed questions for the Tavistock and Anthropological classifications.

For the standard questions, OPCS, CRE and MCB were not significantly different from one another but they were significantly different from all other means.

Once again, even though the analysis of variance was significant for the university group ($F=14.04$, $df=9,108$, $p<.01$) (for analysis of variance summary table see Appendix 6b) and for the school leavers/nurses group ($F=14.40$, $df=9,122$, $p<.01$) (for analysis of variance summary table see Appendix 6c) the significant differences between individual means were not the same for both groups.

For the university group (see Appendix 7a) only for the CRE and MCB were the differences between their standard and reversed questions significant ($p<.01$) - indicating that for these two classifications only, the standard question was thought to be significantly better.

For the school leavers/nurses sample (see Appendix 7b) the OPCS, CRE and MCB classifications had significant differences between their standard and reversed questions, with the standard question considered more suitable.

For the three largest ethnic groups (UK, West Indian and Indian) the standard questions for the OPCS, CRE and MCB classifications

were seen as more suitable than their reversed questions, and also more suitable than the questions for the Tavistock and Anthropological classifications (see Appendix 5).

Objections to the classification/question

The analysis of variance on the third opinion question (regarding the extent to which respondents would object if an employer were to present them with a classification of that sort) was significant for the total sample ($F=24.04$, $df=9,231$, $p<.01$) (for analysis of variance summary table see Appendix 8a).

The pattern of significant differences between individual means (Table 5.5) is similar to that of Table 5.4.

The means for the OPCS, CRE and MCB classifications on the standard question were not significantly different from each other.

Table 5.5 Significant Differences Between Means for Total Sample on Objections to the Classification

		STANDARD QUESTION					REVERSED QUESTION				
		OPCS	CRE	TAVI	MCB	ANTH	OPCS	CRE	TAVI	MCB	ANTH
OPCS	SQ		NS	**	NS	**	*	**	**	*	**
CRE	SQ			**	NS	**	*	**	**	**	**
TAVI	SQ				**	NS	**	**	NS	**	NS
MCB	SQ					**	*	**	**	**	**
ANTH	SQ						**	**		**	NS
OPCS	RQ							NS	*	NS	**
CRE	RQ								NS	NS	**
TAVI	RQ									NS	NS
MCB	RQ										**
ANTH	RQ										

The Tavistock and Anthropological classifications were objected to significantly more when compared with the OPCS, CRE and MCB classification.

The analyses of variance for the university sample ($F=15.20$, $df=9,108$, $p<.01$) (for analysis of variance summary table see Appendix 8b) and school leavers/nurses sample ($F=13.67$, $df=9,122$, $p<.01$) (for analysis of variance summary table see Appendix 8c) were significant. However, examination of the differences between individual means again reveal differences in the preferences of the two groups.

For the university sample's three most favourably rated classifications, OPCS, CRE and MCB (Appendix 9a) the CRE and MCB classifications were objected to less than the OPCS classification ($p<.05$). There was no significant difference between the CRE and MCB classifications.

For the school leavers/nurses sample on these same three classifications, (which also represented their most favourably rated ones), there were insignificant differences between all three (Appendix 9b). For both the groups the standard questions on these three classifications was significantly more preferable to their corresponding reversed questions.

Furthermore, both groups found the Tavistock and Anthropological classifications significantly more objectionable than the OPCS, CRE and MCB.

Examined by ethnic group all groups objected most to the 'Anthropological' and 'Tavistock' classifications (see Appendix 5).

5.1.3.2 Examination of respondents' comments

In examining the respondents' comments it was expected that there should be some relationship between them and the responses to each of the three opinion questions.

In all 251 comments were made and most of these were made on the least popular classifications on the opinion questions, ie. the Tavistock and Anthropological classifications. The least number of comments was made on the MCB classification.

Although some comments are specific to particular classifications, content analysis revealed that 80% of them could be placed into four broad categories, thus enabling some meaningful comparison between classifications to be conducted.

These four categories were:

(i) comments which stated, in effect, that people should not be judged on their racial or ethnic backgrounds but by their abilities and qualifications (there were 31 such comments in all)

(ii) comments stating that more categories should be included (49 comments)

(iii) comments stating that the classification should be scrapped completely (80 comments)

(iv) comments stating that alterations should be made to the terms or questions used (40 comments)

Table 5.6 shows how these four categories of comments were distributed over the five classifications. The Tavistock and Anthropological classifications received a lot of comments in each of the categories. In particular many comments called either for alteration of their terms or for their complete removal.

Table 5.6 Percentage of Each Category of Comments Made for Each Classification

COMMENT	OPCS	CRE	TAVI	MCB	ANTH
Abilities & Oualifications should be used	32%	16%	35%	3%	14%
More categories necessary	25%	25%	31%	11%	8%
Classification not needed/ should be scrapped	14%	12%	32%	7%	35%
Term/Question should be changed	-	10%	45%	10%	35%

The comments on the Tavistock tended to focus on the use of the word 'coloured' for one of the categories. This term was felt to be offensive by many people, in particular Africans, for whom the term has very specific connotations. A lot of people said the term 'black' should be used, and others, especially people from the Far East, felt that more categories should be included.

For the Anthropological classification in particular there were many comments calling for its total removal. The terms were found offensive by many and were not understood by others.

Surprisingly 25% of the comments asking for more categories were on the OPCS classification which is the classification with most categories. Many of these comments, however, came from people from Africa who said that the category of 'African' could be broken down into individual countries; similarly, some people also felt that the category 'English, Welsh, Scottish, Irish' should be four categories and not one.

Few comments were made on the MCB classification, but most of those that were, referred either to increasing the number of categories or changing the terms. Some felt, for example, that it was not possible to be a British West Indian - you could only be British <u>or</u> West Indian, and so these terms ought to be separated.

A lot of miscellaneous comments were made which were specific to each classification. As on the OPCS classification some people objected to the fact that the Indian sub-continent categories were split up whereas there was only one African category; that Arabs can also be Africans; that Guyanese people are also West Indians. For the CRE classification it was said that although there was a category of Asian, this usually referred to people from the Indian sub-continent, so there should be a category for people originating from the Far East but who are also Asian. Finally, many people expressed the view, often in conversation, that even though they might object to answering questions of this sort, they would nevertheless answer if a prospective employer were to ask them, in order not to give a bad impression and thus lessen their job prospects.

5.1.3.3 An Examination of How Respondents Classify Themselves

If the classifications, and the categories within them, are clear and unambiguous then it should be expected that people of similar background and origins should classify themselves in the same way.

In order to examine this, people in the sample were grouped together according to the birthplaces of their parents. The hypothesis being tested was that people would classify themselves in a way which corresponded with their actual origins.

This method, however, assumes that the criterion used not only represents the origins of the respondents but also that of their parents, eg. if the parents were born in India they and their offspring would be assumed to be of Indian origin. This may not be true because even though the parents' birthplace could be India, their origins could be British since they may have been born in India whilst their parents were there on service from Britain. This did not appear to be a particularly prevalent problem, however, eg. every school-leaver in the sample was approached individually and none of them stated this was an issue.

The percentage degree of correspondence between individuals' self-classification and their origins as determined by their parents' place of birth on all of the classifications, is shown in Table 5.7. Here the individual groups shown in Table 5.1 have been combined into larger ones according to the general geographic location of their parents' birthplace.

Table 5.7 Percentage Degree of Correspondence Between National Origins (As Defined by Parents' Place of Birth) And Self-Ratings

	Standard Question %	Reversed Question %
Office of Population Censuses and Surveys' Classification		
UK	100	96.9
West Indies	41	70.7
Indian Subcontinent	95.7	78.7
Africa	100	93.5
Far East	100	91.3
Europe	50	50
Other	100	75
Middle East	87.5	37.5
Commission for Racial Equality's Classification		
UK	100	89.1
West Indies	39	46.3
Indian Subcontinent	91.5	97.5
Africa	96.8	80.6
Far East	100	95.7
Europe	10	20
Other	75	0
Middle East	87.5	25
Tavistock Institute's Classification		
UK	100	100
West Indies	97.6	97.6
Indian Subcontinent	91.5	95.7
Africa	87.1	87.1
Far East	91.3	87
Other	100	100
Middle East	87.5	87.5
Multi-Cultural British Classification		
UK	98.4	95.3
West Indies	56.1	73.2
Indian Subcontinent	83	97.9
Africa	93.6	81.8
Far East	100	95.7
Europe	40	50
Other	100	50
Middle East	87.5	87.5
Anthropological Classification		
UK	65.6	64.1
West Indies	70.7	73.2
Indian Subcontinent	40.4	38.3
Africa	48.4	48.4
Far East	34.9	47.8
Europe	70	70
Other	75	100
Middle East	50	50

The response to the 'standard' question tended to have higher degree of correspondence than the 'reversed' one for all groups except the West Indian where the opposite was true. The lower degrees of correspondence came from people born in England (and so classified themselves as English or an equivalent) but whose parents came from the West Indies, Indian sub-continent and Europe. Many of these people, however, but not all, when asked the reversed question put the origins as defined by their parents' birthplace. For example, on the OPCS classification 34.1% of the people whose parents originated from the West Indies classified themselves as English - on the reversed question though this fell to 19.5%. Furthermore, approximately the same proportion of this group classified themselves as 'African'.

There are a number of other points of interest which emerged from the self-rating. On the OPCS classification the category Arab proved confusing. It would appear that some people took Arab as meaning belonging to the Islamic world. Thus for some people from Africa and the Far East this was an acceptable category.

The same could be said of the Asian category in the CRE classification. This category included people who were from the Indian sub-continent, Far East and Middle-East. Thus if this classification was used for monitoring selection, differences in the selection rates of these groups would not be detected.

It should also be noted that there was a very high degree of non-response for the anthropological classification. The non-response for each group being 34.4% for the UK group, 26.8% for

West Indians; 57.4% for Asian, 54.8% for Africans, 39.1% for the Far East group, 42% for the European group and 50% for the Middle East group.

The reasons for this are perhaps revealed in the comments people made particularly those who said they found the terms offensive and so did not complete it and others who said they did not understand them and so could not complete it.

For the Tavistock classification some people from the Middle East had particular difficulty in determining whether they were coloured or white. Also some people, especially those from the West Indian and African groups, created a third category of 'black' which they placed themselves in. Overall, however, the Tavistock Institute's classification had the highest degree of correspondence. This was perhaps to be expected given that it has only two very broad categories.

However, it must be added that the results from the OPCS and MCB classifications were encouragingly high and, in terms of this particular criterion, there appears to be very little to choose between them.

5.1.4 Conclusions

1. The OPCS, CRE and MCB classifications were considered to be the most suitable and least objectionable of the classifications. Of the subsamples, however, the school leavers/nurses preferred the OPCS, MCB and CRE classifications, whereas the university students preferred the CRE and MCB classifications.

The OPCS, CRE and MCB classifications were also objected to least by each of the large ethnic groups.

2. The Anthropological and Tavistock classifications were seen as being significantly less suitable and more objectionable than the rest. The anthropological classification was also poorly understood.

3. The 'standard' question was considered more suitable than the 'reversed' one.

4. From the comments made by respondents, either in writing or verbally, there appeared to be little enthusiasm for monitoring generally although many people would answer such a question if an employer were to ask it.

5. There were particular interpretative problems encountered with people of West Indian origin who see themselves as African. There were also other interpretive problems which needed to be resolved.

5.2 Pilot Study 2 : Test Re-Test Reliability

5.2.1 Objective

This study was designed to examine the extent to which a particular sample's responses, to each of the classifications, remained the same when tested on two separate occasions. This is an important issue in monitoring because organisations will not

monitor their workforce on only one occasion. Instead, they must either repeat the exercise at regular intervals, or have a system of continuous monitoring, in order to ascertain the effectiveness of their equal opportunities policy. It is essential, therefore, that any differences in the make-up of an organisation's workforce should represent real changes and should not be due to the unreliability of the classification being used. Test-retest is a procedure which is commonly used to measure the reliability of selection tests. It has rarely, if ever, been applied to ethnic classifications.

5.2.2 Method

The sample chosen for this study was a post-graduate lecture group. The group was selected because:

(i) by testing and retesting a particular lecture group we could be sure that approximately the same people would be present on each occasion

(ii) this was one of the largest groups with the largest proportion of different ethnic minorities

(iii) it represented an analogous group in terms of educational qualifications to the Executive Officer applicants.

The format of the questionnaire was altered for this study (see Appendix 10). Since there was no need to ask the three opinion questions for this study they were removed. Also, for each classification two columns of boxes were presented and respondents were asked to place their response to the 'standard' question in the first column, and their response to the 'reversed' question in the second.

In order that each individual questionnaire could be identified, and so matched up for comparison when the sample was re-tested, a question on date of birth was also asked.

On the first occasion 67 questionnaires were distributed to the group and 52 were returned completed. They were not told that they were to be retested.

Four weeks later the same group were retested. On this occasion 39 people returned completed questionnaires. Of these, by using the date of birth information, it was possible to match 34 with their first set of questionnaires. The ethnic origins of these 34 people are shown in Table 5.8.

Table 5.8 Ethnic Composition of Test-Retest Sample

English-Welsh-Scottish-Irish	6
Nigerian	3
Indian	7
Pakistani	2
Iranian	4
Chinese	3
Other	9
	34

5.2.3 Results

The agreements between the test-retest scores are shown in Table 5.9. Generally, the 'reversed' questions had lower reliability than the 'standard' question. In most cases though this was because of non-response on one or other or both occasions.

Table 5.9 Percentage Test Re-Test Agreements

		STANDARD QUESTION	REVERSED QUESTION
OPCS	AGREE	94.1	79.4
	DISAGREE	2.95	5.9
	NON RESPONSE	2.95	14.9
CRE	AGREE	76.5	67.6
	DISAGREE	14.7	11.8
	NON RESPONSE	8.8	20.6
TAVI	AGREE	91.2	82.4
	DISAGREE	-	-
	NON RESPONSE	8.8	17.6
MCB	AGREE	76.5	76.5
	DISAGREE	17.6	11.75
	NON RESPONSE	5.9	11.75
ANTH	AGREE	76.5	73.5
	DISAGREE	-	-
	NON RESPONSE	23.5	26.5

Of the 'standard' questions the OPCS had the highest test-retest agreements of 94.1%. The Tavistock classification also had reliability of above 90%, but it had a greater number of non-responses.

The problems of using general categories were highlighted by the disagreements which occurred on the CRE and MCB classifications.

For the CRE classification the major areas of disagreement concerned South Americans who on one occasion described themselves in the 'other' category, but who on the other occasion described themselves as Caribbean .

For the MCB classification it was the inclusion of an 'Asian' category and a Chinese category which was problematic. Some people, for example, categorised themselves as Chinese on one occasion and Asian on the other. So rather than increasing the reliability, which might have been expected with such general categories, they tend instead to increase disagreements.

5.2.4 Conclusions

On the basis of these results it would appear that:

(i) the OPCS classification had the highest test-retest reliability, although the Tavistock and the MCB classifications also had high reliability

(ii) once again the highest degree of non-response was on the anthropological classification, thus confirming the result found in the first study

(iii) it was the general categories, eg. Arab in the OPCS, Caribbean in the CRE and Asian in the MCB which principally cause the unreliability found.

5.3 Overall Conclusions From The Pilot Studies

The main conclusion arising from these pilot studies was that further pilot work needed to be conducted on the classifications before they could be given to the Executive Officer candidates. The classifications to be used (and the improvements which need to be made to them) are discussed. The implications for monitoring of the 'reversed' question and the inter-rater reliability tests are discussed.

5.3.1 Classifications To Use In Further Pilot Work

The two classifications which emerged with the best results from the pilot studies were the OPCS and the MCB. Both classifications performed similarly on all of the criteria to justify further research on them in order to resolve the interpretive problems. Although the CRE classification was not as reliable or well-received as the OPCS and MCB ones, it was decided that it should be included in further pilot studies for two reasons. Firstly, modifications to it could improve its performance and secondly it is an important classification, being the one which the Commission for Racial Equality actually provide as an example for employers to use.

The Tavistock and Anthropological classifications, it was decided, should not be included in further pilot studies, and should not be used for monitoring generally. Many people found the terms of both classifications offensive, and for the Anthropological classification in particular the terms were also poorly understood.

Although the Tavistock classification performed well on some of the other criteria, eg. correspondence of self-ratings with national origin and test-retest agreements, it did not produce results which were any better than the OPCS, and occasionally the MCB, on any of the other criteria. Also both of these classifications, especially the Anthropological one, tended to have a higher rate of non-response than many of the others.

5.3.2 Resolving Interpretative Problems

The most important problem, on the OPCS, CRE and MCB classifications, was that of people of West Indian origin classifying themselves as African or English/British. Since many of the people who classify themselves in this way are Rastafarians the inclusion of such a category on the OPCS classification was considered worthwhile. In addition, the general category of African on the OPCS classification was removed and replaced by a more exhaustive list of specific African nations eg. Nigerian and Ghanaian .

Similarly the term Arab was replaced with specific Middle-East nations, eg. Iran, Iraq. This would hopefully overcome the dilemma of Africans who also see themselves as Arabs.

The MCB and CRE classifications are designed to be more general in nature than the OPCS, and in order to maintain this difference, other solutions were necessary.

Moreover, there were extra problems for these classifications in that some people of Chinese origin in classifying themselves, for

example, on the MCB classification could call themselves Chinese or Asian . It was felt that this problem as with all the other ones, in fact, could be reduced by providing more details and instructions as to what each category in each classification actually meant.

5.3.3 Use of the Reversed question

The use of the 'reversed' question alone was less reliable, and seen as less suitable and more objectionable than the use of the 'standard' question.

Bearing these points in mind it was also felt that further improvements to the classifications would make the use of the 'reversed' question redundant.

However, the value of the reversed question could still be in monitoring situations where decisions are made on the basis of subjective criteria, eg. in interview, and where the perceptions of the decision maker are the most important variables.

5.3.4 Summary of Conclusions

In summary the conclusions which were made on the basis of the three pilot studies were that:

1. the OPCS, CRE and MCB classifications produced the best results overall and would appear to have some usefulness in monitoring situations. However, certain interpretative problems remained, and these could only be resolved through further pilot work. Solutions to the problems included having detailed instructions on all of the classifications defining each of the

categories, and replacing vague general categories on the OPCS classification with more specific ones

2. the Tavistock and Anthropological classifications should be totally rejected for monitoring purposes

3. the 'reversed' question could be used in certain specific situations, eg. interviews.

CHAPTER 6

Comparison of Classifications For Use in Employment Selection: Stage II of Pilot Studies

As a result of the first set of pilot studies some classifications were thought to be of potential use in the Executive Officer study. Certain changes were felt to be necessary to them, however, and this chapter reports on the results of these refinements.

6.1 Choice of Classification

The purpose of this pilot study, as with the first set, was to ascertain which type of classification would provide the most reliable information whilst causing least objections.

The first pilot studies indicated little difference between three classifications - the OPCS, CRE and MCB - but that there were certain interpretive problems associated with each. In an attempt to overcome these, instructions were provided on each of the classifications stating, quite precisely, what was meant by each of the categories. So for example the British Asian category now stated "British Asian ie. persons born in Britain but who are descended from people who had Indian, Pakistani, Bangladeshi or Sri Lankan origins. This category is not intended to include Asians born in African countries".

On the OPCS classification the general categories of African and Arab were replaced by more specific ones eg. Nigerian, Ghanaian, Iranian and Iraqui. In addition a category of Rastafarian was included. This was done firstly because some people, of West Indian origin, who had classified themselves as African in the earlier studies were Rastafarian. Secondly, the removal of the term African meant that this group's primary term for self-classification had been removed. It was important, therefore, to replace it with a category which this group not only identified with but which would also identify their origins. In addition the instructions on the OPCS classification were changed from "Please tick the appropriate box to show the racial or ethnic group to which you belong" to "Please tick the appropriate box to show the racial or ethnic origins of your forebears".

Solutions to the interpretive problems encountered in the initial pilot studies have been produced by other organisations. The solution employed by the National Housing and Dwelling Survey (NHDS) was to have a classification similar to the OPCS one used in the earlier studies but replacing the British and European categories with the term 'white'. Below this are listed the non-white minority groups eg. West Indian, Pakistani. The NHDS classification was included in this second stage of pilot studies to determine the effectiveness of this method.

6.2 Objectives of the Study

The objectives of this pilot study were:-

(i) to assess the extent to which respondents objected to the four classifications

(ii) to assess which classification respondents felt was most suitable for monitoring purposes

(iii) to assess whether the refinements, made as a result of the first pilot studies, had increased the degree of correspondence between respondents' national origins and their replies to each classification.

6.3 Method

A questionnaire (see Appendix 11) was issued to 169 people. It included the three revised classifications, plus the NHDS. The revised classifications included instructions in order to clarify for respondents what each of the categories meant.

Once they had classified themselves, respondents then rated on a 5-point scale:-

(i) the extent to which they would object to each question if an employer were to ask it

(ii) the extent to which they thought each classification was suitable for monitoring purposes.

Space was provided beneath each of the classifications for any comments which respondents wanted to make. In addition, questions were also asked about their own and their parents' country of birth, their age, sex and educational qualifications.

The respondents comprised two sub-samples gathered from the same centres as the first pilot study, ie

(i) school leavers using the careers services in Coventry and Birmingham (n=59)

(ii) university students who provided the closest analogous group to EO applicants (n=110) (see Table 6.1).

Table 6.1 National Origins of Sample as Determined From Parents' Place of Birth

ORIGINS OF PARENTS OF RESPONDENTS	SCHOOL LEAVERS	UNIVERSITY STUDENTS	TOTAL
United Kingdom & Eire	17	25	42
West Indies	22	4	26
India	15	27	42
Pakistan	3	4	7
Sri Lanka	-	3	3
Nigeria	-	7	7
Sudan	-	2	2
Zambia	-	2	2
Zimbabwe	-	4	4
Kenya	-	2	2
Malaysia	-	12	12
Singapore	-	1	1
Hong Kong	1	8	9
Iraq	-	1	1
Iran	-	2	2
Egypt	-	1	1
Other European	1	1	2
Other Non-European	-	4	4
	59	110	169

6.4 Results

6.4.1 Degree of Objections

Table 6.2 Mean Scores for School Leavers, University Students and Total Groups on Degree of Objections

Classification	School Leavers	University Students	Total
CRE	2.39	3.27	2.96
OPCS	2.55	3.03	2.86
MCB	1.88	2.99	2.60
NHDS	2.87	3.32	3.16

The mean responses of each group in terms of their objections to each classification are shown in Table 6.2. A one-way analysis of variance with repeated measures was applied to the means of the total sample and the result was significant ($F=7.93$, $df=3,168$, $p<.01$). (For ANOVA summary table see Appendix 12a.) Tukey's multiple comparison of means (1949) was applied to the data and the only significant difference was between the MCB and NHDS classifications with the MCB classification being objected to less ($p<.01$).

The analysis of variance was also significant when applied to the school-leavers sample ($F=6.36$, $df=3,58$, $p<.01$) but not for the university sample ($F=1.63$, $df=3,109$, $p>.05$). (For ANOVA summary tables see Appendices 10b and 10c.) The school-leavers objected to the MCB classification significantly less than the NHDS one ($p<.05$).

The responses to the opinion questions of the three largest groups in the sample were examined, ie. United Kingdom and Eire, West Indian and Indian. (See Appendix 13 for full details.) The West Indian and Indian groups objected to the MCB classification the least. The UK group objected least to the OPCS classification. All groups objected most to the NHDS classification.

6.4.2 Degree of Suitability

Table 6.3 Mean Scores for School Leavers, University Students and Total Group on Degree of Suitability

Classification	School Leavers	University Students	Total
CRE	2.49	2.72	2.64
OPCS	2.42	2.63	2.56
MCB	2.32	2.66	2.54
NHDS	2.61	2.87	2.78

The mean responses of each group in terms of the suitability of the classifications are shown in Table 6.3. A one-way analysis of variance with repeated measures was applied to the means of the total sample and the result was not significant ($F=.69$, $df=3,168$, $p>.05$). (For ANOVA summary tables see Appendix 14a.) Non-significant results were also obtained when the analysis of variance was applied to the university sample ($F=1.59$, $df=3,109$, $p>.05$) and the school-leaver sample ($F=.75$, $df=3,58$, $p>.05$). (For ANOVA summary tables see Appendices 14b and 14c).

Examined by ethnic group, all groups considered the NHDS to be the least suitable. The MCB classification was considered the most suitable by the West Indian and Indian groups. The UK group considered the OPCS to be most suitable. (See Appendix 11 for complete table of means.)

6.5 Examination of Respondents' Comments

In all 112 comments were made, over one-third (37.5%) of them on the NHDS classification. Content analysis revealed that 91.1% of the comments could be placed into four broad categories which were similar to the four categories in pilot study 1.

These were that:

(i) people should be judged by abilities and qualifications (19 comments)

(ii) more categories should be included (9 comments)

(iii) the classification should be scrapped completely (34 comments)

(iv) alterations should be made to the terms used (40 comments)

Table 6.4 shows how these four categories of comments were distributed over the four classifications. Most of the comments calling for a particular classification to be scrapped, or for its terms of expression to be changed were made about the NHDS classification. In particular, comments were made about the use of the word 'white' in a classification which contained no other reference to colour. Some people felt that if this term was used then the only other category which would be appropriate was 'black'.

Table 6.4 Percentage of Each Category of Comments Made For Each Classification

Comment	CRE	OPCS	MCB	NHDS
Abilities and Oualifications should be used	21.1%	26.3%	26.3%	26.3%
More categories necessary	44.5%	11.1%	33.3%	11.1%
Classification not needed/ should be scrapped	17.7%	23.4%	17.7%	41.2%
Terms/Questions should be changed	17.5%	17.5%	20%	40%

A view was also expressed that the CRE and MCB classifications should contain more categories.

On the OPCS classification people objected to being asked the origins of their forebears, which was the way the question had been rephrased for this pilot study. Very few comments were made about including more categories on the OPCS classification, and only one person commented on the inclusion of the Rastafarian category.

Overall, then, the comments made were less specific than they were in the previous study, which could indicate that respondents objected to the content of the classifications (as opposed to the concept of monitoring) less than they did before. Nevertheless, using colour as one of the categories in a classification still caused most comment, and was objected to.

6.6 An Examination of How Respondents Classify Themselves

The provision of instructions for each of the classifications appears to have reduced the number of interpretive problems found in the first pilot study (see Table 6.5).

The number of minority group respondents classifying themselves as English or British, etc, was negligible.

Table 6.5 Percentage Degree Of Correspondence Between National Origins (as defined by Parents' Place of Birth) and Self-ratings

	% CORRESPONDENCE		% CORRESPONDENCE
CRE		MCB	
UK	97.6	UK	97.6
West Indies	73.1	West Indies	76.9
Indian subcontinent	100	Indian subcontinent	94.2
Africa	100	Africa	100
Far East	95.5	Far East	95.5
Other	62.5	Other	75.0
OPCS		NHDS	
UK	95.2	UK	97.6
West Indies	96.2	West Indies	76.9
Indian subcontinent	97.1	Indian subcontinent	91.3
Africa	84.1	Africa	84.1
Far East	95.5	Far East	100
Other	75.0	Other	62.5

However, about 25% of people of West Indian origin classified themselves as African on the CRE, MCB and NHDS classifications. This did not occur, however, with the OPCS classification (which did not contain a category of African , instead it mentioned the specific nationalities of Nigerian and Ghanaian , plus an additional category - that of Rastafarian). These changes appear to have been successful, and the degree of correspondence for the West Indian group was highest on the OPCS classification.

The relatively low degree of correspondence (76.5%) for the Indian subcontinent group on the MCB classification was due to African Asians (ie. people born in Africa but whose forebears were Indian, Pakistani, etc) classifying themselves as British Asian. If this is accepted as a valid response, since it does identify origins of forebears, then the degree of correspondence increases to 94.2%.

The highest degree of correspondence overall was obtained with the OPCS classification.

6.7 Conclusions

It was only possible to use a maximum of two classifications in the Executive Officer and the ones chosen were OPCS and MCB. They were objected to least by the total sample of school-leavers and university students and also for both classifications (and in particular the OPCS) there was a high percentage degree of correspondence between national origins and self-classification.

There were very little differences in fact between the MCB and the CRE classifications. On the whole the MCB classification was objected to less, but the differences were not significant. More importantly, it was felt that in a situation where differences in test scores are being examined, it would be more useful to have a classification which differentiates between people born in Britain and those born abroad to see whether there are any differences in their test performance.

In using OPCS and MCB classifications, however, further relatively minor modifications were made. In particular, all references to forebears in the OPCS instructions were removed, in response to the objections referring to this.

The use of instructions, however, explaining what each category of any classification meant, appeared to be a useful and practical way of increasing the effectiveness of the classifications. There were fewer interpretive problems in this study than in the ones previously. In this respect, it is interesting to note that very few, if any, of the classifications being used by organisations, for monitoring purposes, provide instructions as to the meaning and definition of the categories. This would appear to be helpful, a very straightforward way of improving the effectiveness and reliability of the classification without having to spend a lot of time attempting to obtain the 'correct' labels for each category.

Finally, as in the first set of pilot studies references to colour were objected to strongly by many people. It is recommended, therefore, that classifications used by organisations should not attempt to refer explicitly to the skin colour of any individuals.

CHAPTER 7

The Executive Officer Study: Method and Sample Details

A detailed investigation looking solely at the use of tests and their effect on ethnic minority selection has not occurred in the Civil Service before. However, such a study was carried out on the tests used in the American Civil Service and that work is reported here and its conclusions have relevance for this study. Details of the questionnaire development, sampling technique and of the sample itself are provided.

7.1 Relevant Research: Discriminatory Effects of Tests in US Civil Service

The Tavistock Institute's monitoring study (1978) within the Civil Service looked at various aspects of Civil Service recruitment, selection and promotion, but it did not focus in any detail on the use of tests. In the United States of America a study was carried out in 1976-77 by the General Accounting Office (GAO) which critically examined the use of tests and the selection of minority groups within the United States Civil Service. The tests considered were the Professional and Administrative Career Examination (PACE), Junior Federal Assistant, Accountant-Auditor and Social Worker Examinations. PACE covers entry into 118 positions - ones which recent college graduates are able to apply for, and can therefore be seen as the American equivalent of the Executive Officer examinations. This fact, together with the nature of the study and the conclusions which it reached, makes it an important one to be considered here.

In the United States, Civil Service recruitment is guided by a number of selection principles. These are that:-

(i) competitive examinations are open to everyone

(ii) examinations based upon testing and other applicant appraisal procedures do not discriminate on the basis of race, sex, religion, national origins or any other factor not directly related to the requirements of the job or jobs involved

(iii) equal opportunities exist for everyone examined, as long as the minimum educational requirements have been met

(iv) selection is from amongst the best qualified (GAO, 1979)

The selection procedure, therefore, combines meritocratic principles with those of equal opportunity. The GAO wanted to know if the examination procedures were achieving the stated policy objective. Secondly, they wished to establish the validity of the tests.

The sample comprised 167,618 people who took the PACE in 1976. There is a pass mark for the test, but to have a realistic chance of obtaining employment a score of at least 90 is needed.

Candidates identified themselves on a classification with three categories: white, Negro or other/unspecified.

58% of the white applicants passed PACE, with 16% scoring above 90. (The statistical significance of the results is not reported.)

56% of 'other' racial minority applicants passed, with 14% scoring 90 or above.

12% of black people passed, but less than 1% with a score of 90 or above.

Black people performed better on the Junior Federal Assistant test than on PACE. However, the proportions of black people who obtained a score which would give them a realistic chance of employment was still lower than the white group - 15% of white people compared to 5% of black people.

On the Accountant-Auditor and Social Worker Examinations of those who passed proportionately as many black people as white people scored in the 'employable' range.

The Office of Personnel Management (OPM), responsible for the development and administration of the examination programmes, were criticised for not monitoring the selection procedure. "As a result, the Office of Personnel Management does not know what impact the selection procedure components have on minorities, nor can it determine whether adverse impact is present based on the total selection process." (GAO 1979).

The OPM stated they had not introduced monitoring because it was felt to be expensive, but also they had not agreed upon the forms, nor had they devised a classification system which could be used in different selection situations, ie. tests, interviews, etc.

Since the analysis showed that certain tests screened out a disproportionately large number of black applicants, it was asked whether a clear relationship existed between test scores and job performance. PACE is used to screen people for 118 Civil Service occupations, and after examining the validation documentation the tests were criticised by the GAO (1979) on the following grounds:

(i) in developing the tests neither the OPM nor its predecessors had followed the Uniform Guidelines (1978) (a set of guidelines developed by a number of legislative and governmental bodies) scrupulously enough

(ii) although used for 118 occupations, detailed job analysis had been carried out on only 27 occupations

(iii) criterion-related studies were carried out in only three of the 12-15 occupations where they might have been possible. Furthermore, these were concurrent rather than predictive validity studies

(iv) race and ethnicity data were not collected, nor did the researchers ensure that individuals in the study were representative of the relevant labour market

(v) possible test bias or test unfairness was not investigated, although this is required by the Uniform Guidelines (1978)

(vi) suitable alternatives to the test or alternative ways of using the test were not sufficiently investigated. Some alternatives were explored by the OPM, but the GAO "...believes more can and should be done to investigate suitable alternatives which have less adverse impact" (1979)

Finally, the GAO (1979) recommended that the OPM should

(i) increase their efforts to comply with the Uniform Guidelines requirement for collecting and maintaining ethnic records

(ii) re-evaluate the PACE validation strategy to ensure it conforms to the Uniform Guidelines (1978)

Also, if further validation studies were to be carried out they should:

(i) reconsider their validation design

(ii) ensure as far as possible that research participants are representative of the relevant labour market

(iii) investigate the fairness of the test for race, sex and ethnic subsamples.

The weakness of the PACE study is that other than the crude ethnic classification and test scores very little other information was available. In the present study there is more information which enables not only a more detailed examination of pass/fail ratios and test scores, but also of possibly identifying factors which might be related to them.

7.2 Development of the Questionnaire

For the Executive Officer study a questionnaire was developed which contained two classifications, the OPCS and MCB, together with questions to obtain other background information. (For complete questionnaire see Appendix 15.) Following the results of the third pilot study the OPCS and MCB classifications were used in this study.

In addition, candidates were asked to provide

(i) details of the country of their birth and, if it was not within the United Kingdom, the number of years they had lived here

(ii) details of the country or countries where they received their primary, secondary and (if any) further education

(iii) their test index number. Each candidate sitting the EO tests is given an index number by the Civil Service which they write on each of their test answer sheets.

The candidates sex, age and educational qualifications were also recorded.

In order to determine which classification candidates preferred they were asked to indicate the extent to which they would have approved or objected to each of the classifications had they been used as part of the Civil Service selection procedure.

7.3 Administration of Questionnaire

Executive Officer applicants are invited by the Civil Service to take the EO selection tests, and to attend a testing session at a particular time and place. The testing sessions last, on average, three hours.

After the last test had been completed the Civil Service test administrator would introduce the person who was to administer the questionnaire (either the researcher, or a member of Chartamede staff). The candidates were asked to follow the instructions on the questionnaire while they were read aloud. It was made

clear to them that the questionnaire did not form part of the selection procedure and that the information would be treated with the utmost confidentiality. The whole procedure usually took less than ten minutes.

7.4 Sampling Procedure and Sample Details

The choice of venues and testing centres from which the sample was to be drawn was an important issue. Obviously a large sample was needed, but it was as important that it contained a high proportion of minority group applicants.

Approximately 20,000 people per year sit the EO selection tests. Unfortunately no figures were available on the attendance of minority groups. It was necessary to guess at those regions where minority group representation would be reasonably high which almost inevitably meant that the chosen centres were urban. These were London, Birmingham, Cardiff, Bristol, Reading, Nottingham, Liverpool, Sheffield and Leicester.

It was estimated that by concentrating on these areas, particularly on London, 10% or more of our total sample might be minority group applicants. It was planned that approximately 4,000 applicants would be given the questionnaire. During the pilot studies the response rate was between 60-70%. It was predicted that response rate in these situations would be higher - approximately 75%. Therefore the total sample was expected to be approximately 3,000 with 300 minority group applicants.

Data gathering began on 1 March 1982 and ended on 1 June 1982. Some testing centres were visited more than once and, in all, 59 testing sessions were attended, 46 of them in London.

7.4.1 Overall Response Rate

In total, 3,755 people were given the questionnaire. 3.94% of people did not complete any part of the questionnaire, and 0.19% gave their index number but did not complete the classifications. Overall, 95.87% of the sample provided information on their origins. The index number, however, was essential in order to match up candidates' information on the questionnaires with their test scores (in order to assure confidentiality candidates' names were not recorded). 3,449 candidates provided their index numbers, plus the additional information, making a response rate of 91.88%, which exceeded the expected response rate.

There are two possible reasons for this very high, and unexpected, response rate. Firstly, and obviously, many candidates, although told the exercise was voluntary, probably felt 'compelled' to help just by the very nature of the situation. Secondly, however, the response rate might have been much lower had not much care been taken to make the classifications as acceptable as possible.

The total number of respondents from the London area alone was 2,851 - 82.63% of the total. 599 (17.37% of the total sample) completed questionnaires were obtained from the other testing centres.

The numbers of minority group members (ie. people who had ticked one of the minority group categories on either the OPCS or MCB classifications) is shown in Table 7.1. Most minority group members (442 in all) came from London which represented 15.5% of the people tested in London.

There were only 32 minority group members in all of the other centres, which represents 5.34% of the sample in these regions.

Table 7.1 Minority Group Representation in Sample

Testing centre	Completed questionnaires	No. of minorities	Minorities as %age
London	2,851	442	15.50
Birmingham	265	23	8.68
Cardiff	27	2	7.41
Bristol	84	0	0.00
Nottingham	35	3	8.57
Liverpool	83	0	0.00
Sheffield	41	1	2.44
Leicester	24	2	8.33
Reading	39	1	2.56

7.5 Comparison of OPCS and MCB Classifications

3,449 people completed at least one classification and provided their index number. Of these very marginally more completed the OPCS classification - 3,405 versus 3,400. The detailed breakdown for each classification is shown in Table 7.2(a) and 7.2(b).

The MCB classification has larger numbers of people overall in its categories - but this was to be expected since the categories are more general. Only one category, that of 'British Chinese', has less than ten people in it, whereas five OPCS categories fall below that figure.

Correspondingly, there are fewer people in the 'Other' categories on the MCB - 1.8% compared with 2.6% of the OPCS sample.

Table 7.2(a) Numbers of People In Each Category of OPCS Classification

CATEGORY	No.	%
English-Welsh-Scottish-Irish	2925	85.9
Italian	20	0.6
Greek Cypriot	23	0.7
Turkish Cypriot	7	0.2
Polish	14	0.4
West Indian	85	2.5
Rastafarian	2	0.1
Nigerian	6	0.2
Ghanaian	47	1.4
Indian	149	4.4
Pakistani	15	0.4
Bangla-Deshi	3	0.1
Iranian	1	0.0
Chinese	17	0.5
Other	91	2.6

Table 7.2(b) Numbers of People In Each Category of MCB Classification

CATEGORY	No.	%
British	2976	87.5
British Asian	39	1.2
British West Indian	52	1.5
British Chinese	6	0.2
Other British	51	1.5
African Asian	54	1.6
Asian	95	2.8
West Indian	32	0.9
African	29	0.9
European	42	1.2
Chinese	14	0.4
Other	10	0.3

One of the problems which is being tackled with these classifications is increasing the reliability and consistency of how people categorise themselves. As was seen in the pilot studies people categorising themselves as, for example, 'West Indian', on one classification might categorise themselves as 'British' on another. If the classifications are well-understood and completed earnestly then it should be possible to look at a person's self-categorisation on one classification (eg. the OPCS) and estimate that individual's categorisation on a completely different classification (in this case the MCB). For example, the British category on the MCB and the 'English-Welsh-Scottish-Irish' (E-W-S-I) category on the OPCS are, by the definitions provided on the questionnaires, identical, so one would hope to find very similar numbers in both. This is in fact the case, although the MCB 'British' category contains 51 more people, an increase of 1.7%. Of these 51, there were 18 people who did not classify themselves on the OPCS system. In the pilot studies West Indians were most likely to classify themselves as British. On this occasion, however, only two West Indians on the OPCS classified themselves as 'British' on the MCB. Instead, 28 of the 33 who had classified themselves as being part of a minority group on the OPCS, but as 'British' on the MCB, were from European groups, ie. 'Italian', 'Polish', 'Turkish Cypriot', 'Greek Cypriot', etc. The most obvious reason for this is the lack of a 'British European' category on the MCB classification which perhaps might have to be included. The modifications made to the classifications in other respects, however, appear to have been successful. For example, of the 87 'West Indians' and 'Rastafarians' on the OPCS only two categorised themselves as 'British' on the MCB.

Overall, the response rate was very high and the numbers of minority group members were satisfactory, if a little low in some of the regions. Judging by the consistency of categorisation across the classifications, the modifications made as a result of the pilot studies appear to have worked, although there may be an argument in favour of the inclusion of a 'British European' category on the MCB.

7.6 Objectives of the Executive Officer Study

The objectives of this study were:-

(i) to examine any significant differences between the groups, in the OPCS and MCB classifications, in

(a) their pass/fail rates and adverse impact

(b) their overall test scores

(c) the three individual tests. Previous research (see Chapter 3) has shown that minority groups perform less well on verbal rather than non-verbal tests. (Jensen 1980, Alleyne 1962, Hegarty and Lucas 1978, Little et al 1968). It is hypothesised that the greatest differences will be found on the test with the greatest degree of verbal material (The Executive Problems Test) and least differences on the numerical test (Enquiries from Management).

(ii) to determine which factors are related to:

(a) passing/failing the the tests and

(b) test performance.

Anastasi (1975) (see Chapter 3) states that tests are inevitably culture-bound and that they should reveal differences between groups. Greater exposure to a particular culture,

therefore, may be reflected in the test scores obtained and research appears to support this view. (For example, minority group children born in the United Kingdom performed better than those born abroad (Yule et al 1975), and of those children born abroad those who entered the UK before the age of five years performed best (McKie and Thomson 1970). However, test performance of immigrant children increases relatively with length of schooling (Watson 1973). It is hypothesised here, therefore, that on the MCB classification that the British-born minority groups should perform better on the tests and have higher pass rates than their equivalent 'overseas-born' minority group. Furthermore, it is hypothesised that test performance will have a negative correlation with the age at which a person entered the UK, and a positive one with number of years spent in the UK

(iii) to examine any significant differences, for each individual group on the OPCS and MCB classifications, in their preference for/objections to the two classifications.

CHAPTER 8

Results: Examination and Comparison of Pass Rates Between Groups on the Two Classifications

This chapter considers the first of the objectives mentioned in Chapter 7.6 ie. the examination of any significant differences between the groups in the OPCS and MCB classifications in terms of their pass/fail rates.

The pass rates for each group on the OPCS classification are considered first, followed by those on the MCB classification. In both cases a X^2 (chi-square) test for independent samples was applied to the data followed by a post hoc comparison of pass rates.

Finally, adverse impact was examined using the 'rule of thumb' used in American courts (ie. adverse impact is considered to exist when one group's rate is only four fifths or less that of the majority group).

8.1 The Scoring of the Tests

There are three tests in the Executive Officer Qualifying Test Battery - Executive Problems, Enquiries from Management and Intelligence, each of which is scored out of different totals in raw score form eg. Enquiries from Management is scored out of a total of 25, whereas Executive Problems is scored out of 20.

The tests are standardised by comparing each candidate's performance, on each of the three tests, with those obtained by the previous year's applicants (the standardisation sample). There are in fact 14 separate standardisation samples which represent seven age categories (ie. 17-20 years of age, 21-25 years, 26-30 years, 31-35 years, 36-40 years, 41-45 years and those above the age of 45) each of which is split into male and female categories.

The Enquiries from Management and Executive Problems test raw scores are converted to standard scores, giving them a mean of 50 and a standard deviation of 10. The Intelligence test scores are standardised to have a mean of 100 and a standard deviation of 25.

In determining whether a candidate has passed the test the standardised scores from the three tests are added together to provide an overall weighted composite, or total, score which has a mean of 200. (By totalling the individual tests in this way the Intelligence test is weighted two and a half times more highly than the other two.) The pass mark is set at 210 and applicants who score above that <u>and</u> who meet the eligibility requirements are sent an application form and are invited to interview.

8.2 Examination of Overall Pass Rates

An individual's total score is calculated by adding up the standardisation scores from the three tests. If the overall total score is 210 or above and the individual meets the other selection requirements, he or she is called to be interviewed. The pass rates for groups on the OPCS classification are considered below, followed by the MCB classification.

8.2.1 Pass Rates for Each Category on OPCS (see Table 8.1)

The overall pass rates for each of the OPCS categories can be seen in Table 8.1.

Table 8.1

Pass Rates for Each OPCS Category

CATEGORY	No. Pass (Total)	%
English-Welsh-Scottish-Irish	998 (2925)	34.1
Italian	3 (20)	15.0
Greek-Cypriot	1 (23)	4.3
Turkish-Cypriot	0 (7)	0.0
Polish	5 (14)	35.7
West Indian	8 (87)	9.2
Nigerian	0 (6)	0.0
Ghanaian	12 (47)	25.5
Indian	12 (149)	8.1
Pakistani	1 (15)	6.7
Bangladeshi	0 (3)	0.0
Chinese	1 (17)	5.9
Other	25 (92)	27.2

A X^2 (chi-square) test for independent samples was applied to the data. (In order to meet the assumptions of the X^2 test the categories of 'Turkish-Cypriot', 'Nigerian' and 'Bangladeshi' were amalgamated with the category of 'Other'.) The test indicated that, overall, there were differential pass/fail rates for different categories (X^2=94.22, df=10, $p<.01$).

The highest pass rates were those of Polish (35.7%) and 'English-Welsh-Scottish-Irish' (34.1%). A post hoc comparison of pass rates was applied to the data (Marascuillo and McSweeney 1979). The 'E-W-S-I' group pass rate was significantly higher than those of 'Greek Cypriots' ($p<.01$), 'West Indians' ($p<.01$), 'Indians' ($p<.01$) and 'Chinese' ($p<.05$). There were no other significant differences. (For complete table see Appendix 16).

8.2.2 Pass Rates for Each Category on the MCB (see Table 8.2)

The overall pass rates for each of the MCB categories can be seen in Table 8.2. (The category of 'Chinese' was amalgamated with the 'Other' category in order to meet the assumptions of the X^2 test.) The X^2test applied to the data was significant (X^2=117.51, df=11, $p<.01$) indicating that there were differential pass rates for the different groups.

Table 8.2

Pass Rates for Each MCB Category

CATEGORY	No. Pass (Total)	%
British	1006 (2976)	33.8
British Asian	7 (39)	17.9
British West Indian	7 (52)	13.5
British Chinese	1 (6)	16.7
Other British	12 (51)	23.5
African Asian	6 (54)	11.1
Asian	6 (95)	6.3
West Indian	1 (32)	3.1
African	1 (29)	3.4
European	11 (42)	26.2
Chinese	0 (14)	0.0
Other	6 (10)	60.0

A post hoc comparison of pass rates (Marascuillo and McSweeney 1979) revealed that the 'British' pass rate was significantly higher than those of 'African Asian', 'Asian', 'West Indian' and 'African'. (For complete table see Appendix 17.)

One point of interest here is that there is a tendency for the pass rates of the British minority groups to be higher than their corresponding foreign-born groups. None of these differences is significant, however.

On comparing the two classifications it can be seen that although there is a significant difference between the 'British' and 'African' pass rates, there are no significant differences between the 'E-W-S-I' group and the individual African national groupings on the OPCS.

8.2.3 Pass Rates and Adverse Impact

Adverse impact is described as the disproportionate rejection rate of one subgroup by comparison with the rest of the individuals being assessed for selection (Runnymede Trust 1978).

In the United States the courts use a 'rule of thumb' to decide whether the issue of adverse impact should be considered viz when one group's selection rate is only four-fifths or less that of the majority group. On the OPCS classification the four-fifths rate is 27.28% and only the categories of 'Polish' and 'Other' achieve it.

On the MCB classification none of the groups achieve the four-fifths selection rate of the British group.

It would appear, therefore, that, if the American rules of thumb are applied, that the Civil Service Executive Officer selection tests are showing considerable adverse impact against many minority groups, and that the issue of indirect discrimination is legitimately raised.

8.3 Summary of Results on Pass/Fail Rates

The results obtained from an examination of the pass/fail rates of the various groups can be summarised as follows:

(i) on the OPCS classification the E-W-S-I pass rate is significantly higher than those of 'Greek-Cypriot', 'West Indian', 'Indian' and 'Chinese'. On the MCB classification the 'British' pass rate was significantly higher than those of 'African Asian', 'Asian', 'West Indian' and 'African'. There were no significant differences between the pass rates of the minority groups, nor between the 'British' group and the British-born minorities.

(ii) when the pass rates are examined in relation to the four-fifths rule used by American courts there appears to be considerable adverse impact against minority groups.

The factors which are related to the pass rates have not been examined here. These are considered in more detail in Chapter 10.

CHAPTER 9

Results: Examination and Comparison of Test Scores Between Groups on the Two Classifications

This chapter examines objectives i (b) and i (c) mentioned in Chapter 7.6 ie. to examine any significant differences in overall total test score and individual test scores between the groups in the OPCS and MCB classifications.

The total test score is considered first followed by an examination of the Intelligence Test, Enquiries from Management Test and the Executive Problems Test.

For the total test score and for each individual test the results are presented for the OPCS groups first and then the MCB groups. In each case a one-way analysis of variance was applied to the data, followed by a multiple comparison of means using the modified Least Significant Difference (LSD) procedure (Nie et al 1975).

As with Chapter 8, no attempt is made at this stage to determine which factors are related to test performance. This analysis is performed in Chapter 10.

9.1 Comparison of Total Scores

9.1.1 Comparison of Total Scores For OPCS Categories

Table 9.1 provides the mean and standard deviation for each OPCS group. A one-way analysis of variance was applied to the

data and was significant ($F=23.56$, $df=12,3392$, $p<.01$). (For analysis of variance tables see Appendix 18.)

Table 9.1 Overall Total Mean Score for Each OPCS Category

CATEGORY	MEAN (X)	STANDARD DEVIATION (S.D.)
English-Welsh-Scottish-Irish	198.96	35.16
Italian	193.85	35.17
Polish	193.06	34.22
Other	187.78	45.34
Ghanaian	186.51	42.33
Chinese	173.76	43.95
Bangladeshi	169.67	37.10
Greek-Cypriot	164.35	44.88
Nigerian	163.47	28.49
Turkish-Cypriot	161.00	48.50
Indian	158.31	39.44
West Indian	158.17	38.56
Pakistani	157.40	38.54

In order to discover the significant differences between the mean scores the modified Least Significant Difference (LSD) multiple comparison of means was applied (Nie et al 1975) and the results are shown in Table 9.2. The 'E-W-S-I' mean score was significantly higher than those of the 'Greek-Cypriot', 'Indian', 'West Indian' and 'Pakistani' groups. The mean scores of the 'Italian', 'Other' and 'Ghanaian' groups were significantly higher than those of the 'Indian' and 'West Indian' groups.

Table 9.2 Significant Differences Between OPCS Groups For Mean Total Score

	Italian	Polish	Other	Ghanaian	Chinese	Bangladeshi	Greek-Cypriot	Nigerian	Turkish-Cypriot	Indian	West Indian	Pakistani
English-Welsh Scottish-Irish	NS	NS	NS	NS	NS	NS	**	NS	NS	**	**	**
Italian		NS	NS	NS	NS	NS	NS	NS	NS	*	*	NS
Polish			NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
Other				NS	NS	NS	NS	NS	NS	**	**	NS
Ghanaian					NS	NS	NS	NS	NS	**	**	NS
Chinese						NS	NS	NS	NS	NS	NS	NS
Bangladeshi							NS	NS	NS	NS	NS	NS
Greek-Cypriot								NS	NS	NS	NS	NS
Nigerian									NS	NS	NS	NS
Turkish-Cypriot										NS	NS	NS
Indian											NS	NS
West Indian												NS
Pakistani												

** $p < .01$

* $p < .05$

NS Not Significant

9.1.2 Comparison of Total Scores For MCB Categories

The means and standard deviations for the total score for each MCB category can be found in Table 9.3. The one-way analysis of variance applied to this data was significant ($F=23.75$, $df=11,3388$, $p<.01$) (For analysis of variance summary table see Appendix 19.) Table 9.4 shows the significant differences between the individual means.

Table 9.3 Overall Total Mean Score For Each MCB Category

CATEGORY	MEAN (X)	STANDARD DEVIATION (S.D.)
British	199.20	35.32
Other British	194.78	39.08
Other	192.20	39.27
British Chinese	186.00	59.46
European	182.95	47.74
Chinese	169.71	30.50
African Asian	168.43	35.78
British-Asian	164.33	45.96
British West Indian	162.31	40.87
Asian	158.67	35.78
West Indian	153.12	37.54
African	152.59	38.77

Table 9.4 Significant Differences Between MCB Groups For Mean Total Score

	Other British	Other	British Chinese	European	Chinese	African-Asian	British-Asian	British West Indian	Asian	West Indian	African
British	NS	NS	NS	NS	NS	**	**	**	**	**	**
Other British		NS	NS	NS	NS	NS	NS	*	**	**	**
Other			NS	NS	NS	NS	NS	NS	NS	NS	NS
British Chinese				NS	NS	NS	NS	NS	NS	NS	NS
European					NS	NS	NS	NS	NS	NS	NS
Chinese						NS	NS	NS	NS	NS	NS
African-Asian							NS	NS	NS	NS	NS
British-Asian								NS	NS	NS	NS
British West Indian									NS	NS	NS
Asian										NS	NS
West Indian											NS
African											

** $p < .01$ NS Not Significant

* $p < .05$

The 'British' group mean was significantly higher than those of the 'African Asian', 'British Asian', 'British West Indian', 'Asian', 'West Indian' and 'African' groups.

The 'Other British' group mean was significantly higher than the 'British West Indian', 'Asian', 'West Indian' and 'African' groups.

9.2 Comparison of Intelligence Test Scores

9.2.1 Comparison of Intelligence Test Scores For OPCS Categories

The standardised scores for this test are calculated so that they have a mean of 100 and a standard deviation of 25. Table 9.5 presents the means and standard deviations for each of the OPCS categories on this test.

Table 9.5 Intelligence Test Mean Scores For Each OPCS Category

CATEGORY	MEAN (X)	STANDARD DEVIATION (S.D.)
English-Welsh Scottish-Irish	98.42	26.77
Italian	94.30	21.09
Polish	93.71	20.76
Other	91.78	25.65
Ghanaian	90.38	25.24
Chinese	82.82	20.36
Bangladeshi	79.00	23.90
Turkish-Cypriot	78.29	28.70
Nigerian	75.97	14.66
West Indian	75.95	21.71
Greek-Cypriot	75.27	20.88
Indian	72.27	21.29
Pakistani	71.27	20.92

A one-way analysis of variance on this data was significant ($F=17.61$, $df=12,3392$, $p<.01$) (For ANOVA summary table see Appendix 20.) Table 9.6 shows the significant differences between mean scores. The 'E-W-S-I' group mean was significantly higher than the means of 'Greek-Cypriot', 'West Indian', 'Indian' and 'Pakistani' groups. The only other significant differences were between the 'Other' group mean and those of 'West Indian' and 'Indian' groups.

The pattern of significant differences for the 'E-W-S-I' and 'Other' groups is very similar to that obtained from the analysis of total scores. This is perhaps to be expected, since this test is weighted two and a half times more highly than the other two when calculating the total score.

Table 9.6 Significant Differences Between OPCS Groups For Intelligence Test Score

	Italian	Polish	Other	Ghanaian	Chinese	Bangladeshi	Turkish-Cypriot	Nigerian	West Indian	Greek-Cypriot	Indian	Pakistani
English-Welsh-Scottish-Irish	NS	NS	NS	NS	NS	NS	NS	NS	**	**	**	*
Italian		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
Polish			NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
Other				NS	NS	NS	NS	NS	*	NS	**	NS
Ghanaian					NS	NS	NS	NS	NS	NS	NS	NS
Chinese						NS	NS	NS	NS	NS	NS	NS
Bangladeshi							NS	NS	NS	NS	NS	NS
Turkish-Cypriot								NS	NS	NS	NS	NS
Nigerian									NS	NS	NS	NS
West Indian										NS	NS	NS
Greek-Cypriot											NS	NS
Indian												NS
Pakistani												

** $p<.01$ * $p<.05$ NS Not Significant

9.2.2 Comparison of Intelligence Test Scores For MCB Categories

Table 9.7 presents the means and standard deviations for the MCB categories on the Intelligence Test.

Table 9.7 Intelligence Test Mean Scores For Each MCB Category

CATEGORY	MEAN (X)	STANDARD DEVIATION (S.D.)
British	98.39	26.78
Other	98.20	27.70
Other British	94.57	22.09
British Chinese	91.50	28.20
European	87.88	26.28
Chinese	80.21	12.64
British West Indian	77.77	22.86
British Asian	77.31	25.88
African Asian	77.26	19.68
West Indian	73.72	22.43
Asian	72.27	19.99
African	71.38	21.76

The one-way analysis of variance was significant ($F=17.52$), $df=11,3388$, $p<.01$) (For ANOVA summary table see Appendix 21.) Table 9.8 shows the significant differences between group means.

The 'British' group mean was significantly higher than those of 'British Asian', 'British West Indian', 'African Asian', 'Asian' 'West Indian' and 'African' group. The only other significant difference was between the 'Other British' and 'Asian' means.

Compared with their overall test score the British minority groups all performed relatively better on this test, apart from the 'British West Indian' group.

Table 9.8 Significant Differences Between MCB Groups For Intelligence Test Score

	Other	Other British	British Chinese	European	Chinese	British West Indian	British Asian	African Asian	West Indian	Asian	African
British	NS	NS	NS	NS	NS	**	**	**	**	**	**
Other		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
Other British			NS	NS	NS	NS	NS	NS	NS	*	NS
British Chinese				NS	NS	NS	NS	NS	NS	NS	NS
European					NS	NS	NS	NS	NS	NS	NS
Chinese						NS	NS	NS	NS	NS	NS
British West Indian							NS	NS	NS	NS	NS
British Asian								NS	NS	NS	NS
African Asian									NS	NS	NS
West Indian										NS	NS
Asian											NS
African											

** $p < .01$
* $p < .05$
NS Not Significant

9.3 Comparison of Enquiries From Management Test Scores

9.3.1 Comparison of Enquiries From Management Test Scores For OPCS Categories

The standardised scores for this test have been calculated to have a mean of 50 with a standard deviation of 10. Table 9.9 shows the means and standard deviations for each OPCS category.

Table 9.9 Enquiries From Management Test Mean Scores For Each OPCS Category

CATEGORY	MEAN (X)	STANDARD DEVIATION (S.D.)
Polish	51.35	10.99
English-Welsh-Scottish-Irish	50.76	9.96
Italian	50.40	9.19
Ghanaian	49.98	10.25
Other	48.37	11.03
Nigerian	47.50	8.17
Bangladeshi	47.00	7.81
Chinese	46.59	11.95
Greek-Cypriot	44.88	12.63
Indian	44.34	10.58
Pakistani	44.20	9.04
West Indian	42.83	9.80
Turkish-Cypriot	40.71	10.58

The one-way analysis of variance was significant ($F=11.38$, $df=12,3392$, $p<.01$) (For summary table see Appendix 22.) Table 9.10 shows the significant differences between the mean scores. There were only three significant differences: the majority group being higher than those of the 'Greek-Cypriot', 'West Indian' and 'Indian' groups.

Table 9.10 Significant Differences Between OPCS Groups For Enquiries From Management Test Score

	English-Welsh Scottish-Irish	Italian	Ghanaian	Other	Nigerian	Bangladeshi	Chinese	Pakistani	Greek-Cypriot	Indian	West Indian	Turkish-Cypriot
Polish	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
English-Welsh-Scottish-Irish		NS	NS	NS	NS	NS	NS	NS	*	**	**	NS
Italian			NS	NS	ND	NS	NS	NS	NS	NS	NS	NS
Ghanaian				NS	NS	NS	NS	NS	NS	NS	NS	NS
Other					NS	NS	NS	NS	NS	NS	NS	NS
Nigerian						NS	NS	NS	NS	NS	NS	NS
Bangladeshi							NS	NS	NS	NS	NS	NS
Chinese								NS	NS	NS	NS	NS
Pakistani									NS	NS	NS	NS
Greek-Cypriot										NS	NS	NS
Indian											NS	NS
West Indian												NS
Turkish-Cypriot												

** $P < .01$

* $P < .05$

NS Not Significant

9.3.2 Comparison of Enquiries From Management Test Scores For MCB Categories

Table 9.11 presents the means and standard deviations for each MCB group.

Table 9.11 Enquiries From Management Test Mean Scores For Each MCB Category

CATEGORY	MEAN (X)	STANDARD DEVIATION (S.D.)
British	50.90	9.96
Other British	50.41	10.13
Other	48.50	12.12
European	48.38	12.12
Chinese	47.29	8.55
British Chinese	46.33	16.17
African Asian	45.39	9.60
Asian	44.96	11.42
British Asian	44.15	10.31
African	43.38	9.52
British West Indian	43.02	10.15
West Indian	41.59	9.54

The one-way analysis of variance was significant ($F=11.71$, $df=11,3388$, $p<.01$) (see Appendix 23 for ANOVA summary table.)

Table 9.12 shows the significant differences between the mean scores. The 'British' mean was significantly higher than those of 'British Asian', 'British West Indian', 'Asian', 'West Indian' 'African' and 'African Asian' groups.

Again, the pattern of significant differences is similar to that seen with other analyses. However, the range of mean scores was far closer for this test, and the differences between British-born minority groups are smaller than for any of the other tests.

Table 9.12 Significant Differences Between MCB Groups For Enquiries From Management Test Score

	Other British	Other	European	Chinese	British Chinese	African Asian	Asian	British Asian	African	British West Indian	West Indian
British	NS	NS	NS	NS	NS	*	**	**	**	**	**
Other British		NS	NS	NS	NS	NS	NS	NS	NS	NS	*
Other			NS	NS	NS	NS	NS	NS	NS	NS	NS
European				NS	NS	NS	NS	NS	NS	NS	NS
Chinese					NS	NS	NS	NS	NS	NS	NS
British Chinese						NS	NS	NS	NS	NS	NS
African Asian							NS	NS	NS	NS	NS
Asian								NS	NS	NS	NS
British Asian									NS	NS	NS
African										NS	NS
British West Indian											NS
West Indian											

** $p < .01$ NS Not Significant
* $p < .05$

9.4 Comparison of Executive Problems Test Score

9.4.1 Comparison of Executive Problems Test Scores For OPCS Categories

As with the Enquiries From Management test, Executive Problems scores have been standardised so that they have a mean of 50 and a standard deviation of 10. The means and standard deviations for the OPCS categories are presented in Table 9.13.

Table 9.13 Executive Problems Test Mean Scores For Each OPCS Category

CATEGORY	MEAN (X)	STANDARD DEVIATION (S.D.)
English-Welsh-Scottish-Irish	49.78	10.34
Italian	49.15	9.66
Polish	48.00	9.31
Other	47.63	11.89
Ghanaian	46.15	11.17
Chinese	44.35	15.78
Greek-Cypriot	44.24	13.26
Bangladeshi	43.67	6.11
Turkish-Cypriot	42.00	13.17
Pakistani	41.93	11.78
Indian	41.77	11.97
Nigerian	40.00	9.03
West Indian	39.37	11.11

The one-way analysis of variance was significant ($F=16.90$, $df=12,3392$, $p<.01$) (For ANOVA summary tables see Appendix 24.) The significant differences are shown in Table 9.14.

There were as many significant differences overall as for the Intelligence Test - but fewer significant differences involving the majority group. The 'E-W-S-I' mean was significantly higher than the 'West Indian' and 'Indian' group means. The 'Other' group mean was also significantly higher than the 'West Indian' and 'Indian' means.

The 'Italian' and 'Ghanaian' means were significantly higher than the 'West Indian' mean score.

Table 9.14 Significant Differences Between OPCS Groups For Executive Problems Test Score

	Italian	Polish	Other	Ghanaian	Chinese	Greek-Cypriot	Bangladeshi	Turkish-Cypriot	Pakistani	Indian	Nigerian	West Indian
English-Welsh-Scottish-Irish	NS	NS	NS	NS	NS	NS	NS	NS	NS	**	NS	**
Italian		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	*
Polish			NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
Other				NS	NS	NS	NS	NS	NS	**	NS	**
Ghanaian					NS	NS	NS	NS	NS	NS	NS	*
Chinese						NS	NS	NS	NS	NS	NS	NS
Greek-Cypriot							NS	NS	NS	NS	NS	NS
Bangladeshi								NS	NS	NS	NS	NS
Turkish-Cypriot									NS	NS	NS	NS
Pakistani										NS	NS	NS
Indian											NS	NS
Nigerian												NS
West Indian												

** $p < .01$
* $p < .05$
NS Not significant

9.4.2 Comparison of Executive Problems Test Scores For MCB Categories

The means and standard deviations for each MCB group are provided in Table 9.15.

Table 9.15 Executive Problems Test Mean Score For Each MCB Category

CATEGORY	MEAN (X)	STANDARD DEVIATION (S.D.)
British	49.91	10.39
Other British	49.86	11.29
British Chinese	48.17	16.41
European	46.69	13.02
African Asian	45.78	9.76
Other	45.50	16.38
British Asian	42.87	12.48
Chinese	42.21	14.57
Asian	41.44	11.45
British West Indian	40.52	10.71
African	37.83	11.16
West Indian	37.81	11.43

The one-way analysis of variance was significant ($F=18.99$, $df=11,3388$, $p<.01$) (For ANOVA summary table see Appendix 25.) There were considerably more significant differences on this test than for any of the others (see Table 9.16). The 'British' and 'Other British' scores were both significantly higher than those of 'British West Indian', 'Asian', 'West Indian' and 'African'. Additionally, the 'British' mean score was significantly higher than the 'British Asian' and 'Chinese' scores.

The 'African Asian' and 'European' means were also significantly higher than those of the 'West Indian' and 'African' groups.

Table 9.16 Significant Differences Between MCB Group For Executive Problems Test Score

	Other British	British Chinese	Other	European	African Asian	British Asian	Chinese	Asian	British West Indian	African	West Indian
British	NS	NS	NS	NS	NS	**	*	**	**	**	**
Other British		NS	NS	NS	NS	NS	NS	**	**	**	**
British Chinese			NS	NS	NS	NS	NS	NS	NS	NS	NS
European				NS	NS	NS	NS	NS	NS	*	*
African Asian					NS	NS	NS	NS	NS	*	*
Other						NS	NS	NS	NS	NS	NS
British Asian							NS	NS	NS	NS	NS
Chinese								NS	NS	NS	NS
Asian									NS	NS	NS
British West Indian										NS	NS
African											NS
West Indian											

** $p < .01$
* $p < .05$
NS Not Significant

9.5 Summary of Results on Differences Between Test Scores

The results obtained from an examination of the test scores can be summarised as follows:

(i) all of the one-way analyses of variance performed were significant

(ii) for the overall total mean score, on the OPCS classification, the 'E-W-S-I' mean score was significantly higher than those of the 'Greek-Cypriot', 'Indian', 'West Indian' and 'Pakistani' groups. The mean scores of the 'Italian', 'Other' and 'Ghanaian' groups were significantly higher than those of the 'Indian' and 'West Indian' groups. On the MCB classification the 'British' group mean was significantly higher than those of 'African Asian', 'British Asian', 'British West Indian', 'Asian', 'West Indian' and 'African' groups. The 'Other British' group was also significantly higher than the 'British West Indian', 'Asian', 'West Indian' and 'African' groups.

(iii) the above pattern of significant differences remained virtually the same when each of the tests was examined. The minority groups, however, appeared to perform relatively better on the Enquiries From Management test. Indeed, on the OPCS classification there were the fewest number of significant differences, but this could be due to the smaller numbers in some of the categories in this classification. The pattern of significant differences for the MCB groups on this test was similar to the pattern on other tests.

CHAPTER 10

Results: Examination of Factors Related to Pass/Fail and Test Performance

This chapter examines objective (ii) outlined in Chapter 7.6. This is to determine which other factors are responsible for (a) passing/failing the tests and (b) test performance.

The factors relating to pass/fail rates are examined first. Contingency coefficients were computed to determine the relationship between each of the variables and pass/fail. This was followed by a discriminant analysis in order to determine whether it was possible to distinguish between those people who had passed the tests and those who had failed.

In examining the factors relating to test scores, correlation coefficients were computed in order to determine the relationship between the independent variables and the total test score. Multiple regression analysis was then carried out on the data in order to determine the amount of variance in total scores which could be accounted for by the independent variables. To establish the reliability of the results tests for multicollinearity and double cross validations were carried out.

The factors which were significantly related to the test scores were partialled out, or controlled for, in an analysis of covariance, and the changes in the mean test scores for each group were identified.

10.1 Factors Related to Passing or Failing the Tests: Contingency Coefficient Results (see Table 10.1)

In order to measure the degree of association between each of the classificatory variables and the passing or failing of the tests, contingency coefficients were computed. The contingency coefficient was used because "It is uniquely useful when we have only categorical (nominal scale) information about one or both sets of [these] attributes" (Siegel 1956 p.196).

Table 10.1 shows the results of the contingency coefficient analysis. Each of the variables, except for sex and age, were significantly correlated with passing or failing the tests. The highest correlation was between education and pass/fail (c=.14, $p < .01$). Although the correlations are low, they are significant due to the large sample size. Age and sex (the only non-significant variables) are the ones for which separate norm tables have been developed. One important result of this, however, is that it could lead to a suppression of the effects of other variables which are related to age in particular ie. the variables of age on entering the UK and number of years lived in the UK.

Table 10.1 Correlations Between Classificatory Variables and Pass/Fail: Contingency Coefficient Results

Classificatory Variables	Contingency Coefficent (c)
Education	.14**
Country of Secondary Education	.12**
Country of Primary Education	.11**
Age on entering the UK	.11**
Number of years lived in the UK	.09**
Country of Further Education	.09**
Age	.05
Sex	.03

** $p < .01$

Furthermore age on entering the UK appears to be a relatively more important variable than either the number of years a person has lived in the UK and age.

The contingency coefficients for the countries where individuals received their primary, secondary and further education are all significant. It is reasonable to expect, however, that there will be a relationship between these variables and a person's self-categorisation on the OPCS and MCB classifications.

10.2 Discriminant Analyses on Pass Rates

Discriminant analyses were carried out to determine whether it was possible to distinguish between those people who had passed the tests and those who had failed.

Discriminant analysis is a multivariate statistical procedure for describing the classification of individuals into two groups (on the dependent variable). The analysis yields one or more equations (discriminant functions) similar to regression equations. The discriminant function would be a weighted combination of these measures that maximises the difference in results for people in the two groups (in this case whether they passed or failed). In stepwise discriminant analysis, one or more predictor variables are added or omitted at each step to determine which combination of variables provides the best discrimination among the groups.

For each set of analysis the following data are given:

(i) Wilks' Lambda (Λ) - this is an inverse measure of the discriminating power in the original (dependent) variables which has not yet been removed by the discriminating variables. Therefore the larger lambda is the less variance has been accounted for

(ii) Rao's V is a generalised distance measure. The variable selected is the one which contributes the largest increase in V when added to the previous variables. This amounts to the greatest overall separation of groups. In the following tables only variables which increased V by a significant amount are shown.

In addition, discriminant analysis is also a powerful classification technique ie. the process of identifying the likely group membership (pass or fail) of an individual when the only information known is that individual's scores or values on the discriminating variables.

Nine classificatory variables were entered into the analysis. These were: OPCS or MCB classification information; sex; age; the number of years the candidate had lived in the UK; the age at which he/she entered the UK; educational qualifications and the countries where the candidates received their primary, secondary and further education.

Many of these were nominal or categorical variables and so had to be converted into dummy variables. As Nie et al (1975) explain "A set of dummy variables is 'created' by treating each category of a nominal variable as a separate variable and assigning

arbitrary scores for all cases depending upon their presence or absence in each of the categories" (p.274) (their emphasis).

The results are presented first for analysis including the OPCS classification and secondly for the MCB classification.

10.2.1 Discriminant Analysis Including the OPCS Classification

In the stepwise analysis (Table 10.2) six variables brought about a significant change in Rao's V; the OPCS groups of 'E-W-S-I', 'Indian' and 'West Indian', educational qualifications, receiving secondary education in the UK and receiving primary education in India. A Wilk's Lambda of .94 was obtained, indicating that only 6% of the variance of the dependent variable was accounted for.

Table 10.2 Discriminant Analysis Including OPCS Classification

Variable	Wilks' Lambda	Rao's V	Change in V
E-W-S-I	.98	88.62	88.62**
Indian	.97	94.44	5.82**
West Indian	.97	102.50	8.06**
Education qualifications	.95	189.20	86.70**
UK Secondary Education	.94	199.49	10.29**
Indian Primary Education	.94	203.59	4.10*

** $p<.01$
* $p<.05$

Discriminant analysis can also be used to classify people and predictions made as to whether they will pass or fail according to the information available on the discriminating variables.

Table 10.3 shows the results of predicting outcomes based on the equation produced in the discriminant analysis. Of the people who actually failed only 49.7% are predicted correctly. Of the people who passed, however, 71.7% of people are correctly identified. Overall the percentage of people correctly classified was 57.78%.

Table 10.3 Results of Predicting Passes and Failures

		Predicted Outcome	
		Fail	Pass
Actual	Fail	49.7% (1085)	50.3% (1097)
Outcome	Pass	28.3% (357)	71.7% (905)

Percentage correctly classified = 57%

10.2.2 Discriminant Analysis Including the MCB Classification

In the stepwise analysis (Table 10.4) five variables brought about significant changes in Rao's V: 'British', 'Other British', educational qualifications, receiving primary education in India and age on entering the UK. A Wilks' Lambda of .95 was obtained, indicating that only 5% of the variance of the dependent variable was accounted for.

Table 10.4 Discriminant Analysis Including MCB Classification

Variable	Wilks' Lambda	Rao's V	Change In V
British	.98	76.92	76.92**
Other British	.98	86.30	9.38**
Educational qualifications	.95	162.91	82.91**
Indian Primary Education	.95	178.31	9.10**
Age on entering UK	.95	182.33	4.02*

** $p < .01$

* $p < .05$

The results of predicting outcomes are presented in Table 10.5. The equation produced in the discriminant analysis only correctly predicts 48.5% of the people who actually failed and 72.2% of people who passed, with an overall percentage of people correctly classified of 57.17%.

Table 10.5 Results of Predicting Passes and Failures

		Predicted Outcome	
		Fail	Pass
Actual Outcome	Fail	48.5% (1058)	51.5% (1124)
	Pass	27.8% (351)	72.2% (911)

Percentage correctly classified = 57.17%

The results of the discriminant analysis, therefore, indicates that two important factors in determining whether someone passes or fails are the group to which the individual belongs and their educational qualifications. It would also appear that the country where the person received their primary education is another factor to be considered.

It should be noted, however, that only 5% of the variance in pass/fail rates was accounted for by those variables. Furthermore the attempts to predict whether a person passed or failed produced results which were little better than would be expected by chance.

10.3 Factors Related to Test Performance: Correlations Between Test Scores and Independent Variables

In order to determine the degree of association between the classificatory variables and total test score correlation coefficients were computed (see Table 10.6). Significant positive correlations were obtained for country of primary education ($r = .21$, $p < .01$), country of secondary education ($r = .18$, $p < .01$) and educational qualifications ($r = .14$, $p < .01$) and a significant negative correlation was obtained for age on entering the UK ($r = -.20$, $p < .01$). This would indicate therefore that the people who tend to score better on the tests are those with higher educational qualifications and who entered the UK at an earlier age.

Table 10.6 Correlations Between Independent Variables and Total Test Score

Independent Variables	Correlation Coefficient (r)
Age on entering the UK	-.20**
Country of primary education	.21**
Country of secondary education	.18**
Education	.14**
Sex	.05
Number of years spent in UK	.03
Age	-.03

** $p < .01$

In order to ascertain more precisely the amount of variance in test scores accounted for by the independent variables, multiple regression analyses were carried out.

10.4 Multiple Regression Analysis on Test Scores

Multiple regression analysis enables the relationship between a dependent variable (in this case the overall weighted composite or total score) and a number of independent variables to be more precisely determined. In the multiple regression analyses carried out nine classificatory variables were entered into the equation; OPCS or MCB self classification, sex, age, the number of years the candidates had lived in the UK, the age at which he/she had entered the UK, educational qualifications, and the countries where the candidates had received their primary, secondary and further education.

Multiple regression is a statistical tool through which the relationships between a dependent variable and a set of independent variables can be analysed. As McNemar (1969) points out there are two principal uses for the technique "(i) it yields the optimum weighting for combining a series of variables in predicting a criterion and provides an indication of the accuracy of the subsequent predictions; (ii) it permits the analysing of variation into component parts" (p.206/207).

The principal uses therefore are for explaining and predicting. The accuracy of predictions can be determined by carrying out double cross validations.

The initial concern of this project was in determining what proportion of the variance could be accounted for by the variables in the study, and not in ascertaining the exact importance of each of the variables. For this reason, therefore, the variables were entered into the regression analyses in different

orders to see if any increases in the amount of variance accounted for could be obtained.

With the multiple regression analyses the problem of the nominal variables was overcome by using criterion scaling. Pedhazur (1982) states that dummy coding of nominal variables "becomes unwieldly when the number of variables is large or even with a small number of variables when each consists of many categories" (p.387/388).

He goes on to say that "A categorical variable is said to be criterion scaled when it is transformed to a single vector in which each individual's score is equal to the criterion mean of the group to which he or she belongs" (p.388). Criterion scaling is particularly useful when a variable selection procedure, such as a stepwise regression, is being used.

Two types of multiple regression analysis were carried out:

(i) stepwise analysis with OPCS or MCB information entered into the equation first

(ii) OPCS and MCB information entered into the equation last, with all other variables entered stepwise into the analysis.

There are, however, certain assumptions which have to be met in using multiple regression analysis. Included in these are that:

"1. The sample is drawn at random.

2. Each array of Y for a given combination of Xs follows the normal distribution.

3. The regression of Y and Xs is linear.

4. All the Y arrays have the same variance."

(Nie et al, 1975, p.341.)

If the sample size is large (as in this study) the assumption of a normal distribution can be relaxed. (Nie et al, 1975.)

The homogeneity of variances and linearity assumptions were seen to be met after the residuals were directly examined.

One problem which needs to be addressed in multiple regression analysis is that of correlation amongst the independent variables. Where there are high intercorrelations amongst variables there may also be difficulties in the estimation of the regression statistics. This problem, known as multicollinearity, "refers to the absence of orthogonality in a set of independent variables" (Pedhazur, 1982, p.233), was tested for first of all.

10.5 Testing for Multicollinearity

In testing for multicollinearity a number of regression analyses were carried out. In each case the dependent variable was one of the independent variables used in predicting the total test scores. In this way it was possible to determine the variance in each variable which could be accounted for by the other independent variables.

The highest degree of multicollinearity, not surprisingly, was found between the variables of age, number of years spent in the UK, and age on entering the UK. (In each case $R^2 = 1.00$ because if you know the value of two of the variables the third one can be calculated. Eg. if you know that someone's age is 25 and that they have been living in the UK for 15 years then the third

variable - the age at which the person entered the UK - is easily calculated.) There was also high multicollinearity between the classifications and country of primary and secondary education. (R^2 = .51.)

To a certain extent, however, this is to be expected. Often a person's 'ethnicity' is determined by their place of birth. The country of birth will, usually, also be the place where the child goes to primary and secondary school. The problem here, therefore, is not so much one of multicollinearity as of repetition. Gordon (1968) makes a distinction between redundancy (or high correlations between independent variables) and repetitiveness (or the number of redundant variables). The data presented here, therefore, might contain a number of variables which are measuring the same factor in different ways. However, although there are advantages in having repetitive variables there are also some hazards as Pedhazur (1982) points out. "Researchers frequently introduce multicollinearity by using multiple indicators for variables in which they have greater interest or which they consider more important from a theoretical point of view. This is not to say that multiple indicators are not useful. On the contrary, they are of utmost importance. When, however, multiple indicators are used in a regression analysis they can play havoc with the regression statements." (p.242). To control for the effects of multicollinearity the variables of countries of primary and secondary education were excluded from the analysis. In addition, only one of the variables of age, number of years spent in the UK and age at which the person entered the UK was included in the analysis at any one time.

10.5.1 Results of Testing for Multicollinearity For OPCS Classification (see Table 10.7)

Age and the number of years spent in the UK were not significant when included in the analysis. Age on entering the UK did have a significant effect, however. Table 10.7 shows the results of a stepwise analysis, 11% of the variance was accounted for (which incidentally was the same as when the highly correlated variables were included in the analysis). The change in R^2, however, was significant only for OPCS and educational qualifications. When OPCS was entered last into the analysis (for table see Appendix 26) age on entering the UK and the OPCS classification each accounted for 4% of the variance.

Table 10.7 Stepwise Multiple Regression Analysis Including OPCS With Highly Correlated Variables Excluded

Variable	R Square	F	R Square Change
OPCS	.08	161.26	.08**
Educational qualifications	.11	84.21	.03*

** $p < .01$
* $p < .05$

10.5.2 Results of Testing For Multicollinearity for MCB Classification (see Table 10.8)

For the MCB classification the variables of MCB and educational qualifications accounted for 10% of the variance.

When MCB was entered into the equation last (for table see Appendix 27) age on entering the UK and the MCB classification each accounted for 4% of the variance.

The classifications and educational qualifications were important factors in determining a person's test score. These results are similar to those obtained for the discriminant analysis of pass/fail rates.

Table 10.8 Stepwise Multiple Regression Analysis Including MCB with Highly Correlated Variables Excluded

Variable	R Square	F	R Square Change
MCB	.08	141.02	.08**
Educational qualifications	.10	79.42	.02*

** $p<.01$ * $p<.05$

10.6 Predicting the Total Score: Results of the Double Cross Validation

Double cross validation was employed in attempting to predict the total score since this is recognised as "the most rigorous approach to the validation of results from regression analysis in a predictive framework" (Pedhazur 1982, p 150).

In double cross validation the sample is split into two random sub-samples, and a regression equation is calculated for each one. The regression equation obtained in one sub-sample is then

applied to the predictor (or independent) variables of the other sub-sample. In each sub-sample, therefore, a predicted set of criterion scores is obtained. A Pearson correlation is then carried out between the observed and predicted criterion scores. The correlation coefficient obtained is analogous to an R^2. Therefore four R^2s are obtained - two obtained directly in each sub-sample, and two calculated on the basis of regression equations obtained from alternate sub-samples. The difference between the directly obtained R^2 and the calculated R^2 indicates shrinkage between the two. If the difference between the two is considerable, the equation cannot readily be used for predictive purposes.

10.6.1 Double Cross Validations with the OPCS Classification

The total sample was split into two random sub-samples, and double cross validations carried with the OPCS classification entered first into the regression analysis.

For the first sample the directly obtained R^2 was .12 and the calculated R^2 was only .03.

For the second sample an R^2 of .12 was directly obtained whereas the calculated R^2 equalled .02. There appears, therefore, to be considerable shrinkage in the R^2 here. (See Appendix 28.)

Similar results were obtained when the OPCS information was entered last into the equation. (See Appendix 29.)

10.6.2 Double Cross Validations with the MCB Classification

As with the OPCS classification the total sample was split into two random sub-samples and double cross validations carried out with the MCB classification entered first into the regression analysis. For the first sample the directly obtained R^2 was .12, and the calculated R^2 was .02. For the second sample there was an obtained R^2 of .10 and a calculated R^2 of .01 (see Appendix 30). Similar results were obtained when the MCB classification information was entered last into the regression analysis. (For summary table of results see Appendix 31.) The results of the double cross validation indicate that the regression equations obtained cannot usefully be employed in predicting the total scores.

10.7 Analysis of Covariance

Factors such as sex, educational qualifications, country of primary education and age of entry into the UK appear to be important in determining the total score. The examination of test scores was considered to be a most important aspect of this study and so an analysis of covariance (ANCOVA) was carried out to examine the effects on the mean test scores of controlling for (or partialling out) the above variables.

In any experimental study accuracy of the results will be increased by controlling any extraneous variability. This can be done in two ways, (i) by experimental (or direct control); (ii) by statistical (or indirect control) (Wildt and Ahtola 1978).

Direct control involves maintaining extraneous variables at constant levels eg. by grouping subjects into homogeneous blocks. The control is achieved through the design of experimental procedures. Indirect control relates to the analysis of the study rather than its design. It is achieved by first of all measuring one or more concomitant variables in addition to the independent variables of primary interest. "By partitioning out the amount of variability in the dependent variable accounted for by the concomitant variables, the experimenter is able to more accurately assess the influence of the other dependent (experimental) variables" (Wildt and Ahtola 1978). In this case the independent variable is the OPCS or MCB classifications. The concomitant variables - the ones whose effects were to be partitioned out were sex, educational qualifications, country of primary education and age on entering the UK.

10.7.1 ANCOVA Results for OPCS Classification

For the OPCS classification, the overall effects of the covariates were significant ($F=49.40$, $df=4,3318$, $p<.01$), and the effects of each individual covariate, except for country of primary education, was significant. The covariates account for 5% of the variance in total scores. Despite removing the effects of these variables, the effect of the OPCS classification was still significant ($F=19.72$, $df=12,3318$, $p<.01$) and accounts for 6% of the total variance. (For ANCOVA summary table see Appendix 32.)

The effect of partialling out these variables, however, increases the total mean score for most of the OPCS groups except for the 'E-W-S-I' and 'Italian' and 'Ghanaian' groups. (The reduction in

score for the first two groups is less than one point, whereas it is 8.98 for the 'Ghanaian' group.) The largest increase in total mean score was obtained by the 'Nigerian' group (8.62). (See Table 10.9.)

Similar patterns of results were obtained when ANCOVA was carried out on each of the individual tests. (For adjusted mean test scores see Table 10.9). (For ANCOVA summary tables see Appendices 33, 34 and 35.)

Visual examination of the adjusted means, however, shows that the partialling out of the variables has a greater relative effect on the mean scores obtained on the Executive Problems Test.

10.7.2 ANCOVA Results for MCB Classification

The effects of the covariates overall were significant ($F=44.60$, $df=4,3315$, $p<.01$) and they accounted for 5% of the variance. (For ANCOVA summary table see Appendix 36.) Even with the effects of the covariates partialled out the MCB classification was still significant ($F=19.61$, $df=11,3315$ $p<.01$) and overall accounted for 5.5% of the variance.

The effect of partialling out these variables, however, was to increase the total mean score for the 'overseas-born' minority groups and to reduce it for all the British groups (but by less than one point in every case). Generally speaking the differences between the original and adjusted total mean scores are smaller for the MCB classification than for the OPCS. (The highest increase in total mean score, for example, was 4.09 points obtained

Table 10.9 Comparison of Original Mean Test Scores with Adjusted Means For OPCS Groups

	Total Test Score		Intelligence Test		Enquiries from Management		Executive Problems	
	Adjusted	Original	Adjusted	Original	Adjusted	Original	Adjusted	Original
English-Welsh-Scottish-Irish	198.64	198.96	98.18	98.42	50.71	50.76	49.68	49.78
Italian	193.56	193.85	94.14	94.30	50.39	50.40	49.01	49.15
Greek Cypriot	165.14	164.35	75.75	75.27	45.07	44.88	44.40	44.24
Turkish Cypriot	163.86	161.00	79.99	78.29	41.29	40.71	42.71	42.00
Polish	193.97	193.06	94.44	93.71	51.19	51.35	48.27	48.00
West Indian	158.82	158.17	76.28	75.95	42.89	42.83	39.63	39.37
Nigerian	172.09	163.47	81.99	75.97	49.01	47.50	42.45	40.00
Ghanaian	177.53	186.51	86.56	90.38	48.80	49.98	43.53	46.15
Indian	163.92	158.31	76.10	72.27	45.21	44.34	43.53	41.77
Pakistani	161.58	157.40	74.12	71.27	44.80	44.20	43.30	41.93
Bangladeshi	175.25	169.67	82.88	79.00	47.75	47.00	45.58	43.67
Chinese	178.95	173.76	86.50	82.82	47.19	46.59	45.63	44.35
Other	189.89	187.78	93.21	91.78	48.68	48.37	48.33	47.63

by the 'Chinese' group.) (For table showing the adjusted total mean scores and other test mean scores see Table 10.10.) Similar patterns of results were obtained when ANCOVA was carried out on each of the individual tests. (For ANCOVA summary tables see Appendices 37, 38 and 39.)

As with the ANCOVA on the OPCS classification, the partialling out of the variables has a greater relative effect on the mean scores obtained on the Executive Problems Test.

10.8 Summary of Results Examining the Factors Related To Pass/Fail and Test Performance

The results obtained from examining the factors related to pass/fail rates and test performance can be summarised as follows.

1. The variables which are significantly correlated with pass/fail rates are educational qualifications, age on entering the UK, countries of primary, secondary and further education, and number of years lived in the UK.

2. The discriminant analysis, however, revealed that the factors in this study account for only 6% of the variance in pass/fail rates.

3. The variables which are significantly correlated with test performance are educational qualifications and age on entering the UK (this latter being a negative correlation).

Table 10.10 Comparison of Original Mean Test Scores with Adjusted Means For MCB Groups

	Total Test Score		Intelligence Test		Enquiries from Management		Executive Problems	
	Adjusted	Original	Adjusted	Original	Adjusted	Original	Adjusted	Original
British	198.94	199.20	98.20	98.39	50.83	50.90	49.79	49.91
British Asian	164.06	164.33	77.28	77.31	44.12	44.15	42.97	42.87
British West Indian	161.91	162.31	77.35	77.77	42.90	43.02	40.64	40.52
British Chinese	185.15	186.00	91.47	91.50	46.23	46.33	48.50	48.17
Other British	194.96	194.78	94.60	99.37	50.43	50.41	49.88	49.86
African Asian	171.46	168.43	79.67	77.26	46.23	45.39	47.24	45.78
Asian	162.31	158.67	74.90	72.27	45.89	44.96	43.16	41.44
West Indian	154.69	153.12	74.83	73.72	41.99	41.59	38.64	37.81
African	156.11	152.59	75.37	71.38	44.37	43.38	39.44	37.83
European	184.22	182.95	88.68	87.88	48.68	48.38	47.05	46.69
Chinese	173.80	169.71	83.34	80.21	48.40	47.29	44.09	42.21
Other	194.60	192.20	99.92	98.20	49.09	48.50	46.49	45.50

4. Multiple regression analysis was carried out but variables which had a high degree of multicollinearity variables were removed from the analysis. However, only approximately 11% of the variance in test scores was accounted for by the variables in the study.

5. Double cross validations using multiple regressions were carried out, indicating that the regression equations produced would not be useful for predictive purposes.

6. The analysis of covariance was carried out to determine the effects on mean test scores when the variables of sex, educational qualifications, country of primary education and age on entering the UK were excluded. The total mean score of all the OPCS groups (except for 'E-W-S-I', 'Italian' and 'Ghanaian') all increased. On the MCB classification the total mean scores of all the 'British' groups were slightly reduced, but the scores of all the overseas-born groups were increased.

7. The mean test scores most affected, relatively speaking, were those of the Executive Problems Test.

CHAPTER 11

Results: Examination and Comparison of Candidates' Reactions to the Classifications

An important part of the Executive Officer Study was to determine candidates' reactions to each of the classifications (objective (iii) in Chapter 7.6). They were asked to categorise themselves on two classifications - the OPCS and MCB. For each classification they were also asked how they would react if that classification was actually used to monitor candidates. They had to indicate their opinion on a five-point scale ranging from 'Strongly Agree' (1) to 'Strongly Object' (5).

Correlated t-tests were carried out which examined

(i) each OPCS group's reactions to the two classifications

(ii) each MCB group's reactions to the two classifications.

11.1 OPCS Groups' Reactions to the Classifications

Table 11.1 presents the means for each OPCS group's reactions to the OPCS and MCB classifications.

The groups which objected least to both the OPCS and MCB classifications were those of 'Polish' and 'E-W-S-I'. The groups which objected most strongly to the classifications were those of 'Turkish-Cypriot' and 'Pakistani'.

Table 11.1 Mean Scores for Reactions to the Classifications for OPCS Groups

OPCS Category	Reactions to OPCS		Reactions to MCB	
	$\bar{X}$	SD	$\bar{X}$	SD
E-W-S-I	3.64	.99	3.62	.99
Italian	4.25	.71	4.10	.83
Greek-Cypriot	3.95	.74	3.95	.84
Turkish-Cypriot	4.71	1.25	4.71	1.25
Polish	3.62	1.11	3.39	1.05
West Indian	4.11	.96	4.00	.97
Nigerian	3.80	1.00	3.80	1.10
Ghanaian	3.71	.80	3.78	.83
Indian	4.16	1.03	4.13	1.06
Pakistani	4.40	.97	4.67	.91
Bangladeshi	4.33	.58	4.33	.58
Chinese	3.65	1.17	3.82	1.09
Other	4.13	.99	4.09	.96

Correlated t-tests were conducted to identify any significant differences in each group's mean reactions to the two classifications. The only significant difference was for the 'E-W-S-I' group ($t=3.47$, $n=2834$, $p<.01$) which preferred the MCB classification. However, the actual difference between the two means was small, but the result was significant because of the large sample size.

11.2 MCB Groups' Reactions to the Classifications

Table 11.2 presents the means for each MCB group's reactions to the OPCS and MCB classifications. The 'British' and 'Chinese' groups objected to both classifications the least, and the groups which objected most strongly were those of 'Asian' and 'West Indian'. The differences in the reactions to the classifications were significant for the 'British' ($t=3.63$, $n=2873$, $p<.01$), 'British Asian' ($t=2.24$, $n=34$, $p<.05$) and 'Other British' groups ($t=2.20$, $n=48$, $p<.05$). In each case the MCB classification was objected to least.

Table 11.2 Mean Scores for Reactions to Classifications for MCB Groups

MCB Category	Reactions to OPCS		Reactions to MCB	
	$\bar{X}$	SD	$\bar{X}$	SD
British	3.65	.99	3.63	.99
British Asian	3.97	.93	3.80	.99
British West Indian	4.00	1.00	3.96	1.00
British Chinese	4.00	.75	4.00	.75
Other British	4.27	1.11	4.13	1.21
African Asian	4.16	.99	4.12	.97
Asian	4.27	.93	4.31	.86
West Indian	4.27	.95	4.35	.97
African	4.04	1.02	3.96	.93
European	3.87	1.20	3.92	1.19
Chinese	3.71	.89	3.79	.96
Other	4.00	1.07	4.00	1.07

11.3 Conclusions

The conclusions examine two areas; firstly the differences between the majority and minority groups' reactions to the classifications; and secondly, the differences in preference of the two classifications.

Overall, the majority group on the two classifications (ie. the 'E-W-S-I' and 'British' groups) tended to object to the classifications less than the minority groups. The mean scores for their reactions to the classifications ranged from 3.62 to 3.65. A score of '3' indicated a 'neutral' reaction and so it would appear that the majority group would not feel strongly about the use of either classification for monitoring purposes.

Apart from the 'Polish' group's reactions to the OPCS classification all the minority groups objected to the classifications more than the majority group. This is a finding which confirms the results of the earlier pilot studies (see Chapters 5 and 6).

On this evidence, therefore, if an organisation was to introduce ethnic monitoring of its workforce it could expect to obtain more initial resistance from the minority groups rather than the majority. This would appear to have important implications about the ways in which ethnic monitoring is introduced to people.

The second result, which should be noted, is that where there were significant differences between any individual group's preference for classifications, the MCB one was always objected to the least. It is also interesting to note that all the 'British'

minority groups 'preferred' the MCB classification (although the difference was only significant for the 'British Asian' group).

Overall, therefore, the MCB classification could be considered to be the most preferred classification. However, the differences between classifications, although significant for some groups, were small and most groups tended to react very similarly to both classifications.

CHAPTER 12

Conclusions and Overview of the Project

This final chapter has three parts to it. Firstly there is a review of the results relating to pass/fail rates, test scores and ethnic classifications. Secondly there is a critical evaluation of the methodology of the whole project and an indication of where further research work is necessary. The third part briefly presents the overall conclusions.

12.1 Main Conclusions To Be Drawn From The Study

This section reviews the results of the factors related to pass/fail rates and test scores; examines what the differences in test scores mean and whether the EO test battery is fair; and evaluates the implications for ethnic monitoring.

12.1.1 Review of the Results on Factors Related to Pass/Fail Rates and Test Scores

There were a number of factors which were significantly related to both the passing of EO test battery and to the test scores (see Tables 10.1 and 10.6). These were educational qualifications, the age at which the person entered the UK, and the countries where the person received their primary and secondary education. (The correlations between age on entering the UK and test scores were negative.) The correlation coefficients, however, were all low (none were above .20) and they were significant primarily because of the large sample size. Age and sex did not correlate

significantly with either passing/failing the tests or test scores. This was not surprising, however, because separate norms are used for age and sex. The result of this, however, would be to suppress the effects of variables related to sex and, in particular, age. These variables include the number of years lived in the UK and age on entering the UK. Despite this, however, age on entering the UK was still correlated significantly with passing/failing and with test scores.

With educational qualifications those with higher educational qualifications tended to perform better. This is a phenomenon which has been observed elsewhere, eg. examination of the norms of some commercially available tests reveal that graduates/ undergraduates tend to receive higher test scores on many types of tests than people who are less well qualified (eg. Nyfield 1980).

For age on entering the UK, those who entered the UK at an earlier age tended to perform better. This result tends to support the hypothesis that exposure to or familiarity with a particular culture might affect the test scores which an individual obtains. This reflects the earlier findings of McFie and Thomson (1970) and Watson (1973). McFie and Thomson (1970) found that West Indian children who had arrived in Britain before five years of age, performed better (on the verbal component of the Wechsler Intelligence Scale for Children) than those who arrived after that age. Similarly, Watson (1973) found that the performance of immigrant children on tests increases relatively with the length of schooling.

The X^2 for independent samples was applied to the data on pass rates for each group on both the classifications (see Tables 8.1 and 8.2). In both cases the X^2 were significant.

Furthermore, on the OPCS classification the pass rates of the 'E-W-S-I' group was significantly higher than those of 'Greek Cypriots' ($p<.01$), 'West Indians' ($p<.01$), 'Indians' ($p<.01$) and 'Chinese' ($p<.05$). (See Appendix 16.)

On the MCB classification the 'British' group pass rate was significantly higher than those of 'African Asian' ($p<.01$), 'Asian' ($p<.01$), 'West Indian' ($p<.01$), and 'African' ($p<.01$). Each of these minority groups might legitimately raise the issue of adverse impact. (See Appendix 17.)

In order to determine the amount of variance in the total score accounted for by the independent variables, multiple regression analyses were carried out. Certain problems had to be addressed, however, in conducting these analyses. In particular these were:

(i) the order in which the variables were entered into the regression analysis and

(ii) multicollinearity

Nie et al (1975) point out the difficulties of ordering the variables in the regression equation and highlight that the order in which variables are entered could affect the amount of variance which is accounted for. For this purpose, therefore, the variables were entered into the regression equation in different orders, so as to determine whether, by doing so, the total amount of variance explained could be increased.

A more difficult problem, however, was that of high multicollinearity amongst some of the variables to be entered into the regression equation. In particular, multicollinearity was found between the variables of countries of primary, secondary and further education and the OPCS and MCB classifications, and also between the variables of age, age on entering the UK and number of years lived in the UK.

The variables of countries of primary, secondary and further education and the OPCS and MCB classifications could be seen as examples of repetitiveness or as multiple indicators of the same variable.

In order to overcome this problem the repetitive variables (apart from the OPCS and MCB classifications) were excluded from the multiple regression analysis.

For age, age on entering the UK and number of years lived in the UK only one variable was entered into the regression on any one occasion, which meant that a number of separate regression analyses had to be carried out.

As Pedhazur (1982) states, however, there is no perfect or widely accepted solution to multicollinearity between variables (p.245).

The results of the multiple regression analyses indicated that group membership (as determined by the OPCS and MCB classifications) explained between 4% and 8% of the variance in test scores. Overall, however, the total amount of variance accounted

for by the variables included in the analysis was between 10% and 11% (see Tables 10.7 and 10.8 and Appendices 26 and 27).

Therefore, the amount of variance which the variables in this study explained, individually or in total, was only small. Any further study would need to examine other variables which might account for greater variance than was explained here.

Furthermore, the regression equations produced did not prove useful for predictive purposes. When double cross validations were conducted, considerable shrinkage was evident (see Appendices 28, 29, 30 and 31).

With the discriminant analysis group membership, educational qualifications and some countries of primary education brought significant changes in Rao's V. However, these variables only accounted for approximately 5% of the variance in pass/fail rates (see Tables 10.2 and 10.4). When variables in the discriminant analysis equation were used to classify whether people would pass or fail the tests the results were little better than would be expected by chance (see Tables 10.3 and 10.5).

12.1.2 Review of Results on Differences Between Groups' Test Scores

On each of the tests, and for the total test scores some groups performed consistently, and significantly, less well than others. On the OPCS classification the groups of 'West Indian', 'Indian', 'Greek-Cypriot' and 'Pakistani' performed significantly less well than the 'E-W-S-I' group. On the MCB classification the 'African Asian', 'British Asian', 'British West Indian', 'Asian', 'West Indian' and 'African' groups all performed significantly less well than the

'British' group. (Each of these groups, apart from the 'African Asian' and 'British Asian' ones, also performed significantly worse than the 'Other British' group.)

However, significant differences were only found between the groups with larger sample sizes. This obscures, to a certain extent, those groups which performed consistently well on all the tests. On the OPCS classification this included the 'Italian', 'Polish', 'Ghanaian' and 'Other' groups. On the MCB classification the 'Other British', 'British Chinese', 'European' and 'Other' groups also performed well. However, only on few occasions were significant differences found between them and the lower scoring groups. For example, on the Enquiries from Management test (see Table 9.9), the 'Polish' group was the highest scoring one on the OPCS classification. However, no significant differences were found between this group and any of the others.

Of the non-European minority groups, the 'Ghanaians' performed well (eg. see Tables 9.1 and 9.2). It was not possible to identify conclusively why this group should have performed so much better than other African groups but it is an area which deserves further investigation.

On the MCB classification there was a trend for the British-born groups (as defined on the MCB classification) to perform relatively better than their overseas-born counterparts. These differences were not significant, however. It should be noted that the British-born minority groups (in particular 'British Asians' and 'British West Indians'), still performed significantly less well than the 'British' group on many of the tests. This finding tends to

support those of Yule et al (1975) where a group of West Indian children born in the UK scored significantly higher on WISC than those born in the West Indies, but whose scores were still well below the national average.

This might also be seen as lending some support to the 'cultural familiarity' hypothesis. In fact Anastasi (1976) felt that as the culture of a country affects behaviour differences, exposure to a particular culture will, and indeed should, be detected by the tests. Nevertheless, she also stressed that differences between people of different backgrounds should only be considered if they affect performance in the job.

The analysis of covariance, however, indicated that the effect on mean test scores of partialling out other variables was, for many groups, only very small. Other factors need to be examined therefore, which might possibly have a greater contribution to make in the overall variance in test performance (see Tables 10.11 and 10.12).

Previous research (eg. Jensen 1980, Coleman et al 1966) appeared to indicate that verbal tests could well be biased against at least some minority groups and that verbal reasoning tests might possibly show more bias. With the three Civil Service tests, however, it was not possible to identify any one test which might be more biased against minority groups. Nevertheless there was a tendency for minority groups to obtain lower scores on the Executive Problems Test.

12.1.3 What Do Differences in the Test Scores Mean?

Having identified the differences in the test scores, the issues which have to be addressed are what do the differences mean and what inferences can be drawn from them?

Thorndike (1982), identifies three types of inference which can be made. Firstly, that the test is biased at item level. Secondly, that people from the lower scoring groups will generally do less well in the job for which they are being selected than people from the higher scoring group. Thirdly, that some groups have inherently less potential for carrying out certain kinds of jobs.

Bias at the item level was reviewed by Thorndike (1982) who felt that where bias occurred, this was not likely to be the cause.

The second inference, however, is one which is made by the Civil Service in using the EO test battery. The validation studies conducted (Paul and Kuplens 1977) and outlined in Chapter 2, indicate however, that the EO tests have only low, if significant, correlations with job performance. Furthermore, it is not possible to determine whether the tests are adequately valid for all groups.

If the third inference were accepted it would mean that not only do tests predict future job performance but also that they determine inherent factors differing in the three groups ie. that the differences are inherited or genetic. There are two important considerations here. Firstly, this is a very controversial and difficult area to research as it is virtually impossible to identify, isolate and define relevant environmental and motivational factors on the one hand and genetic characteristics on the

other. Secondly, this inference assumes that there is a clear, demonstrable link between test scores and job performance. In this case, it is argued, further information is required before this can be said to be equally true for the minority groups.

12.1.4 An Examination of the EO Test Battery's Fairness Under the Race Relations Act (1976)

One question which has to be addressed is whether the EO test battery is fair under the Race Relations Act (1976). The section of the Act which applies to the use of tests is that which outlines indirect discrimination. This is defined as when a person "applies to that other a condition or requirement which he applies equally to persons not of the same racial group as that other but (i) which is such that the proportions of persons of the same racial group as that other who can comply with it is considerably smaller than the proportion of persons not of that racial group who can comply with it, and (ii) which he cannot show to be justifiable irrespective of the colour, race, nationality or ethnic or national origins of the persons to whom it applied and (iii) which is to the detriment of that other because he cannot comply with it" (S.1.(i)(1)).

As mentioned in Chapter 2, cases of indirect discrimination will revolve around the meaning of 'justifiable'. Following the American example justifiability has become associated with necessity ie. necessary for a business to operate and is not merely convenient or desirable.

If there are disparities in the selection rates for different groups an employer will be potentially at risk, and so could be called upon to justify their procedures.

However, if the American 'four-fifths' rule of thumb was applied, then virtually all the minority groups on both classifications could claim adverse impact against them (see Chapter 8).

Given the differences in the pass rates of the various groups the Civil Service could be called upon to justify the use of the EO test battery.

From the figures produced in this study, it is not possible to state whether the tests are indirectly discriminatory. The figures indicate that the proportion of many of the minority groups who are able to pass the test is considerably smaller than the proportion of the majority group who are able to do so. However, it is not possible from this data alone to state whether the tests are being justifiably used. One way of establishing the tests' justifiability is to establish their validity. Tests with low validity coefficients have a higher probability of being indirectly discriminatory. Nevertheless, tests with high validity might still not be fair eg. the test might be differentially valid, that is to say that the tests might be significantly more valid for some groups than it is for others.

12.1.5 The Implications of the Project for Ethnic Monitoring

Prior to the 1981 census the OPCS decided not to include a question on ethnic origins (OPCS 1980) (see Chapter 4). The implications behind this were that:

(i) the form of the question was inappropriate

(ii) it would not be reliable

(iii) it would not be acceptable.

These are the issues faced by any organisation wishing to introduce ethnic record keeping, and the ones which were addressed in this project.

The pilot studies on ethnic monitoring indicated a number of points, including that:

(i) it is possible to obtain classifications which are accepted by respondents

(ii) the reliability of the classifications can be improved by the provision of instructions

(iii) different classifications might be useful in different circumstances

(iv) certain terms and expressions should be avoided wherever possible.

(i) Acceptance by Candidates

Many of the groups of people when faced with the classifications either objected to their use or reacted neutrally to them.

The reactions to any particular classification is dependent to a certain extent upon the nature of the sample. For example, in the third pilot study (Chapter 6) the school-leaver group objected less to each of the classifications than the university sample.

However, in the Executive Officer study, and in the earlier pilot studies the minority groups tended to object more to the use of the classifications than the majority group, indicating more resistance amongst minority groups to ethnic monitoring. This has very serious implications in terms of the way an organisation should introduce ethnic monitoring. Time and care must be spent on introducing the idea of monitoring, in particular to the minority groups within the organisations. The way in which this is done, and the classifications which are used, could affect the response rates and ultimately the overall effectiveness of the results. Judging from people's comments about the classifications in the pilot studies it would appear that there are misapprehensions and misgivings about ethnic monitoring (see Tables 5.6 and 6.4). In particular people are afraid that the information will be used to make selection decisions. These fears need to be allayed if any monitoring exercise is to gain widespread acceptance and support within an organisation.

(ii) The Reliability of the Responses on the Classifications

Much attention and time has been devoted, by organisations such as the OPCS, to obtaining a classification which is reliable. The research has examined different classifications and experimented with categories within those classifications. However, the 'perfect' classification, ie. one which is universally accepted and which reliably identifies minority groups, has not been obtained. Unreliability of responses on a classification could be due to a number of factors which include:

(i) ambiguity in the categories ie. respondents might not know what particular categories are referring to

(ii) having categories on a classification which overlap, thus giving some respondents a choice of categories to place themselves in

(iii) having respondents who deliberately lie about their origins.

Apart from directly supervising people when they complete the classifications there is little that can be done to prevent the last problem. There is, however, a solution which can be applied to overcome the first two problems.

The solution used in this study was the straightforward one of using instructions to define closely what was meant by each of the categories on any classification. By providing these guidelines or definitions, respondents are left with little doubt as to how they should respond, as the ambiguity in any particular classification will have been reduced. The reliability of classifications, therefore, depends more on the amount of guidance given to respondents than on the actual categories within the classifications. If the instructions are clear, then almost all classifications should achieve good reliability. 100% reliability will probably not be attainable as the CRE (1978) admitted when they stated that "in whichever form the question is asked it is likely that some individuals will find it difficult to reply because of the complex nature of their origins and that some replies will be misleading" (p.12). However, it is clear that by these straightforward methods reliability can be considerably increased.

Pilot Study II (see Chapter 5), however, has highlighted a problem which has rarely, if ever, been tackled with ethnic classifications - that of test-retest reliability. It would appear that the classifications as they are presently being used might well be unreliable in a test-retest sense. The use of instructions and definitions might be expected to improve this, but it is an area which deserves further research.

(iii) The Use of Different Classifications in Different Situations

It could be that certain types of classifications will be better in certain situations than others. With self-classifications it is possible to have a very comprehensive list of categories describing a person's origins, but when using selection tests, however, it might be useful and pertinent to split minority groups into those who were born in Britain and those who were born abroad. Certainly the more general categories of the MCB classification are more amenable to having statistical analyses applied to them as each category has larger numbers of people.

Finally, the reversed question could be used in situations where the self-perception of the individual might not be as important as how they are perceived by others.

(iv) Terms and Expressions to be Avoided

References to colour and anthropological classifications should not, as far as possible, be used in ethnic monitoring. Classifications containing references to colour, and those based on anthropological categories were objected to strongly (see Table 5.5). Furthermore, in the pilot studies all tended to have higher rates of non-response. The Committee on Race Relations Research (1975)

stated that terms used in classifications should not be offensive to respondents (see Chapter 4). The pilot studies enabled respondents' reactions to be carefully considered in order to produce classifications which were acceptable. The high response rate in the Executive Officer study (see Chapter 7), which was much higher than had been anticipated, might well be related to the fact that respondents were faced with terms and questions which they found less offensive.

12.2 A Critical Examination of the Project

This section examines and evaluates the project as a whole, and in particular its strengths, weaknesses, and future research which needs to be conducted.

12.2.1 The Executive Officer Test Battery Results

From an equal opportunities viewpoint the most significant statistics which emerged were the pass/fail rates of the different groups. However, it could be argued that the sample used in the Executive Officer study represented only approximately 17.5% of the total applicants of 20,000.

Some groups however were, in terms of absolute numbers, under-represented in the Executive Officer sample. This was particularly evident on the OPCS classification (see Table 7.2(a)).

Furthermore, the majority of the sample was obtained from the London testing centre (82.63% of the total sample), and perhaps a

differently composed sample would have been obtained had the other testing centres been sampled more often.

Nevertheless, the sample used in this study possibly contained a higher proportion of ethnic minority groups than would be represented in the total sample. Data gathering only took place in urban areas where higher numbers of ethnic minorities might reasonably be expected. Indeed 13.95% of the total sample identified themselves as belonging to a minority group and in this respect, therefore, it could be said that the results are generalisable to the total sample.

However, in examining the factors associated with test scores it would have been very useful to have included more variables in the study. In particular, two factors would have proved interesting to examine. Firstly, the social class or background of the applicants. Secondly, it was not known which candidates and how many, were internal Civil Service candidates, and the effects that this had on candidates' test scores ie. did internal candidates perform better on the tests and also what proportion of the ethnic minority candidates were internal?

These were factors, however, which the Civil Service refused permission to examine.

More research is required, moreover, to reexamine the factors related to test scores. It would appear that those people who entered the UK at an earlier age and who have better educational qualifications tend to perform better. The age of entry into the UK emerged as a significant factor despite the fact that its

effects may well have been suppressed due to the use of separate age norms. The importance of this variable therefore may be underestimated in this study. Furthermore, there were some indications that British-born minority tend to perform better than their foreign-born counterparts.

12.2.2 Examination of the Whole Selection Procedure

The Executive Officer study examined the pass/fail rates of the various groups taking the test battery. The tests, however, are not the only hurdle in the selection procedure. Those candidates who are successful in the tests are invited to be interviewed, and only after this is the decision taken as to whether candidates are accepted or not. A complete review of the Executive Officer selection procedure must include an examination of how minority group applicants fare at this stage, which would obviously require monitoring of the interviewing stage. The areas which might be examined include (i) the person specification used by interviewing panels - in particular are they used and if so are there any criteria which could possibly be biased against minority group applicants; and how consistently do interviewing panels use the criteria? (ii) the treatment of candidates at interviews - in particular do interviewers treat ethnic minority applicants differently? (iii) interviewers' awareness of how bias may creep into their selection decisions - in particular what training have interviewers received not only in interviewing but also in becoming aware of their biases and how these affect interview decisions?

12.2.3 Examining Differential Validity in the EO Test Battery

The issue of differential validity is legitimately raised with the EO test battery. Differential validity exists where "(a) there is a significant difference between the correlation coefficient of a selection device and a criterion obtained for one ethnic group and the correlation of the same device with the same criterion obtained for the other group. And (b) the validity coefficients are significantly different from zero for one or both groups" (Boehm 1972, p.33).

In this case there are significant differences between the test scores obtained by the majority group and some minority groups. It could be that differential validity exists here, and so needs to be investigated further, possibly following the guidelines laid down by Katzell and Dyer (1977).

Taking a wider perspective, the issue of differential validity should be one which concerns all psychometricians. The range of occupational tests commercially available is beginning to increase and psychologists need to be anticipating and examining those issues of test use which will become important in the future. Indeed, since 1968 there have been many papers published on differential validity and selection fairness, but virtually all of them have been American. This is a neglected area in occupational psychology within the UK which deserves far more attention than it is currently receiving and this is in spite of the fact that the equal opportunities legislation has been in force for ten years now.

12.2.4 Researching Methods Of Making Selection Fair: Job Analysis and Work Sample Tests

More research effort needs to be placed into making selection fairer and more effective. In Guion's (1976) 'Tenets of Orthodoxy' (see Chapter 3) the foundations of a good selection procedure depended on a systematic, thorough, job analysis. Unfortunately, many organisations do not undertake detailed job analyses before implementing a selection procedure.

In America, a job analysis would be seen as an essential pre-requisite to any test design and implementation. The Uniform Guidelines (1978) state that "There should be a job analysis. This job analysis should show the work behaviour(s) required for successful performance of the job or group of jobs being studied, the critical or important work behaviour(s) in the job or group of jobs being studied and an identification of the construct(s) believed to underline successful performance of these critical or important work behaviour(s) in the job or jobs in question. Each construct should be named and defined, so as to distinguish it from other constructs. If a group of jobs is being studied the jobs should have in common one or more critical or important work behaviours at a comparable level of complexity" (Section 14D(2)).

No such set of detailed guidelines exists in the UK. The Codes of Practice published by the Commission for Racial Equality (1984) and the Equal Opportunities Commission (1985) are the nearest equivalent, but they lay down very general rules about good personnel practice. There is a need for a set of guidelines

produced for and to be used by psychologists, personnel specialists and others involved in selection and recruitment which details the procedures which should be followed when introducing or implementing new selection procedures.

Job analysis also represents one way of examining selection methods which have less adverse impact but which are also valid. An area where these twin aims might be achieved is that of work sample tests. In general, work sample tests involve the applicant in performing a task or set of tasks which are thought (usually on the basis of a job analysis) to have direct, central relevance to the job in question. The task(s) performed are often sampled from the range of tasks performed by job holders.

Robertson and Kandola (1982) reviewed research work on work sampling with regard to adverse impact and concluded that the results have been encouraging. The research can be split into two areas: first those studies which have concentrated purely on the selection rates of minority and majority groups, and secondly those studies which compare validities for the two groups.

The work of Schmidt et al (1977) belongs in the first category. They looked at the adverse impact of a written paper and pencil test compared to that of a specially designed work sample test. On the written test 9% of the minority group scored above the majority group mean, but on the performance test 24% were above the majority mean. Similar results were also obtained by Cascio and Phillips (1979) in a comprehensive study on the selection rates of minority and majority groups using work sample tests.

These studies, though, suffer from the fact that they do not report any validity data. Some studies, however, have addressed themselves specifically to validity.

Kesselman and Lopez (1979) working with accountancy personnel, compared the validities of minority and majority groups for a custom-built work sample test, with the commercially available measure of mental ability (yielding verbal, numerical and total scores). Their hypothesis that a selection instrument based on a job analysis would give higher validities with respect to performance criteria and at the same time reduce the probability of unfair discrimination when compared with a selection instrument not utilising job analysis information was borne out. The validity for the minority group (n=28) using the custom-built test was 0.37 compared with 0.18 for the commercially available test, and the coefficients for the majority group (n=24) were 0.56 and 0.2 respectively. They conclude that the findings of the investigation demonstrate that a well constructed content-valid test based on a thorough job analysis can tap important determinants of job success in which racial differences are small or non- existent. These results were echoed in studies by Grant and Bray (1970) and Field, Bayley and Bayley (1977).

Schmidt et al (1977) feel that work sample tests give minority groups a better chance of being selected because they manage to get at those abilities which are required of the job, whereas traditional measures, "may be tapping precisely those determinants of job success on which racial differences are largest, conversely they may fail to tap important determinants of job

success on which such differences are smaller or perhaps non-existent" (p.188).

Despite the generally encouraging results of most of the studies mentioned above there has been little other work regarding the use of work sample tests in reducing adverse impact, and little published work at all in Britain. There can be little doubt that research regarding the discriminatory effect of the present selection procedures is valuable and that the work sample test allows an opportunity to reduce adverse impact whilst at the same time maintaining or even increasing validity.

12.2.5 Implications of the Research for Ethnic Monitoring

The pilot studies on ethnic monitoring indicated that there are certain terms and expressions which should be avoided in classifications because people object to them. They also indicated, however, that the most effective way of improving the reliability of classifications is to provide clear instructions to respondents on how they should complete the classification, and also defining what the categories mean. To date, no classifications provide these instructions - the assumption being that people know which category on a classification they belong in. However, this assumption does not allow for the fact that some people might identify with more than one category on a classification, each of which is legitimate to that individual. This in fact is the major source of unreliability with the classifications. By using the methods outlined in this study it will be possible for organisations to introduce more subtle and informative forms of ethnic classifications.

However, the samples used in the pilot studies were comparatively small, when you consider that the total population in fact consists of all people above school-leaving age. More research, therefore, needs to be carried out examining the issues explored in the pilot studies but using larger samples. In particular greater effort needs to be placed in examining the test-retest reliability of classifications, which is an issue which has rarely, if ever, been tackled with regard to ethnic monitoring. Furthermore, a greater period of time between testing and retesting of the sample would perhaps be desirable.

12.3 Overall Conclusions

The work on ethnic monitoring, in the pilot studies and the Executive Officer study, indicated that by being careful in their construction it is possible to obtain classifications which are acceptable, and reliable. In doing so, however, there are certain expressions and classifications which should be avoided as they are found offensive by respondents. Overall, minority groups tended to object more to the use of ethnic classifications. This has implications about the ways in which organisations should introduce ethnic record keeping into the workplace.

The results of the Executive Officer study indicated that some minority groups, in particular those of Indian, West Indian, Pakistani and Greek-Cypriot (on the OPCS classification) and those of African Asian, British Asian, British West Indian, Asian, West Indian and African (on the MCB classification) all performed significantly less well than the majority groups on the EO tests.

The multiple regression and discriminant analyses which were carried out indicated that the factors of group membership (as determined by the OPCS and MCB classifications), educational qualifications and age on entering the UK accounted for significant, though small, amounts of the variance in test scores and pass/fail rates.

The project also highlighted the need for further research in particular areas including:

(i) examining test-retest reliability of ethnic classifications using larger samples

(ii) examining the factors which are related to test performance in order to identify why certain minority groups perform less well. Certain factors were identified in this study, but wider-ranging studies including a larger number of variables are required before satisfactory explanations can be provided.

Finally, there is a wider need for research on the issues of fairness and bias on the use of tests in the United Kingdom. Even though the anti-discrimination legislation has been in force for nearly ten years very little research has been conducted by British psychologists on the very important issues of differential validity and test bias. These are issues which have initiated large amounts of research and discussion in the United States of America, and although the results of these studies are extremely interesting a more specific body of work relating to the position of ethnic minority groups within the United Kingdom is needed.

The issues of equality of opportunity and fairness cannot be ignored, and, although the American experience teaches us that hard and fast solutions are hard to come by, these are the challenges which British psychologists must address in the near future if they are to have a significant voice in the debates yet to come.

APPENDIX 1

Examples of Questions in the Executive Officer Qualifying Test Battery

Civil Service Commission

EXAMPLES OF QUESTIONS IN THE EXECUTIVE OFFICER QUALIFYING TEST BATTERY

The Executive Officer Qualifying Test Battery consists of the following different tests:

1. Enquiries from Management
2. Intelligence Test
3. Executive Problem Test

A general description of these tests and some examples of the types of questions in them are given in this leaflet. The answers to these examples are given on the back page. No other examples are available.

As the actual tests may not be in precisely the form indicated in this leaflet, you should carefully read the instructions before you begin. In particular, you should note how the answers are to be shown.

The tests are generally difficult and it is expected that even the best candidates will gain less than full marks. Within each test the questions are arranged roughly in order of difficulty with the easiest first. In this leaflet too the examples of each test have been arranged in approximate order of difficulty. The easiest examples are intended to be simpler than the questions in the tests to let you see clearly the types of reasoning involved, and the most difficult examples are designed to be as hard as the hardest questions in the tests.

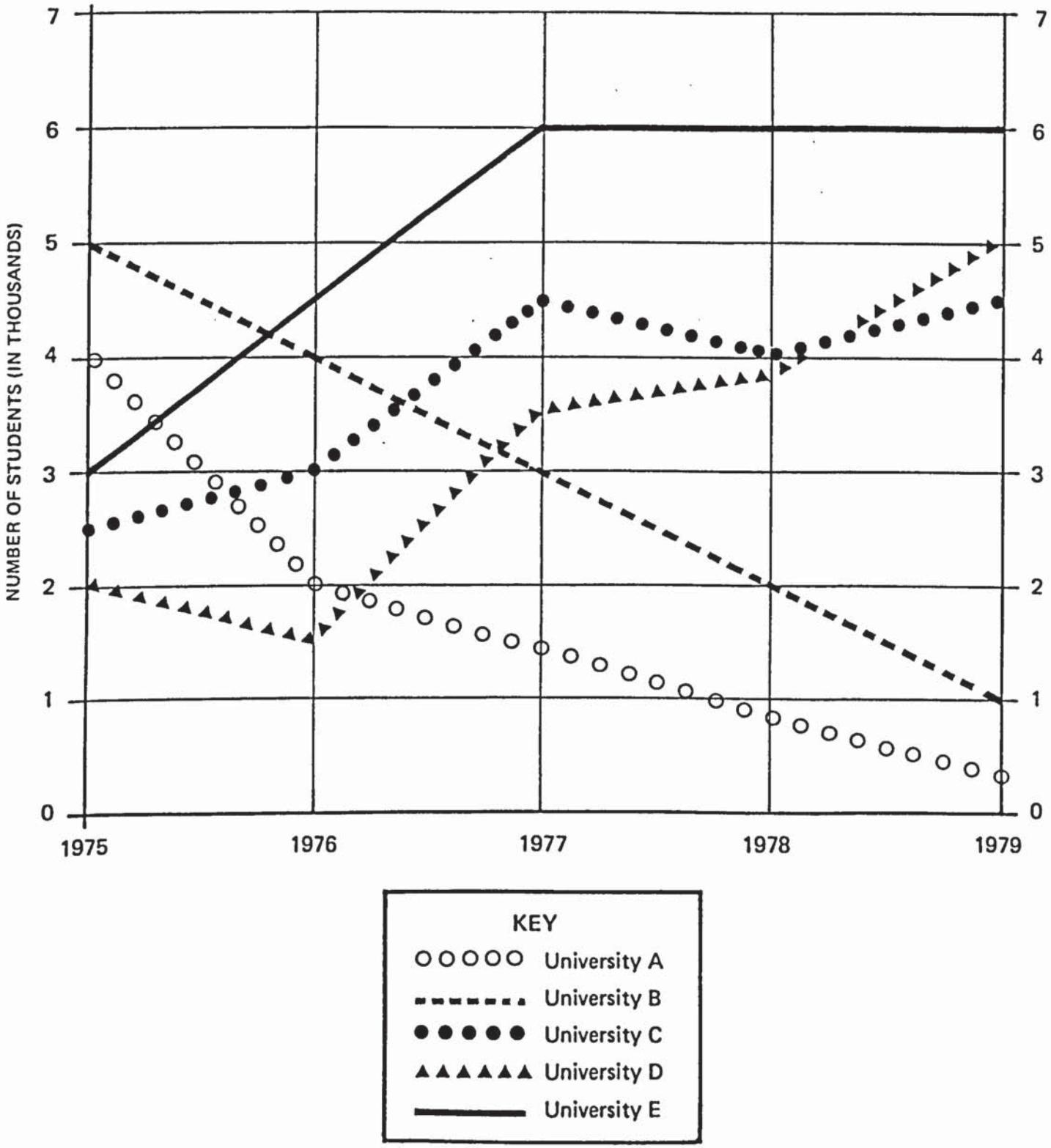
When you take the tests it is possible that you will not have time to attempt all the questions in each one. This need not matter at all. Just work steadily, doing as many questions as you can within the time allowed for each test. Remember to avoid spending an undue amount of time on a question you find difficult; it may be better to go on to the next question.

ENQUIRIES FROM MANAGEMENT

This is a test of numeracy lasting 25 minutes, in which you are required to understand and deal with arithmetical figures. It consists of 5 graphs and tables, each with a number of questions attached. The questions are answered by putting a tick through the correct one of 5 possible answers.

EXAMPLE 1

NUMBER (IN THOUSANDS) OF STUDENTS AT FIVE UNIVERSITIES: FIGURES FOR 1975 – 1979



QUESTIONS

1. Which university had the fewest students in 1979?

A B C D E

2. What was the total number (in thousands) of students at these 5 universities in 1975?

14 15 16½ 17 16,500

3. Which university showed a decrease at a constant rate from 1975 to 1979 in its student population?

A B C D E

4. Which university had the same number of students in 3 consecutive years?

A B C D E

5. Which university increased its student population in each of the years between 1976 and 1979?

A B C D E

EXAMPLE 2

NUMBER OF RAINCOATS, UMBRELLAS AND HATS SOLD BY FIVE SHOPS IN JUNE AND DECEMBER 1979

Shop	June			December		
	Raincoats	Umbrellas	Hats	Raincoats	Umbrellas	Hats
A	10	7	0	45	60	46
B	8	6	1	35	52	34
C	6	7	0	32	55	33
D	10	8	0	52	75	43
E	8	12	0	46	80	46

QUESTIONS

1. Which shop sold the same number of raincoats as hats in December 1979?

A B C D E

2. In June 1979, one shop had the least difference in its sales of raincoats and umbrellas. Which shop was it?

A B C D E

3. The number of umbrellas sold by one shop in December 1979 was 6 times its sales of raincoats in June 1979. Which shop was it?

A B C D E

4. Which shop had the greatest increase in the actual number of umbrellas sold in December 1979 as compared with June 1979?

A B C D E

5. What was the percentage increase in the total sales of raincoats in December 1979 as compared with June 1979?

42% 400% 500% 600% 4,000%

EXAMPLE 3

NUMBER OF JOB APPLICANTS WHO PASSED AND FAILED THE SELECTION TEST USED BY FIVE EMPLOYERS FROM 1975 TO 1979. (EACH JOB APPLICANT WAS ONLY TESTED ONCE AND BY ONLY ONE EMPLOYER.)

Employer	Result of test	Year				
		1975	1976	1977	1978	1979
1	Passed	87	68	71	58	93
	Failed	88	79	87	85	92
2	Passed	4	4	5	11	12
	Failed	7	6	7	15	13
3	Passed	2	2	3	3	3
	Failed	4	4	2	6	3
4	Passed	10	11	9	17	15
	Failed	12	12	11	25	14
5	Passed	5	7	6	11	10
	Failed	6	4	8	14	20

QUESTIONS

1. What was the total number of job applicants who answered the test in 1975 and 1976?

   197 200 225 322 422

2. In which year did more job applicants pass than fail the test given by Employer 1?

   1975 1976 1977 1978 1979

3. Which employer gave the test to more job applicants in 1976 than in 1975?

   1 2 3 4 5

4. Of the total number of successful job applicants in 1978, what percentage was given the test by Employer 3?

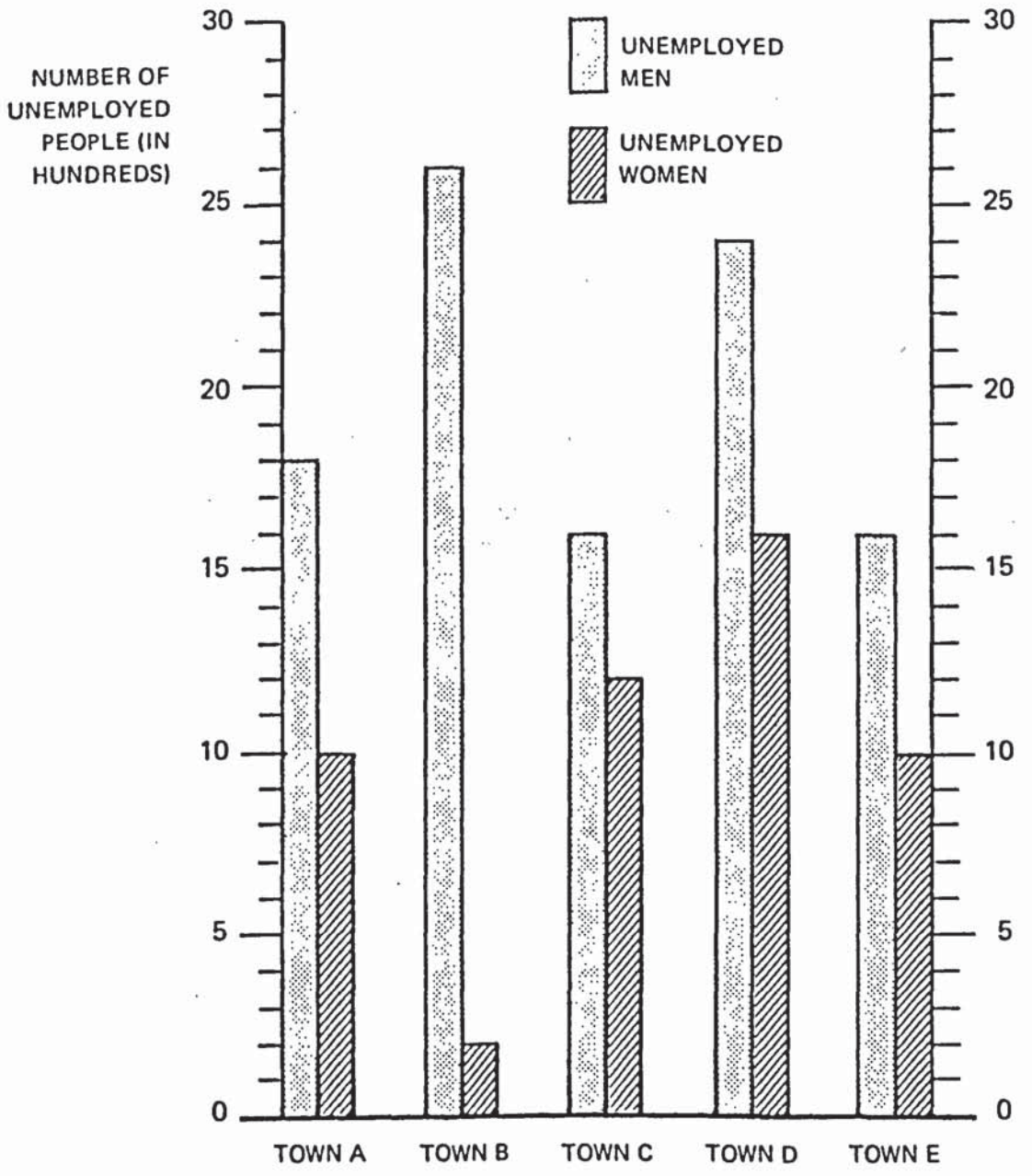
   3% 6% 9% 11% 17%

5. What was the ratio of the number of job applicants who passed the test given by Employer 2 in 1975 to the number who failed the test given by Employer 5 in 1979?

   1:6 1:5 1:4 7:20 2:5

EXAMPLE 4

NUMBER (IN HUNDREDS) OF MEN AND WOMEN WHO WERE UNEMPLOYED ON 1 JULY 1979 IN FIVE TOWNS



QUESTIONS

1. Which town had the greatest number of unemployed people on 1 July 1979?

A B C D E

2. Which 2 of these 5 towns had the same number of unemployed women on 1 July 1979?

A and C A and E C and D C and E D and E

3. What was the total number (in thousands) of unemployed men in these 5 towns on 1 July 1979?

10 15 50 100 10,000

4. What proportion of the total number of unemployed women in these 5 towns on 1 July 1979 was in Town E?

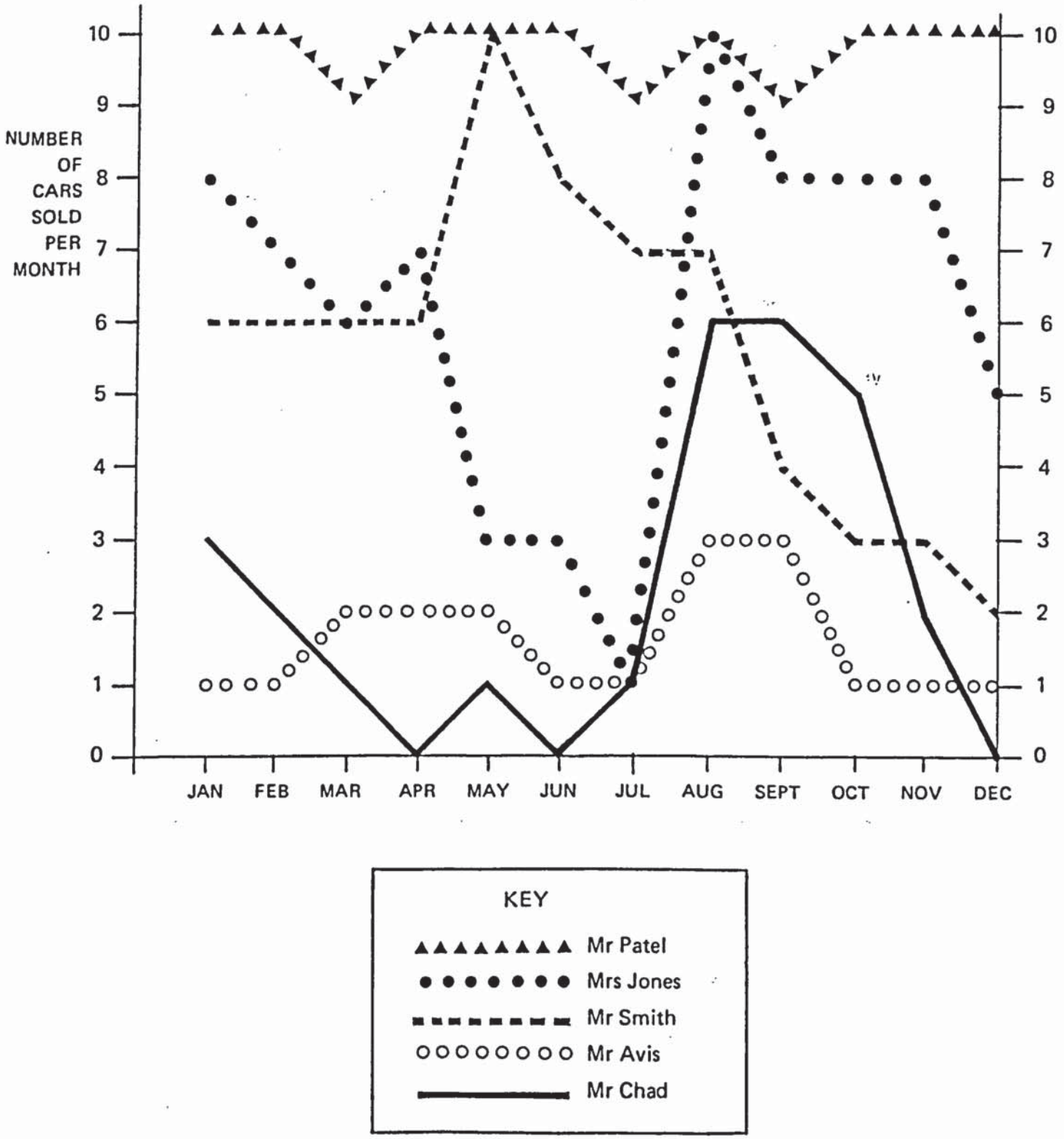
1/20 1/15 8/75 1/5 8/25

5. On 13 July 1979, the number of unemployed men and women in Town D increased by 25% and 150% respectively over the numbers unemployed on 1 July 1979. What was the total number of unemployed people (in hundreds) in Town D on 13 July 1979?

30 40 46 62 70

EXAMPLE 5

NUMBER OF CARS SOLD BY FIVE GARAGE-OWNERS IN EACH MONTH OF LAST YEAR



QUESTIONS

1. Whose car sales were the most consistent throughout the year?

Mr Avis Mr Chad Mrs Jones Mr Patel Mr Smith

2. Who sold the same number of cars in each of 4 consecutive months?

Mr Avis Mr Chad Mrs Jones Mr Patel Mr Smith

3. In which month did 3 of the garage-owners sell the same number of cars?

March April May July August

4. In which month did the car sales of Mrs Jones and Mr Smith differ most?

May June July October November

5. Which month showed the largest increase in total car sales over total car sales one month before?

April May July August November

INTELLIGENCE TEST

This test lasts 45 minutes, and is intended to measure intelligence, accuracy, and common sense. It consists of 3 separately-timed parts, each lasting 15 minutes and each containing short-answer problems.

PART 1

In this part you have to solve 17 problems. You do not need any knowledge of arithmetical formulae or algebra.

EXAMPLES

1. Which one of the following words would look the same if the 2 middle letters were turned from top to bottom?

1. KILT 2. AXIS 3. DROP 4. FACT 5. SHUN

Answer:

2. Five cars, L, M, N, O, and P, are travelling along a motorway in that order with L leading. If P moves in between O and N, and L changes places with O, and then M changes places with N, what is their new order?

Answer:

3. If Kate's birthday is September 20 and the following Wednesday week will be October 1, on what day of the week is Kate's birthday?

Answer:

4. If OXFI is the word "rail" in code, what is the word "bus" in the same code?

Answer:

5. The first act in a variety show was a minute late in starting, and this and every subsequent act lasted a minute longer than anticipated. If there was no interval between the acts and the last act was 13 minutes late in starting, how many acts were there?

Answer:

6. The 3 words below are in code. One stands for "sting", one for "roads", and one for "giver". Write the decoded words in the appropriate places below.

1.	JWTYK	*Answers:*	1. .....
2.	ADPBJ		2. .....
3.	KMDGA		3. .....

7. Each chair in a room has 2 cushions on it, and each cushion is covered in only one colour of cloth. One third of the cushions are blue, one half are red, and the remaining 5 cushions are green. How many chairs are there in the room?

Answer:

8. Alan, Brenda, Colin, Dora, and Eric each have only one car, and their 5 cars include at least one in each of the following colours: red, yellow, green, brown. Alan's car is a yellow Ford. The only Morris is green. Brenda's car is the same colour as Dora's and the same make as Eric's. Colin's car is brown. What is the colour of Eric's car?

Answer:

9. A, B, C, D, and E are 5 ropes. Each is exactly a whole number of metres in length, and none of them is the same length as any other of them. A is as long as B (the shortest rope) and C put together, and C, D, and E together are one metre more than twice the length of A. If B is 3 metres long, C is 4 metres long, and D is longer than E, how long is E?

Answer:

PART 2

In this part there are 30 sentences, in each of which 2 words need to be interchanged in order to make the sentence read sensibly. You have to pick out these words and put a tick through each.

EXAMPLES

1. After discussing Lord Trent's objections to the proposal, the statement issued a meeting recording their decision to postpone any action on this matter.

2. Some planning developments permit a householder to carry out whatever authorities he wishes, irrespective of the possible dangers to himself and to others.

3. Intense meditation is commonly experienced before speaking in public, but anxiety can help overcome this problem.

4. American experts at the conference argued that carrots contain a previously unknown body which is not digested by the human substance.

5. Noticing Jane's discomfort, he filled her glass with an acute drink to relieve the effervescent feeling of dryness in her throat.

6. The administrators have continually to seek authority from a mundane level even for such higher matters as work and supplies.

7. In view of all these sources of friction, spokesmen of differing civil sympathies concur in pessimism and suggest that political war is a real possibility.

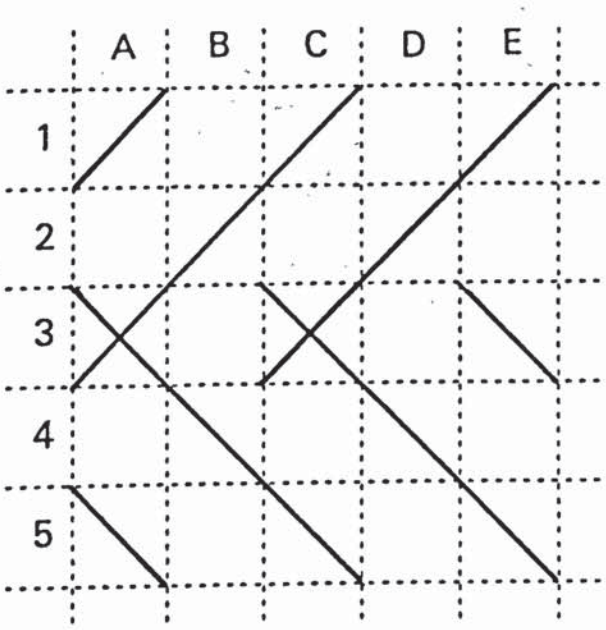
8. The great scholar who is a good leader is necessarily equipped but the two do not doubly go together.

9. Naturally towns present special problems: in the past our villages have grown slowly and planned to take the shape of larger units.

PART 3

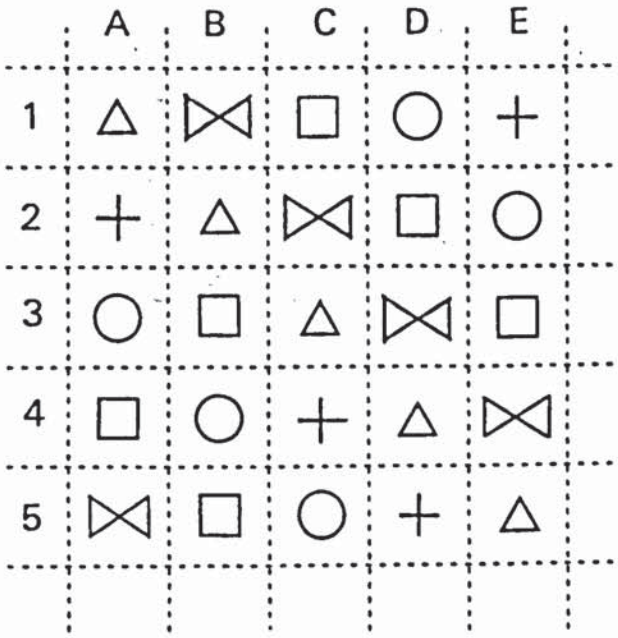
In this part each of the 27 questions consists of a design, or part of a pattern, covering a number of squares: in one square the design is incorrect. You have to find this square and write its row number and column letter as your answer.

EXAMPLES



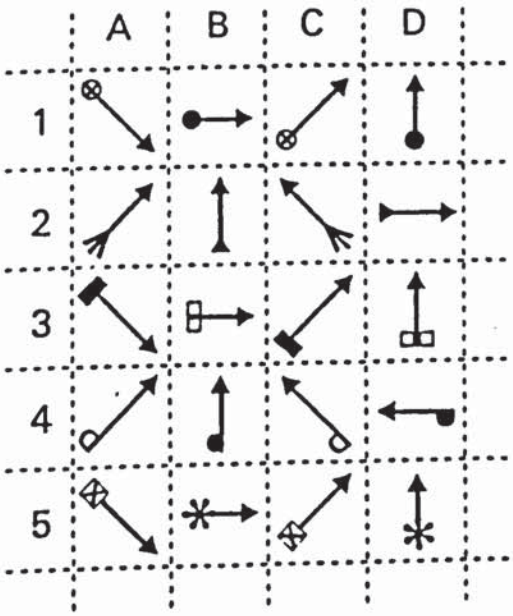
Answer:

(2)



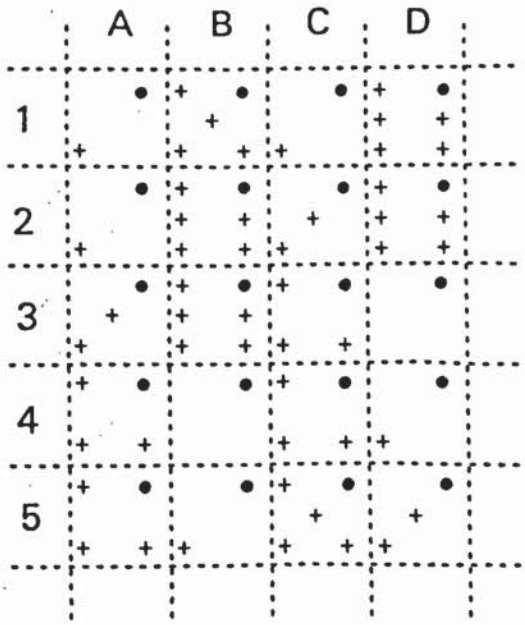
Answer:

(3)



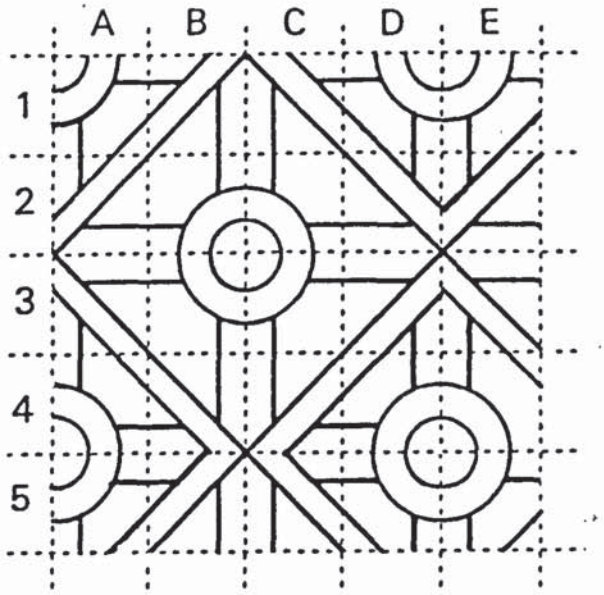
Answer:

(4)



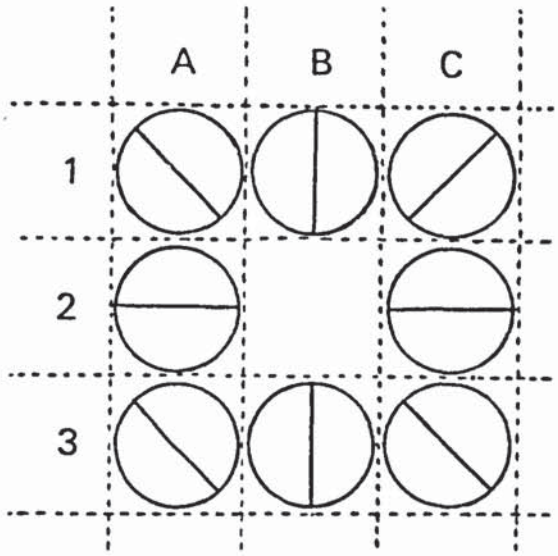
Answer:

(5)



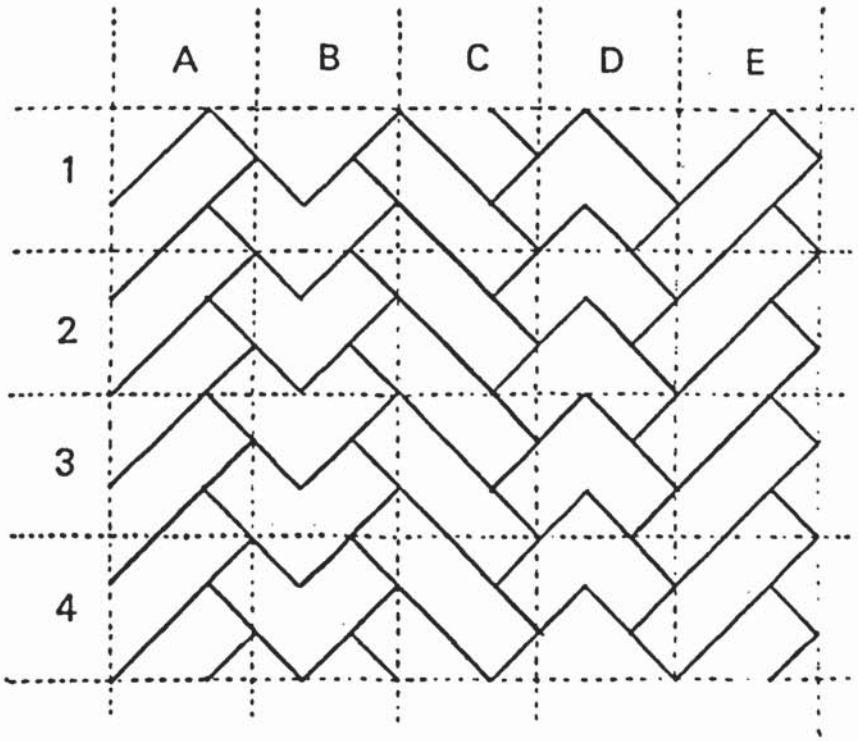
Answer:

(6)



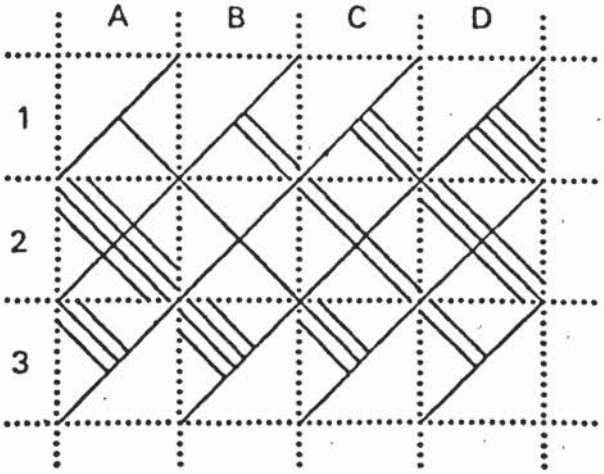
Answer:

(7)



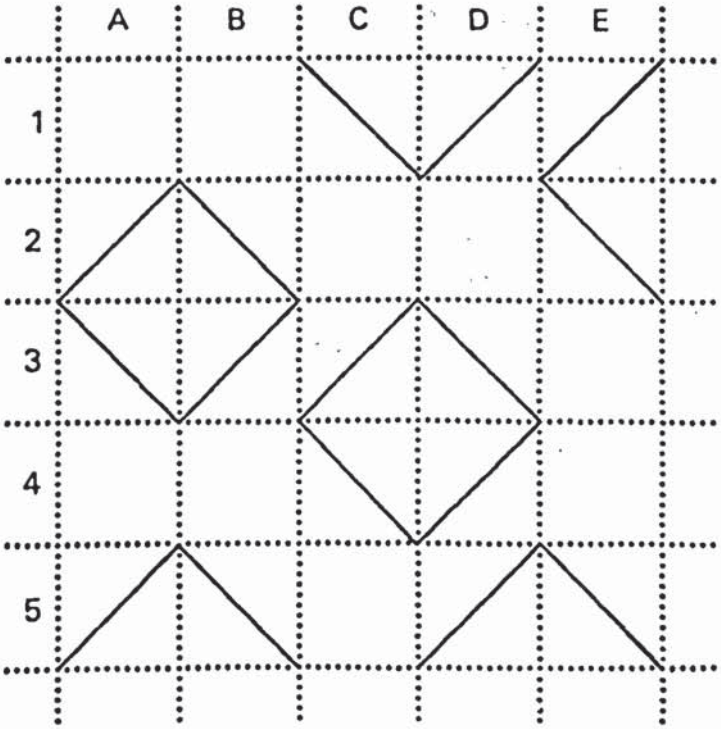
Answer:

(8)



Answer:

(9)



Answer:

EXECUTIVE PROBLEM TEST

This test lasts 1½ hours and consists of 20 short-answer problems of the kind that can arise in administration.

EXAMPLES

1. You have to arrange bookings for the use of a tennis court. In a particular week, there are 8 pairs of players who each wish to book it for a morning or an afternoon of only one day, and no pair wants to play tennis with any other pair. The table shows the names of the players in each pair, whether they want to use the court in the morning or in the afternoon, and the days which would suit them.

Table

Pair of players	*Morning or afternoon required*	*Suitable days*
Mr Brown and Mr Black	Morning only	Tuesday
Mr Lipman and Miss Johnston	Morning or afternoon	Wednesday
Mr Kamul and Miss Jackson	Morning or afternoon	Tuesday
Mr Jones and Mr Smith	Afternoon only	Monday
Miss Kerr and Mrs Ruddell	Afternoon only	Wednesday
Miss O'Shea and Miss Arrell	Morning only	Friday
Miss Allen and Miss Reilly	Morning or afternoon	Monday
Mr Thompson and Mrs Pope	Morning or afternoon	Friday

In order that you may fit them all in, will you offer morning or afternoon to each of the 4 pairs who have expressed no preference for one period over the other?

Answer: Mr Lipman and Miss Johnston ..... Mr Kamul and Miss Jackson .....
Miss Allen and Miss Reilly ..... Mr Thompson and Mrs Pope .....

2. You are a school secretary making arrangements for the annual outing. The children have been given details of 5 possible outings, called A, B, C, D, and E, and have been asked to place them in their order of preference. When the orders of preference were scrutinized it was found that they could be arranged in 7 groups. The table shows the order of preference of each of the 7 groups of children. It will not be possible to arrange 5 different outings; but it will be possible to arrange 2, and the headmaster has told you to choose the 2 which will permit every child to have either his first or his second choice. In order to do this, which 2 will you choose?

Table

Group of children	*Outings in order of preference*				
1	B	A	C	E	D
2	E	D	B	C	A
3	A	B	E	C	D
4	C	B	A	E	D
5	D	E	C	A	B
6	A	B	D	C	E
7	A	D	B	E	C

Answer: ..... and .....

3. Imagine that you work in a public library. The readers whose names are shown in the table below have asked for the books indicated against their names. Where a book has been requested in this way by more than one reader, it is the library's policy not to lend any reader more than one of the books which he or she has requested if that would mean that another reader could not have any of the books for which he or she has asked. In order that each of the readers shown may have one of the books which he or she has requested, which book will you lend to Mrs Quinn?

Table

Name of reader	*Books requested*				
Mrs Nye	A	B	C	E	F
Miss Orr	B	D	G		
Mr Price	B	D	G		
Mrs Quinn	B	D	E	G	
Miss Robinson	A	C	D	E	F
Mr Smith	A	C	E	F	G
Mrs Taylor	B	G			

Answer: .....

4. You are the head of a large section in an office and you are allowed £30,000 per year to buy essential computer equipment. You need 9 items and their costs (in thousands of pounds) are as follows:

18, 7, 7, 20, 5, 4, 7, 6, 16

You notice that your budget will enable you to buy these 9 items over a period of 3 years, but in each year you must spend exactly the £30,000 allowed for that year. What are the costs (in thousands of pounds) of the other 2 items which you will buy in the year in which you get the item costing £4,000? (Assume that there is no inflation taking place.)

Answer: 4, ... and .....

5. You are the manager of a warehouse, and have to make arrangements for receiving a consignment of household goods. The warehouse is constructed in such a manner that along each side of a central gangway, which runs the full length of the building, there are 4 storage bays. These bays are lettered A, B, C, and D on the left-hand side, and E, F, G, and H on the right-hand side. Each bay is used to store one particular product, and no other. It is not possible to unload goods on one side of the gangway and transfer them to a bay on the other side, because the arrangement of the lifting gear will not permit this. The gang of men working at the warehouse can unload 4 lorry-loads of any kind of goods each day and store them in the correct bay.

A lorry can carry 18 cubic metres of goods. A lorry-load is always made up of one single product, and the lorry is loaded as near to full capacity as is possible. The following table gives information about the goods, including the amount which you must now take into the warehouse.

Table

Item	*Volume of each crated item (cubic metres)*	*Number of items to be received*	*Storage bay letter*
Washing machines	0.60	87	H
Large freezers	1.25	42	C
Medium freezers	0.90	60	D
Slim-line refrigerators	0.33	108	F
Normal refrigerators	0.28	64	G
Gas cooker with eye-level grill	0.56	96	A
Gas cooker with low-level grill	0.33	162	B
Electric cooker	0.33	108	E

What is the minimum number of days over which you must spread the unloading of the goods for:

a. bays A to D, and
b. bays E to H?

Answer: a. ..... b. .....

6. The Government of the island of Inogu has established a tourist office through which foreign tourists can book their accommodation on the island. But the firms which cater for tourists have asked that, instead of being situated in the capital, the tourist office should operate through a number of local offices in the main resorts, so that its staff will be on hand to deal with queries and complaints from tourists. The Government has agreed to try out a pilot scheme on these lines in 2 of the main resorts next year, provided that it is financed by the tourist industry. It has been agreed that every hotel will contribute £25 towards the cost of the scheme, every guest house £15, and every private house which takes in visitors £6; and those which are situated in the resorts selected to have the local offices will contribute twice these amounts. The total contributions from all the towns must cover the rent of the 2 local offices and the salaries of the 2 additional staff (one in each office) who will be needed to man them. These salaries are £1,200 each. The table shows the names of the 6 holiday resorts, the numbers of hotels, guest houses, and private houses offering accommodation in each, and the cost of renting a suitable office in each. You have to choose the 2 resorts to have the local tourist offices. You have to make ends meet financially, and subject to this you should choose resorts making large financial contributions in preference to those making small ones. Which 2 resorts will you choose?

Table

Resort	*Number of hotels*	*Number of guest houses*	*Number of private houses*	*Rent of office*
Jakara	4	16	26	£400
Kabalu	2	7	20	£150
Lakeija	3	14	18	£250
Morogu	2	6	10	£150
Nikasta	4	15	10	£350
Orampo	4	12	14	£200

Answer: ..... and .....

7. You have to arrange a teaching schedule, using one room, for a 3-day period, a Monday, a Tuesday, and a Wednesday. The time available is from 0900 to 1200 hours and from 1400 to 1700 hours on each day, and 2 sessions of each of the following subjects are to be given:

English (2- hour subject)
French (2-hour subject)
German (2-hour subject)
Biology (1-hour subject)
Geography (1-hour subject)
History (1-hour subject)

You have 6 instructors, lettered from A to F, who teach as follows:

A... English and History
B... French and Geography
C... German and Biology
D... English and Geography
E... French and Biology
F... German and History

Each instructor is to give one session of each of his 2 subjects. No subject may be given twice on the same day, and no instructor may give both his subjects on the same day.

Instructors A and D are not available on Tuesday; instructors B and F are not available on Wednesday.

You decide to arrange each day's sessions in alphabetical order according to subject. You further decide to hold the History sessions on Monday and Tuesday. What do you put down as the first subject on Wednesday afternoon, and which instructor do you tell to give it?

Answer: Subject ..... Instructor .....

ANSWERS

Enquiries from Management

Example 1

1–A: 2–16½: 3–B: 4–E: 5–D.

Example 2

1–E: 2–C: 3–A: 4–E: 5–400%.

Example 3

1–422: 2–1979: 3–4: 4–3%: 5–1:5.

Example 4

1–D: 2–A and E: 3–10: 4–1/5: 5–70.

Example 5

1–Mr Patel: 2–Mr Smith: 3–July: 4–May: 5–August.

Intelligence Test

Part 1

1–2: 2–ONMPL: 3–Saturday: 4–YRP: 5–13: 6–ROADS, GIVER, STING: 7–15: 8–Red: 9–5 metres.

Part 2

1–statement, meeting: 2–developments, authorities: 3–meditation, anxiety: 4–body, substance: 5–acute, effervescent: 6–mundane, higher: 7–civil, political: 8–necessarily, doubly: 9–naturally, planned.

Part 3

1–3E: 2–3B: 3–2D: 4–5D: 5–4E: 6–3A: 7–1D: 8–3C: 9–1A.

Executive Problem Test

1–Mr Lipman and Miss Johnston MORNING, Mr Kamul and Miss Jackson AFTERNOON, Miss Allen and Miss Reilly MORNING, Mr Thompson and Mrs Pope AFTERNOON: 2–B and D: 3–E: 4–20 and 6: 5–a.3, b.2: 6–Lakeija and Orampo: 7–French, E.

(650) Dd8201187 40M 4/81 HPLtd So'ton G3927

APPENDIX 2

Questionnaire Given to Respondents

of Pilot Study I on

Suitability/Objectionability of Classifications

1. Please state the country of birth of
   (i) yourself
   (ii) your mother
   (iii) your father.

2. Using the level of two 'A' levels and five 'O' levels as a guide, would you say that your educational qualifications are probably

   lower than that level? 1☐
   greater than or equal to that level? 2☐

3. My age is between

   16-20 1☐
   21-30 2☐
   31-40 3☐
   41-50 4☐
   51-60 5☐
   60+ 6☐

4. Sex

   M 1☐
   F 2☐

Col. nos

1. Please tick the appropriate box to show the racial or ethnic group to which you belong.

English, Welsh, Scottish, Irish	01 ☐	
Italian	02 ☐	
Greek/Greek Cypriot	03 ☐	
Turkish/Turkish Cypriot	04 ☐	
Polish	05 ☐	
Other European	06 ☐	13-14
West Indian or Guyanese	07 ☐	
African	08 ☐	
Indian	09 ☐	
Pakistani	10 ☐	
Bangladeshi	11 ☐	
Arab	12 ☐	
Chinese	13 ☐	
Other non-European (including mixed non-European descent)		

Please state 14

For the purposes for which it is to be used, I feel that this classification is

Very suitable ☐ ☐ ☐ ☐ ☐ Not suitable 15
1 2 3 4 5

This type of question is

Very suitable ☐ ☐ ☐ ☐ ☐ Not suitable 16
1 2 3 4 5

If an employer were to present me with a classification of this sort I would object to it

Not at all ☐ ☐ ☐ ☐ ☐ very much 17
1 2 3 4 5

How do you think this classification could be improved ?

Col. No.

2. I would describe my ethinic origins as

African*	1 ☐	
Asian	2 ☐	
Caribbean**	3 ☐	18
UK, European, Irish	4 ☐	
Other European	5 ☐	

Other countries - please state .....

For the purposes for which it is to be used I feel that this classification is

Very suitable ☐ ☐ ☐ ☐ ☐ Not suitable 19
1 2 3 4 5

This type of question is

Very suitable ☐ ☐ ☐ ☐ ☐ Not suitable 20
1 2 3 4 5

If an employer were to present me with a classification of this sort I would object to it

Not at all ☐ ☐ ☐ ☐ ☐ very much 21
1 2 3 4 5

How do you think this classification could be improved ?

* This category is not intended to include West Indians who may identify their origins as African. Equally, it should not include Europeans or Asians whose families have settled in African countries.

** This category is intended to include Asians who were born in the West Indies, or whose recent forebears were born there.

Col. No.:

3. I would describe myself as

Coloured [ ] 1 White [ ] 2 22

For the purposes for which it is to be used, I feel that this classification is

Very suitable [ ] [ ] [ ] [ ] [ ] Not suitable 23

1 2 3 4 5

This type of wuestion is

Very suitable [ ] [ ] [ ] [ ] [ ] Not suitable 24

1 2 3 4 5

If an employer were to present me with a classification of this sort I would object to it

Not at all [ ] [ ] [ ] [ ] [ ] Very much 25

1 2 3 4 5

How do you think this classification could be improved ?

Col. Nos

4. I would describe myself as

British	01	☐
British Asian	02	☐
British West Indian	03	☐
British African	04	☐
British Chinese	05	☐
African Asian	06	☐
Asian	07	☐
West Indian	08	☐
African	09	☐
European	10	☐
Chinese	11	☐
Other - please state	12	

26-27

For the purposes for which it is to be used, I feel that this classification is

Very suitable ☐ ☐ ☐ ☐ ☐ Not suitable
1 2 3 4 5

28

This type of question is

Very suitable ☐ ☐ ☐ ☐ ☐ Not suitable
1 2 3 4 5

29

If an employer were to present me with a classification of this sort I would object to it

Not at all ☐ ☐ ☐ ☐ ☐ Very much
1 2 3 4 5

30

How do you think this classification could be improved ?

Col. Nos

5. I would describe my racial group as

Caucasian	1	☐
Negroid	2	☐
Mongoloid	3	☐

31

For the purposes for which it is to be used, I feel that this classification is

Very suitable					Not suitable
1	2	3	4	5	

32

This type of question is

Very suitable					Not suitable
1	2	3	4	5	

33

If an employer were to present me with a classification of this sort I would object to it

Not at all					Very much
1	2	3	4	5	

34

How do you think this classification could be improved ?

Col Nos.

6. From your name and visual appearance only, what do you think another person might assume your racial or ethnic origins to be ?

English Welsh, Scottish, Irish	01	
Italian	02	
Greek/Greek Cypriot	03	
Turkish/Turkish Cypriot	04	
Polish	05	
Other European	06	
West Indian or Guyanese	07	
African	08	
Indian	09	
Pakistani	10	
Bangladeshi	11	
Arab	12	
Chinese	13	
Other non-European (including mixed non-European descent)		

Please state .....14....

35-36

For the purposes for which it is to be used, I feel that this classification is

Very suitable [1] [2] [3] [4] [5] Not suitable

37

This type of question is

Very suitable [1] [2] [3] [4] [5] Not suitable

38

If an employer were to present me with a classification of this sort I would object to it

Not at all [1] [2] [3] [4] [5] very much

39.

How do you think this classification could be improved ?

Col. No.

7. From your name and visual appearance only, what do you think another person might assume your racial or ethnic origins to be ?

African	1 ☐
Asian	2 ☐
Caribbean	3 ☐
UK, European, Irish	4 ☐
Other European	5 ☐
Other countries - please state	6

40

For the purposes for which it is to be used, I feel that this classification is

Very suitable ☐ ☐ ☐ ☐ ☐ Not suitable

1 2 3 4 5

41

This type of question is

Very suitable ☐ ☐ ☐ ☐ ☐ Not suitable

1 2 3 4 5

42

If an employer were to present me with a classification of this sort I would object to it

Not at all ☐ ☐ ☐ ☐ ☐ Very much

1 2 3 4 5

43

How do you think this classification could be improved ?

Col. Nos.

8. Do you think another person would say you are

Coloured ☐ 1 White ☐ 2. 44

For the purposes for which it is to be used I feel that this classification is

Very suitable ☐ ☐ ☐ ☐ ☐ Not suitable 45

1 2 3 4 5

This type of question is

Very suitable ☐ ☐ ☐ ☐ ☐ Not suitable 46

1 2 3 4 5

If an employer were to present me with a classification of this sort I would object to it

Not at all ☐ ☐ ☐ ☐ ☐ Very much 47.

1 2 3 4 5

How do you think this classification could be improved ?

Ns.5

9. From your name and visual appearance, what do you think another person would assume you to be ?

British	01	☐
British Asian	02	☐
British West Indian	03	☐
British African	04	☐
British Chinese	05	☐
African Asian	06	☐
Asian	07	☐
West Indian	08	☐
European	09	☐
Chinese	10	☐

Other - please state ..... 11 ... 48-49

For the purposes for which it is to be used I feel that this classification is

Very suitable					Not suitable
1	2	3	4	5	

50

This type of question is

Very suitable					Not suitable
1	2	3	4	5	

51

If an employer were to present me with a classification of this sort I would object to it

Not at all					very much
1	2	3	4	5	

52

How do you think this classification could be improved ?

10. From your visual appearance, what do you think another person would assume to be your racial grouping ?

Caucasian	1	☐
Negroid	2	☐
Mongoloid	3	☐

53

For the purposes for which it is to be used I feel that this classification is

Very suitable					Not suitable
1	2	3	4	5	

54

This type of question is

Very suitable					Not suitable
1	2	3	4	5	

55

If an employer were to present me with a classification of this sort I would object to it

Not at all					Very much
1	2	3	4	5	

56

How do you think this classification could be improved.?

APPENDIX 3

ANALYSIS OF VARIANCE TABLES FOR SUITABILITY OF CLASSIFICATION

(A) FOR TOTAL SAMPLE

SOURCE	df	SS	MS	F	
Between groups	9	260.8	28.98	25.11	
Within groups	231	1860.8	8.05		$p < 0.01$
Error	2079	2399.2	1.15		
Total	2319	4520.8			

(B) FOR UNIVERSITY SAMPLE

SOURCE	df	SS	MS	F	
Between groups	9	136.3	15.14	13.37	
Within groups	108	843.7	7.81		$p < 0.01$
Error	972	1101.4	1.13		
Total	1089	2081.5			

(C) FOR SCHOOL LEAVER/NURSES SAMPLE

SOURCE	df	SS	MS	F	
Between groups	9	174.9	19.43	17.11	
Within groups	122	1013.1	8.30		$p < 0.01$
Error	1098	1247.4	1.14		
Total	1229	2435.4			

APPENDIX 4

Significant Differences Between Means for University Sample on Suitability of Classifications

		STANDARD QUESTION					REVERSED QUESTION				
		OPCS	CRE	TAVI	MCB	ANTH	OPCS	CRE	TAVI	MCB	ANTH
OPCS	SQ		NS	**	NS	**	NS	NS	**	*	**
CRE	SQ			**	NS	**	**	**	**	**	**
TAVI	SQ				**	NS	**	**	*	**	*
MCB	SQ					**	*	NS	**	**	**
ANTH	SQ						**	**	NS	**	NS
OPCS	RQ							NS	**	NS	NS
CRE	RQ								**	NS	**
TAVI	RQ									**	NS
MCB	RQ										NS
ANTH	RQ										

Significant Differences Between Means for School Leavers/ Nurses Sample on Suitability of Classifications

		STANDARD QUESTION					REVERSED QUESTION				
		OPCS	CRE	TAVI	MCB	ANTH	OPCS	CRE	TAVI	MCB	ANTH
OPCS	SQ		**	**	NS	**	**	**	**	**	**
CRE	SQ			NS	*	**	NS	NS	NS	NS	**
TAVI	SQ				**	**	NS	NS	NS	NS	**
MCB	SQ					**	**	**	**	**	**
ANTH	SQ						**	**	**	**	NS
OPCS	RQ							NS	NS	NS	**
CRE	RQ								NS	NS	**
TAVI	RQ									NS	**
MCB	RQ										**
ANTH	RQ										

APPENDIX 5

Mean Scores By Ethnic Group For Opinion Measures of Suitability and Objectionability

(a) Suitability of Classification

		UK	West Indian	Indian
OPCS	Standard	2.31	2.78	2.80
CRE	Standard	2.79	2.83	2.78
TAVI	Standard	2.99	3.43	3.42
MCB	Standard	2.71	2.55	2.60
ANTH	Standard	3.54	4.46	3.71
OPCS	Reversed	3.00	3.09	3.19
CRE	Reversed	3.03	3.26	3.04
TAVI	Reversed	2.86	3.42	3.59
MCB	Reversed	2.88	3.24	3.22
ANTH	Reversed	3.31	4.24	3.62

(b) Suitability of Question

		UK	West Indian	Indian
OPCS	Standard	2.41	2.87	2.97
CRE	Standard	2.91	2.95	3.06
TAVI	Standard	3.18	3.47	3.64
MCB	Standard	3.01	2.68	2.58
ANTH	Standard	3.57	4.49	3.75
OPCS	Reversed	3.20	3.29	3.36
CRE	Reversed	3.31	3.41	3.25
TAVI	Reversed	3.14	3.64	3.69
MCB	Reversed	2.93	3.26	3.23
ANTH	Reversed	3.31	4.26	3.68

(c) Objections to Classifications

		UK	West Indian	Indian
OPCS	Standard	2.29	2.74	2.88
CRE	Standard	2.74	2.78	2.72
TAVI	Standard	3.11	3.41	3.54
MCB	Standard	2.77	2.55	2.52
ANTH	Standard	3.36	4.00	3.52
OPCS	Reversed	2.84	3.00	3.15
CRE	Reversed	3.03	3.24	3.12
TAVI	Reversed	2.88	3.40	3.43
MCB	Reversed	2.92	3.16	3.15
ANTH	Reversed	3.22	3.89	3.49

APPENDIX 6

ANALYSIS OF VARIANCE TABLES FOR SUITABILITY OF QUESTION

(A) FOR TOTAL SAMPLE

SOURCE	df	SS	MS	F	
Between groups	9	229.0	25.44	23.68	
Within groups	231	1878.1	8.13		$p < 0.01$
Error	2079	2233.9	1.07		
Total	2319	4341.1			

(B) FOR UNIVERSITY SAMPLE

SOURCE	df	SS	MS	F	
Between groups	9	131.2	14.57	14.31	
Within groups	108	859.7	7.96		$p < .01$
Error	972	990.3	1.01		
Total	1089	1981.2			

(C) FOR SCHOOL LEAVER/NURSES SAMPLE

SOURCE	df	SS	MS	F	
Between groups	9	141.4	15.74	14.40	
Within groups	122	1013.2	8.31		$p < .01$
Error	1098	1199.8	1.09		
Total	1229	2354.8			

APPENDIX 7

Significant Differences Between Means for School Leavers/ Nurses Sample on Suitability of Question

		STANDARD QUESTION					REVERSED QUESTION				
		OPCS	CRE	TAVI	MCB	ANTH	OPCS	CRE	TAVI	MCB	ANTH
OPCS	SQ		*	**	NS	**	**	**	**	**	**
CRE	SQ			*	NS	**	NS	**	*	NS	**
TAVI	SQ				**	**	NS	NS	NS	NS	**
MCB	SQ					**	**	**	**	*	**
ANTH	SQ						**	**	**	**	NS
OPCS	RQ							NS	NS	NS	**
CRE	RQ								NS	NS	**
TAVI	RQ									NS	**
MCB	RQ										**
ANTH	RQ										

Significant Differences Between Means for University Sample on Suitability of Question

		STANDARD QUESTION					REVERSED QUESTION				
		OPCS	CRE	TAVI	MCB	ANTH	OPCS	CRE	TAVI	MCB	ANTH
OPCS	SQ		**	**	*	**	NS	NS	**	NS	**
CRE	SQ			**	NS	**	**	**	**	**	**
TAVI	SQ				**	NS	**	**	NS	**	*
MCB	SQ					**	NS	NS	**	**	**
ANTH	SQ						**	**	NS	**	NS
OPCS	RQ							NS	**	NS	**
CRE	RQ								**	NS	**
TAVI	RQ									*	NS
MCB	RQ										NS
ANTH	RQ										

NS - Not Significant * - $p < .05$ ** - $p < .01$

APPENDIX 8

ANALYSIS OF VARIANCE TABLES FOR EXTENT OF OBJECTIONS TO CLASSIFICATION

(A) FOR TOTAL SAMPLE

SOURCE	df	SS	MS	F	
Between groups	9	225.8	25.09	24.04	
Within groups	231	2675.8	11.58		$p < 0.01$
Error	2079	2169.9	1.04		
Total	2319	5071.5			

(B) FOR UNIVERSITY SAMPLE

SOURCE	df	SS	MS	F	
Between groups	9	136.9	15.21	15.20	
Within groups	108	1170.9	10.84		$p < 0.01$
Error	972	972.4	1.00		
Total	1089	2280.3			

(C) FOR SCHOOL LEAVER/NURSES SAMPLE

SOURCE	df	SS	MS	F	
Between groups	9	129.8	14.42	13.67	
Within groups	122	1498.9	12.28		$p < 0.01$
Error	1098	1156.6	1.05		
Total	1229	2785.4			

APPENDIX 9

Significant Differences Between Means for University Sample on Objectionability to Classification

		STANDARD QUESTION					REVERSED QUESTION				
		OPCS	CRE	TAVI	MCB	ANTH	OPCS	CRE	TAVI	MCB	ANTH
OPCS	SQ		*	**	*	**	*	**	**	**	**
CRE	SQ			**	NS	**	**	**	**	**	**
TAVI	SQ				**	*	**	**	NS	**	**
MCB	SQ					**	*	**	**	**	**
ANTH	SQ						**	**	NS	**	NS
OPCS	RQ							NS	**	NS	*
CRE	RQ								**	NS	*
TAVI	RQ									**	NS
MCB	RQ										NS
ANTH	RQ										

Significant Differences Between School Leavers/ Nurses Sample on Objectionability to Classification

		STANDARD QUESTION					REVERSED QUESTION				
		OPCS	CRE	TAVI	MCB	ANTH	OPCS	CRE	TAVI	MCB	ANTH
OPCS	SQ		NS	**	NS	**	**	**	**	**	**
CRE	SQ			**	NS	**	NS	**	*	*	**
TAVI	SQ				**	**	NS	NS	NS	NS	**
MCB	SQ					**	NS	**	**	**	**
ANTH	SQ						**	**	**	**	NS
OPCS	RQ							NS	NS	NS	**
CRE	RQ								NS	NS	**
TAVI	RQ									NS	**
MCB	RQ										**
ANTH	RQ										

NS - Not Significant * - $p < .05$ ** - $p < .01$

APPENDIX 10

Questionnaire Given to Respondents of Pilot Study II - Test-Retest Reliability

1. Please state the country of birth of

(i) yourself 4-

(ii) your mother 6-

(iii) your father. 8-

2. Using the level of two 'A' levels and five 'O' levels as a guide, would you say that your educational qualifications are probably

lower than that level? 1☐ 10

greater than or equal to that level? 2☐

3. My age is between

16-20	1☐
21-30	2☐
31-40	3☐
41-50	4☐
51-60	5☐
60+	6☐

11

4. Sex

M	1☐
F	2☐

12

5. Your date of birth _____

For each classification tick in (i) column a, the appropriate box to show the racial or ethnic group to which you belong, (ii) column b, the racial or ethnic group another person, e.g. an interviewer, might assume you to be from your name and visual appearance.

CLASSIFICATION 1

	Column a	Column b	
English, Welsh, Scottish, Irish	01 ☐	01 ☐	
Italian	02 ☐	02 ☐	13-14
Greek/Greek Cypriot	03 ☐	03 ☐	
Turkish/Turkish Cypriot	04 ☐	04 ☐	
Polish	05 ☐	05 ☐	
Other European	06 ☐	06 ☐	
West Indian or Guyanese	07 ☐	07 ☐	15-16
African	08 ☐	08 ☐	
Indian	09 ☐	09 ☐	
Pakistani	10 ☐	10 ☐	
Bangladeshi	11 ☐	11 ☐	
Arab	12 ☐	12 ☐	
Chinese	13 ☐	13 ☐	
Other non-European (including mixed non-European descent)			

Please state 14 ..... 14 .....

CLASSIFICATION 2

	Column a	Column b	
African	1 ☐	1 ☐	17
Asian	2 ☐	2 ☐	
Caribbean	3 ☐	3 ☐	18
UK, European, Irish	4 ☐	4 ☐	
Other European	5 ☐	5 ☐	
Other countries - please state 6		6	

Interviewer code

CLASSIFICATION 3

	column a	column b
Coloured	1	1
White	2	2

75

CLASSIFICATION 4

	column a	column b
British	01	01
British Asian	02	02
British West Indian	03	03
British African	04	04
British Chinese	05	05
African Asian	06	06
Asian	07	07
West Indian	08	08
African	09	09
European	10	10
Chinese	11	11
Other - please state	12	12

76-77

CLASSIFICATION 5

	column a	column b
Caucasian	1	1
Negroid	2	2
Mongoloid	3	3

78

APPENDIX 11

Questionnaire Given to Respondents in Stage II of Pilot Studies

1. Please state the country of birth of

   (i) yourself
   (ii) your mother
   (iii) your father

2. Using the level of two 'A' levels and five 'O' levels as a guide, would you say that your educational qualifications are probably

   lower than that level?
   greater than or equal to that level?

1. I would describe my ethnic origins as

☐ African	ie. persons born in, or whose recent forebears were born in Africa. This category is not intended to include West Indians who may identify their origins as African or Europeans or Asians whose families settled in Africa.
☐ Asian	ie. persons born in or whose recent forebears were born in the Caribbean. This category is not intended to include Asians who were born in the West Indies or whose recent forebears were born there.
☐ UK, European and Irish	ie. English, Welsh, Scottish, Irish.
☐ Other European	ie. excluding the UK and Eire, eg. Cyprus, Eastern Europe, EEC countries, Spain, Malta, etc.
☐ Other	ie. persons born in or originating from countries not included in the above categories, including people of mixed descent who choose not to classify themselves elsewhere.

2. For the purposes for which it is to be used, I feel that this classification is

very suitable	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	not suitable

3. If an employer were to present me with a classification of this sort, I would object to it

not at all	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	very much

4. How do you think this classification could be improved?

1. Please tick the appropriate box to show the racial or ethnic group to which you belong.

White	☐
West Indian	☐
African	☐
Indian	☐
Pakistani	☐
Bangladeshi	☐
Arab	☐
Chinese	☐

Other (including mixed descent)
please state _____

2. For the purposes for which it is to be used, I feel that this classification is

very suitable	☐	☐	☐	☐	☐	not suitable
	1	2	3	4	5	

3. If an employer were to present me with a classification of this sort, I would object to it

not at all	☐	☐	☐	☐	☐	very much
	1	2	3	4	5	

4. How do you think this classification could be improved?

1. Please tick the appropriate box to show the racial or ethnic group to which your forebears belong (tick one box only).

English, Welsh, Scottish, Irish ☐

This category is intended only for persons whose forebears originally came from these countries. If your forebears' origins are other than these please tick the appropriate box below.

Italian ☐
Greek/Greek-Cypriot ☐
Turkish/Turkish-Cypriot ☐
Polish ☐
West Indian ☐
Rastafarian ☐
Nigerian ☐
Ghanaian ☐
Indian ☐
Pakistani ☐
Bangladeshi ☐
Iranian ☐
Chinese ☐
Other (including mixed descent) please state

2. For the purposes for which it is to be used, I feel that this classification is

very suitable ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 not suitable

3. If an employer were to present me with a classification of this sort, I would object to it

not at all ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 very much

4. How do you think this classification could be improved?

1. Please tick the appropriate box below to show which of the groups you belong to (tick one box only).

British	01	☐	ie. persons who were born in England, Wales, Scotland or Ireland AND whose forebears originally came from these countries.
British Asian	02	☐	ie. persons born in Britain but whose forebears came from India, Pakistan, Bangladesh or Sri Lanka. This category is not intended to include Asians born in African countries.
British West Indian	03	☐	ie. persons born in Britain but whose forebears came from the West Indies.
British Chinese	04	☐	ie. persons born in Britain but whose forebears came from China.
Other British - please state			_____ ie. persons born in Britain but whose forebears came from areas other than those described above.
African Asian	05	☐	ie. persons born in Africa but whose forebears came from India, Pakistan, Bangladesh, or Sri Lanka.
Asian	06	☐	ie. persons born in India, Pakistan, Bangladesh or Sri Lanka.
West Indian	07	☐	ie. persons born in the West Indies.
African	08	☐	ie. persons born in Africa.
European	09	☐	ie. persons born in European countries other than England, Wales, Scotland and Ireland eg. EEC countries, Eastern Europe, Cyprus, Spain, Malta, etc.
Chinese	10	☐	ie. persons born in South-East Asia of Chinese origins.
Other - please state			_____ ie. persons born in countries not included in the above categories including people of mixed descent.

2. For the purpose for which it is to be used I feel that this classification is

very suitable ☐ ☐ ☐ ☐ ☐ not suitable
1 2 3 4 5

3. If an employer were to present me with a classification of this sort I would object to it

not at all ☐ ☐ ☐ ☐ ☐ very much
1 2 3 4 5

4. How do you think this classification could be improved?

APPENDIX 12

Analysis of Variance Tables For Objections To The Classifications

(a) For Total Sample

Source	df	SS	MS	F	
Between groups	3	25.50	8.5	8.94	
Within groups	168	1163.75	6.93		$p<.01$
Error	504	479.50	.95		
Total	675	1668.75			

(b) For University Sample

Source	df	SS	MS	F	
Between groups	3	13.79	4.60	1.63	NS
Within groups	109	129.12	1.19		
Error	327	922.71	2.82		
Total	439	1065.62			

(c) For School-Leavers Sample

Source	df	SS	MS	F	
Between groups	3	11.63	3.88	6.36	
Within groups	58	389.99	6.72		$p<.01$
Error	174	106.62	.61		
Total	235	508.24			

APPENDIX 13

Mean Scores By Ethnic Group For Opinion Measures of Objections to and Suitability of the Classifications

(a) Objections to the Classifications

	UK	West Indian	Indian
CRE	2.61	2.50	3.21
OPCS	2.46	2.63	2.96
MCB	2.53	1.82	2.92
NHDS	2.73	2.86	3.43

(b) Suitability of the Classifications

	UK	West Indian	Indian
CRE	2.47	2.50	2.64
OPCS	2.57	2.54	2.86
MCB	2.57	2.42	2.83
NHDS	2.66	2.63	2.93

APPENDIX 14

Analysis of Variance Tables For Suitability of Classifications

(a) For Total Sample

Source	df	SS	MS	F	
Between groups	3	2.03	.67	.69	
Within groups	168	853.29	5.08		NS
Error	504	491.98	.97		
Total	675	1347.57			

(b) For University Sample

Source	df	SS	MS	F	
Between groups	3	5.39	1.80	1.59	
Within groups	109	562.82	5.16		NS
Error	327	369.61	1.13		
Total	434	937.82			

(c) For School-Leaver Sample

Source	df	SS	MS	F	
Between groups	3	1.53	.51	.75	
Within groups	58	285.98	4.93		NS
Error	174	117.48	.68		
Total	235	404.99			

APPENDIX 15

Questionnaire Given to Respondents in Executive Officer Study

THE UNIVERSITY OF ASTON IN BIRMINGHAM

Gosta Green, Birmingham B47ET/Tel:021 359 3611 Ex 6178

IHD
Interdisciplinary Higher Degrees Scheme
Office

PLEASE DO NOT TURN OVER THIS FRONT PAGE UNTIL YOU HAVE BEEN INSTRUCTED TO DO SO.

The questionnaire before you is part of a research project being carried out by the Runnymede Trust - an educational charity conducting work in race relations - and the University of Aston in Birmingham, concerning equal opportunities in selection.

The aim of an equal opportunity policy is to ensure that no job applicant receives less favourable treatment on the grounds of race, colour, nationality, ethnic or national origins. In order to check whether or not an equal opportunity policy is working in selection, information is needed about applicants' ethnic background. Our research is aimed at finding a suitable method of obtaining that information.

From pilot work carried out already we have developed two different classifications, which we are asking you to fill in.

Your questionnaires will only be handled by the researchers. They will be treated with the utmost confidentiality and at no time will anyone from the Civil Service be given access to them. Thus, your responses to the questionnaire will not be known to the Commission and can have no possible bearing on your progress in this competition.

The link with the tests that you have just finished is that one way of studying how the classifications work will involve the researchers in looking at the average scores obtained by various groups of candidates. But - at no time will the researchers be given access to your test papers. The Civil Service Commission will merely provide them with test scores, after you have been notified of your results, on the basis of your index number. Complete confidentiality is therefore assured.

The questionnaire will only take a few minutes to complete, and your participation in this exercise is purely voluntary. However, bearing in mind that the information will be treated in confidence; will not be available to the Civil Service Commission; will have no effect on the progress of your application; and that the results of this research should benefit applicants for jobs in the future, we hope that you will take part.

Thank you for your co-operation.

1. Please tick the appropriate box to show which of the following groups you belong to (tick one box only).

English, Welsh, Scottish, Irish	01	☐
Italian	02	☐
Greek/Greek Cypriot	03	☐
Turkish/Turkish Cypriot	04	☐
Polish	05	☐
West Indian	06	☐
Rastafarian	07	☐
Nigerian	08	☐
Ghanain	09	☐
Indian	10	☐
Pakistani	11	☐
Bangladeshi	12	☐
Iranian	13	☐
Chinese	14	☐

This category is intended only for persons who are descended from people who originally came from these countries. If you are descended from people whose origins are other than these please tick the appropriate box below.

Other (including mixed descent) Please state _____

2. Had this classification actually been used as part of your selection procedure, would you: (please tick the appropriate box)

☐	☐	☐	☐	☐
strongly object	object	neutral	approve	strongly approve

1. Please tick the appropriate box below to show which of the groups you belong to (tick one box only).

British	01	☐	ie persons who were born in England, Wales, Scotland or Ireland AND who are descended from people who originally came from these countries.
British Asian	02	☐	ie persons born in Britain but who are descended from people who had Indian, Pakistani, Bangladeshi, or Sri Lankan origins. This category is not intended to include Asians born in African countries.
British West Indian	03	☐	ie persons born in Britain but who are descended from people who were of West Indian origin.
British Chinese	04	☐	ie persons born in Britain but who are descended from people who were of Chinese origin.
Other British - please state			_____ ie persons born in Britain but who are descended from people whose origins are other than those described above.
African Asian	05	☐	ie persons born in Africa but who are descended from people whose origins are Indian, Pakistani, Bangladeshi, or Sri Lankan.
Asian	06	☐	ie persons born in India, Pakistan, Bangladesh or Sri Lanka.
West Indian	07	☐	ie persons born in the West Indies.
African	08	☐	ie persons born in Africa.
European	09	☐	ie persons born in European countries other than England, Wales, Scotland and Ireland eg EEC countries, Eastern Europe, Cyprus, Spain, Malta, etc.
Chinese	10	☐	ie persons born in South-East Asia of Chinese origins.
Other - please state			_____ ie persons born in countries not included in the above categories including people of mixed descent.

2. Had this classification actually been used as part of your selection procedure, would you : (please tick the appropriate box)

☐	☐	☐	☐	☐
strongly object	object	neutral	approve	strongly approve

1. My Index number is _____

2. (a) Please state which country you were born in _____

   (b) If you were not born in the UK, please state how many years you have lived in the UK

   _____ years

3. Please state in which country you gained your:

   (a) Primary education _____

   (b) Secondary education _____

   (c) Further education (if any)_____

APPENDIX 16

Significant Differences in Pass Rates Between OPCS Groups

	Polish	E-W-S-I	Other	Ghanaian	Italian	West Indian	Indian	Pakistani	Chinese	Greek-Cypriot
Polish		NS	NS	NS	NS	NS	NS	NS	NS	NS
E-W-S-I			NS	NS	NS	**	**	NS	*	**
Other				NS	NS	NS	NS	NS	NS	NS
Ghanaian					NS	NS	NS	NS	NS	NS
Italian						NS	NS	NS	NS	NS
West Indian							NS	NS	NS	NS
Indian								NS	NS	NS
Pakistani									NS	NS
Chinese										NS
Greek-Cypriot										

NS Not Significant

** $p < .01$

* $p < .05$

APPENDIX 17

Significant Differences in Pass Rates Between MCB Groups

	Other	British	European	Other British	British Asian	British Chinese	British West Indian	African Asian	Asian	African	West Indian
Other		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
British			NS	NS	NS	NS	NS	**	**	**	**
European				NS	NS	NS	NS	NS	NS	NS	NS
Other British					NS	NS	NS	NS	NS	NS	NS
British Asian						NS	NS	NS	NS	NS	NS
British Chinese							NS	NS	NS	NS	NS
British West Indian								NS	NS	NS	NS
African Asian									NS	NS	NS
Asian										NS	NS
African											NS
West Indian											

NS Not Significant

** $p < .01$

* $p < .05$

APPENDIX 18

Analysis of Variance Table for Total Scores By OPCS Categories

Source	df	SS	MS	F	
Between groups	12	457478.58	37623.22	29.02	$p<.01$
Within groups	3385	4388326.99	1296.40		
Total	3397	4839805.57			

APPENDIX 19

Analysis of Variance Table For Total Scores by MCB Categories

Source	df	SS	MS	F	
Between groups	11	415998.26	37818.02	29.03	$p<.01$
Within groups	3388	4342630.22	1302.53		
Total	3399	4758628.48			

APPENDIX 20

Analysis of Variance Table For Intelligence Test Scores By OPCS Categories

Source	df	SS	MS	F	
Between groups	12	143412.3597	11951.03	14.81	p .01
Within groups	3385	2731983.9292	807.09		
Total	3397	2875396.2890			

APPENDIX 21

Analysis of Variance Table For Intelligence Test Scores by MCB Categories

Source	df	SS	MS	F	
Between groups	11	146973.79	13361.25	16.56	p .01
Within groups	3380	2727278.86	806.89		
Total	3391	2874252.66			

APPENDIX 22

Analysis of Variance Table For Enquiries From Management Test Scores By OPCS Categories

Source	df	SS	MS	F	
Between groups	12	14260.31	1188.36	11.79	$p < .01$
Within groups	3385	341138.94	100.78		
Total	3397	355399.25			

APPENDIX 23

Analysis of Variance Table For Enquiries From Management Test Scores By MCB Categories

Source	df	SS	MS	F	
Between groups	11	13522.55	1229.32	12.13	$p < .01$
Within groups	3380	342512.29	101.34		
Total	3391	356034.84			

APPENDIX 24

Analysis of Variance Table For Executive Problems Test Scores By OPCS Categories

Source	df	SS	MS	F	
Between groups	12	21929.24	1827.44	16.43	$p < .01$
Within groups	3385	376613.63	111.26		
Total	3397	398542.86			

APPENDIX 25

Analysis of Variance Table For Executive Problems Test Scores By MCB Categories

Source	df	SS	MS	F	
Between groups	11	22221.81	2020.16	18.10	$p < .01$
Within groups	3380	377241.24	111.61		
Total	3391	399463.05			

APPENDIX 26

Stepwise Multiple Regression Analysis With Highly Correlated Variables Excluded (OPCS Entered Last)

Variable	R Square	F	R Square Change
Age on Entering UK	.04	12.82	.04**
Educational qualifications	.06	84.21	.02*
OPCS	.11	161.26	.05**

APPENDIX 27

Stepwise Multiple Regression Analysis With Highly Correlated Variables Excluded (MCB Entered Last)

Variable	R Square	F	R Square Change
Age on Entering UK	.04	4.07	.04**
Educational qualifications	.06	79.42	.02*
MCB	.11	141.02	.05**

APPENDIX 28

Multiple Regression Analysis: Results of Double Cross Validation For OPCS Classification with OPCS Entered First

Group 1

Variable	R Square	F	R Square Change
OPCS	.09	63.91	.09**
Educational qualifications	.11	42.45	.02*

Pearson r (Calculated R^2) = .03

Group 2

Variable	R Square	F	R Square Change
OPCS	.10	100.41	.10**
Educational qualifications	.11	33.99	.01

Pearson r (Calculated R^2) = .02

** $p < .01$
* $p < .05$

APPENDIX 29

Multiple Regression Analysis: Results of Double Cross Validation With OPCS Entered Last

Group 1

Variable	R Square	F	R Square Change
Country of Primary Education	.06	8.82	.06**
Educational qualifications	.09	42.39	.03*
OPCS	.12	61.10	.03*

Pearson r (Calculated R^2) = .03

Group 2

Variable	R Square	F	R Square Change
Country of Primary Education	.04	2.30	.04*
Educational qualifications	.06	34.39	.02*
OPCS	.12	95.08	.06**

Pearson r (Calculated R^2) = .01

** $p < .01$
* $p < .05$

APPENDIX 30

Multiple Regression Analysis: Results of Double Cross Validation With MCB Entered First

Group 1

Variable	R Square	F	R Square Change
MCB	.09	47.86	.09**
Educational qualifications	.11	38.40	.02*

Pearson r (Calculated R^2) = .02

Group 2

Variable	R Square	F	R Square Change
MCB	.08	71.68	.08**
Educational qualifications	.10	31.37	.02*

Pearson r (Calculated R^2) = .01

** $p < .01$
* $p < .05$

APPENDIX 31

Multiple Regression Analysis: Results of Double Cross Validation With MCB Entered Last

Group 1

Variable	R Square	F	R Square Change
Country of Primary Education	.06	5.92	.06**
Educational qualifications	.09	38.32	.03*
MCB	.12	45.08	.03*

Pearson r (Calculated R^2) = .02

Group 2

Variable	R Square	F	R Square Change
Country of Primary Education	.04	2.39	.04**
Educational qualifications	.06	31.74	.02*
MCB	.10	66.36	.04

Pearson r (Calculated R^2) = .00

APPENDIX 32

Analysis of Covariance for Total Scores By OPCS Categories

Source	df	SS	MS	F	
Covariates	4	319.98	79.99	49.40	$p<.01$
Education	1	37.93	37.93	23.42	$p<.01$
Sex	1	11.82	11.82	7.30	$p<.01$
Country of Primary Education	1	2.97	2.97	1.84	NS
Age on Entering UK	1	261.34	261.34	161.40	$p<.01$
Main effects	12	383.11	31.93	19.72	$p<.01$
OPCS	12	383.11	31.93	19.72	$p<.01$
Explained	16	703.09	43.94	27.14	$p<.01$
Residual	3318	5372.54	1.62		
Total	3334	6075.63	1.82		

APPENDIX 33

Analysis of Covariance For Intelligence Test Scores By OPCS Categories

Source	df	SS	MS	F	
Covariates	4	258.20	64.55	50.99	$p<.01$
Education	1	23.82	23.82	18.82	$p<.01$
Sex	1	9.84	9.84	7.77	$p<.01$
Country of Primary Education	1	1.80	1.80	1.42	NS
Age on Entering UK	1	217.27	217.27	171.63	$p<.01$
Main effects	12	280.31	23.36	18.45	$p<.01$
OPCS	12	280.31	23.36	18.45	$p<.01$
Explained	16	538.51	33.66	26.59	$p<.01$
Residual	3318	4200.36	1.27		
Total	3334	4738.87	1.42		

APPENDIX 34

Analysis of Covariance For Enquiries From Management Test Scores By OPCS Categories

Source	df	SS	MS	F	
Covariates	4	34.34	8.58	16.75	$p<.01$
Education	1	5.79	5.79	11.30	$p<.01$
Sex	1	3.34	3.34	6.51	$p<.01$
Country of Primary Education	1	0.10	0.10	0.20	NS
Age on Entering UK	1	23.57	23.57	45.98	$p<.01$
Main effects	12	4.08	4.08	7.96	$p<.01$
OPCS	12	4.08	4.08	7.96	$p<.01$
Explained	16	5.21	5.21	10.16	$p<.01$
Residual	3318	0.51	0.51		
Total	3334	0.54	0.54		

APPENDIX 35

Analysis of Covariance for Executive Problems Test Scores By OPCS Categories

Source	df	SS	MS	F	
Covariates	4	80.49	20.12	31.46	$p<.01$
Education	1	8.87	8.87	13.87	$p<.01$
Sex	1	1.15	1.15	1.79	NS
Country of Primary Education	1	1.73	1.73	2.71	NS
Age on Entering UK	1	69.10	69.10	108.02	$p<.01$
Main effects	12	80.40	6.70	10.47	$p<.01$
OPCS	12	80.40	6.70	10.47	$p<.01$
Explained	16	160.89	10.06	15.72	$p<.01$
Residual	3318	2122.55	0.64		
Total	3334	2283.44	0.69		

APPENDIX 36

Analysis of Covariance For Total Scores By MCB Categories

Source	df	SS	MS	F	
Covariates	4	320.82	80.20	44.60	$p<.01$
Education	1	38.49	38.49	22.36	$p<.01$
Sex	1	15.91	15.91	9.24	$p<.01$
Country of Primary Education	1	2.96	2.96	1.72	NS
Age on Entering UK	1	254.75	254.75	148.00	$p<.01$
Main effects	11	371.19	33.74	19.61	$p<.01$
MCB	11	371.19	33.74	19.61	$p<.01$
Explained	15	692.00	46.13	26.80	$p<.01$
Residual	3315	5705.95	1.72		
Total	3330	6397.95	1.92		

APPENDIX 37

Analysis of Covariance For Intelligence Test Scores By MCB Categories

Source	df	SS	MS	F	
Covariates	4	257.73	64.43	45.57	$p<.01$
Education	1	24.73	24.73	17.49	$p<.01$
Sex	1	14.11	14.11	9.98	$p<.01$
Country of Primary Education	1	1.56	1.56	1.10	NS
Age on Entering UK	1	209.13	209.13	147.92	$p<.01$
Main effects	11	263.61	23.96	16.95	$p<.01$
MCB	11	263.61	23.96	16.95	$p<.01$
Explained	15	521.34	34.76	24.58	$p<.01$
Residual	3315	4686.83	1.41		
Total	3330	5208.17	1.56		

APPENDIX 38

Analysis of Covariance For Enquiries From Management Test Scores By MCB Categories

Source	df	SS	MS	F	
Covariates	4	35.36	8.84	17.07	$p < .01$
Education	1	5.73	5.73	11.07	$p < .01$
Sex	1	3.68	3.68	7.10	$p < .01$
Country of Primary Education	1	0.05	0.05	0.10	NS
Age on Entering UK	1	23.93	23.93	46.20	$p < .01$
Main effects	11	47.69	4.34	8.37	$p < .01$
MCB	11	47.69	4.34	8.37	$p < .01$
Explained	15	83.05	5.54	10.69	$p < .01$
Residual	3315	1716.88	0.52		
Total	3330	1799.93	0.51		

APPENDIX 39

Analysis of Covariance For Executive Problems Test Scores By MCB Categories

Source	df	SS	MS	F	
Covariates	4	80.11	20.03	31.04	$p<.01$
Education	1	8.81	8.81	13.65	$p<.01$
Sex	1	1.50	1.50	2.33	NS
Country of Primary Education	1	2.28	2.28	3.53	NS
Age on Entering UK	1	67.83	67.83	105.15	$p<.01$
Main effects	11	81.45	7.41	11.48	$p<.01$
MCB	11	81.45	7.41	11.48	$p<.01$
Explained	15	161.56	10.77	16.70	$p<.01$
Residual	3315	2138.56	0.65		
Total	3330	2300.12	0.69		

REFERENCES

ADVISORY COMMITTEE ON RACE RELATIONS RESEARCH (1975)
Race Relations Research: A report to the Home Secretary
London, HMSO

ALLEYNE, H.M.McD. (1962)
The effect of bilingualism on performance in certain intelligence and attainment tests
Unpublished MA thesis, University of London

AMERICAN PSYCHOLOGICAL ASSOCIATION
Standards for educational and psychological tests
Washington DC: APA 1974

AMERICAN PSYCHOLOGICAL ASSOCIATION (1975)
Principles for the Validation and Use of Personnel Procedures
Division of Industrial-Organisational Psychology
Washington DC: APA

ANASTASI, A. (1976)
Psychological Testing (4th Ed)
New York: MacMillan

ARVEY, R.D. (1972)
Some comments on Culture Fair Tests
Personnel Psychology, 25, 433-448

BALLARD, R. & HOLDERN, B. (1975)
The Employment of Coloured Graduates in Britain
New Community, Vol.IV, No. 3, Autumn

BARTLETT, C.J., BOBKO, P., & PINE, S.M. (1977)
Single Group Validity: Fallacy of the facts?
Journal of Applied Psychology, 62, 2, 155-157

BARTLETT, C.J. & O'LEARY, B.S. (1969)
A Differential Prediction Model to Moderate the Effects of Heterogeneous Groups in Personnel Selection and Classification
Personnel Psychology, 22, 1-17

BENEDICT, R. (1942)
Race and Racism
London: Routledge and Kegan Paul

BIESHEUVEL, S. (1949)
Psychological tests and their application to non-European peoples
Yearbook Education, London: Evans Bros.

BINDMAN, G. & GROSZ, S.
Indirect Discrimination and the Race Relations Act
In 'A Review of the Race Relations Act 1976'
London: The Runnymede Trust (1979)

BINDMAN, G. (April 1979)
Indirect Discrimination and the Race Relations Act
New Law Journal

BOBKO, P. & BARTLETT, C.J. (1978)
Sub-group Validities: Differential Definitions and Differential Prediction
Journal of Applied Psychology, 63, 1, 1-11

BOEHM, V.R. (1972)
Negro-White Differences in Validity of Employment and Training Selection Procedures
Journal of Applied Psychology, 56, 1, 33-39

BOEHM, V.R. (1977)
Differential Prediction: A Methodological Artefact?
Journal of Applied Psychology, 62, 2, 146-154

BOEHM, V.R. (1978)
Populations, Preselection and Practicalities: A Reply to Schmidt and Hunter
Journal of Applied Psychology, 63, 1, 15-18

BRELAND, H.M. & IRONSON, C.M. (1976)
Defunis reconsidered: A comparative analysis of alternative admission strategies
Journal of Educational Measurement, 13, 1, 89-99

BRITISH PSYCHOLOGICAL SOCIETY & THE RUNNYMEDE TRUST (1980)
Discriminating Fairly: A Guide to Fair Selection
Leicester: The BPS and
London: The Runnymede Trust

BROTHERTON, C.J., COX, T.R., HOWARTH, C.I.
Studies of Selection in British Industry
Manchester: NCC 1981

BROTHERTON, C. (1980)
Paradigms of Selection Validation: Some comments in the light of British Equal Opportunities Legislation
Journal of Occupational Psychology, 53, 73-79

BULMER, M. (1980)
On the feasibility of identifying 'race' and 'ethnicity' on censuses and surveys
New Community, VIII, 1 and 2 (Spring/Summer 1980)

CAMPBELL, J.T., CROOKS, L.A., MAHONEY, M.M., & ROCK, D.A.
An investigation of sources of bias in prediction of job performance - a 6-year study
Tech Rep PR - 73-37 Princeton, N.J.: Educational Testing Service

CASCIO, W.F., & PHILLIPS, N.F. (1979)
Performance Testing: A rose among thorns
Personnel Psychology, 32, 751-766

CIVIL SERVICE DEPARTMENT
Applications of Race Relations Policy in the Civil Service
HMSO LONDON (1978)

CLEARY, T.A. (1968)
Test Bias: Prediction of Grades of Negro and White Students in Integrated Colleges
Journal of Educational Measurement, 5, 2, 115-124

COCHRAN, A.J. (1981)
Vocational PhDs: Aston's I.H.D. Scheme
University of Aston

COLE, N.S. (1973)
Bias in Selection
Journal of Educational Measurement, 10, 237-255

COLEMAN, J.S. ET AL (1966)
Equality of Educational Opportunity
Washington DC
US Office of Education

COMMISSION FOR RACIAL EQUALITY (1978)
Monitoring on Equal Opportunity Policy - A Guide for Employers
London: CRE

COMMISSION FOR RACIAL EQUALITY (1978)
Equal Opportunity in Employment: A Guide for Employers
London: CRE

COMMISSION FOR RACIAL EQUALITY (1979)
Draft Code of Practice
London: CRE

COMMISSION FOR RACIAL EQUALITY (1984)
Code of Practice
London: CRE

CRONBACH, L.J., YALOW, E. & SCHAEFFNER, G. (1980)
A Mathematical Structure for Analysing Fairness in Selection
Personnel Psychology, 33, 693-704

DANIEL, W.W. (1968)
Racial Discrimination in England
London: Penguin

DARLINGTON, R.B. (1971)
Another look at 'cultural fairness'
Journal of Educational Measurement, 8, 71-82

DISTEFANO, M.K., PRYER, M.W. & CRAIG, S.H. (1978)
Job Relatedness of a Post-Training Job Knowledge Criterion
Used to Assess Validity and Test Fairness
Personnel Psychology, 33, 785-793

EQUAL OPPORTUNITIES COMMISSION (1985)
Code of Practice
London: HMSO

EELLS, K., DAVIS, A.L, HAVIGHURST, R.J., HERRICK, V.E.
& TYLER, R.W. (1951)
Intelligence and Cultural Differences
Chicago University Press: Chicago

EINHORN, M.J. & BASS, A.R. (1971)
Methodological Considerations Relevant to Discrimination
in Employment Testing
Pyschological Bulletin, 75, 4, 261-269

EQUAL EMPLOYMENT OPPORTUNITY COMMISSION (EEOC) (1970)
Guidelines on Employee Selection Procedures
Federal Register, (1970), 35, (1949), 12333-12336

FERGUSON, G.A. (1981)
Statistical Analysis in Psychology and Education
McGraw-Hill (5th Ed.)

FIELD, H.S., BAYLEY, E.A. & BAYLEY, S.M. (1977)
Employment Test Validation for Minority and Non-Minority Production Workers
Personnel Psychology, 30, 37-46

FIELD, S., MUIR, G., REES, T., & STEVENS, P. (1981)
Ethnic Minorities in Britain
London: HMSO

FINCHER, C. (1975)
Differential Validity and Test Bias
Personnel Psychology, 28, 481-500

FLAUGHER, R.C. (1974)
Institutionalised Discrimination and the Issue of Test Bias
Princeton, New Jersey: Educational Testing Service

FLAUGHER, R.C. (1974)
Bias in Testing: A Review and Discussion
Princeton, New Jersey: Educational Testing Service

FLAUGHER, F.C. (1978)
The many definitions of Test Bias
American Psychologist 33, 671-679

GAEL, S. (1974)
On O'Leary's Fair Employment
American Psychologist, 29, 216-217

GAEL, S. & GRANT, D.L. (1972)
Employment Test Validation for Minority and Non-Minority Telephone Company Service Representatives
Journal of Applied Psychology, 56, 2, 135-139

GHISELLI, E.E. (1966)
The Validity of Occupational Aptitude Tests
New York: Wiley

GOODMAN, J. (1980)
Indirect Discrimination and the Employer's Right to Select
Personnel Management, February

GOODMAN, J. & NOVARRA, V. (1977)
The Sex Discrimination Act - A Role for Psychologists
Bulletin of the BPS, 30, 104-105

GRANT, D.L. & BRAY, D.W. (1970)
Validation of Employment Tests for Telephone Installation and Repair Occupations
Journal of Applied Psychology, 54, 1, 7-14

GROSS, A.L. & SU, W (175)
Defining a "fair" or "unbiased" selection model: A question of utilities
Journal of Applied Psychology, 60, 345-351

GUION, R.M. (1965)
Personnel Testing
New York: McGraw-Hill

GUION, R.M. (1976)
Recruiting, Selection, Job Placement
In M.D. Dunnette (ed) Handbook of Industrial and Organisational Psychology
Chicago: Rand McNally

GUION, R.M. (1978)
Scoring of Content domain samples: The problem of fairness
Journal of Applied Psychology, 63, 4, 449-506

HEGARTY, S. & LUCAS, D. (1978)
Able to learn
Windsor: NFER

HEPPLE, B. (1968)
Race Jobs and the Law in Britain
London: Penguin

HITNER, T., KNIGHTS, D., GREEN, E., TORRINGTON, D (1981)
Races at Work: Equal Opportunity Policy and Practice
Employment Gazette, September, 404-409

HOME OFFICE (1975)
Racial Discrimination
Cmnd. 6234
HMSO, London

HOME OFFICE (1976)
Race Relations Act
HMSO, London

HOME OFFICE (1968)
Race Relations Act
HMSO, London

HUBBOCK, J. & CARTER, S. (1980)
Half a chance
London: CRE

HUNTER, J.E. & SCHMIDT, F.L. (1976)
A Critical Analysis of the Statistical and Ethical Implications
of Various Definitions of Test Bias
Psychological Bulletin, 83, 1053-1071

HUNTER, J.E. & SCHMIDT, F.L. (1978)
Differential and Single Group Validity of Employment Tests
by Race: A Critical Analysis of Three Recent Studies
Journal of Applied Psychology, 63, 1, 1-11

INCOMES DATA SERVICE
Race Discrimination
IDS Brief Supplement 23,
London: Incomes Data Service 1979a

INSTITUTE OF PERSONNEL MANAGEMENT
Towards fairness in selection: A code for non-discrimination
London: I.P.M. (1978)

IRVINE, S.H. (1966)
Towards a rationale for testing attainments and abilities
in Africa
British Journal of Educational Psychology, 36, 24-32

JAE-ON, K., KOHOUT, F.J.
Analysis of Variance and Covariance
In Nie et al (1975) Statistical Package for the Social Sciences
McGraw-Hill Inc: 2nd ed.

JENNER, G. (1984)
The CRE Code of Practice
Personnel Management, February

JENSEN, A. (1971)
Do Schools Cheat Minority Children?
Educational Research, 14, 1, 3-28

JENSEN, A.R. (1980)
Bias in Mental Testing
New York: Methuen

KATZELL, R.A., DYER, F.J. (1977)
Differential Validity Revived
Journal of Applied Psychology, 62, 2, 137-145

KATZELL, R.A. & DYER, F.J. (1978)
On Differential Validity and Bias
Journal of Applied Psychology, 63, 1, 19-21

KERLINGER, F.N. & PEDHAZUR, E.J. (1973)
Multiple Regression in Behavioural Research
New York: Holt, Rinehart and Winston Inc

KESSELMAN, G.A. & LOPEZ, F.E. (1979)
The impact of job analysis on employment test validation for minority and non-minority accounting personnel
Personnel Psychology, 32, 91-108

KIRK, R.C. (1968)
Experimental Design: Procedures for the Behavioural Sciences
Belmont California: Brooks/Cole

KIRKPATRICK, J.J., EWEN, R.B., BARRETT, R.S., KATZELL, R.A. (1969)
Testing and Fair Employment
London: University of London Press

LESTER, A., BINDMAN, G. (1972)
Race and Laws
London: Penguin

LINN, R. (1973)
Fair Test Use in Selection
Review of Educational Research
43, 2, 139-161

LINN, R. (1976)
In Search of Fair Selection Procedures
Journal of Educational Measurement
13, 1, 53-58

LITTLE, A., MABEY, C. & WHITAKER, G. (1968)
The Education of Immigrant Pupils in Inner London Primary Schools
Race IX, 4, April

LOMAS, G. (1980)
Race, Colour, Ethnicity - The Collection of Data
New Community, Vol. VIII Nos 1 and 2
Spring-Summer

LOMAS, G. (1980)
Ethnic and Racial Data: What Next
New Community Vol VIII No 3, Winter, 329-335

LUNDSGAARDE, H.P. (1974)
Racial and Ethnic Classifications: An Appraisal of the Role of Anthropology in the Lawmaking Process
New Community Vol 111, No 4, Autumn, 349-358

MACDONALD, I. (1977)
Race Relations: The New Law
London: Butterworths

MCFIE, J. & THOMPSON, J.A. (1970)
Intellectual Abilities of Immigrant Children
British Journal of Educational Psychology, 40, 348-51

MCINTOSH, N. & SMITH, D.J. (1974)
The Extent of Racial Discrimination
Political and Economic Planning, Vol. XL. Broadsheet 547
London: Policy Studies Institute

MCNEMAR, Q. (1975)
On so-called test bias
American Psychologist, 30, 848-851

MARASCUILO, C.A., & MCSWEENEY, M. (1977)
Non-parametric Statistics and Distribution-free Methods for the Social Sciences
New York

MAYS, R. (1980)
Bias in the selection of Civil Service Administrators
Paper presented at British Psychological Society, London Conference

MILLER, K. (Ed) (1975)
Psychological Testing in Personnel Assessment
London: Gower Press

MILLER, K.M. & HYDES, K.M. (1971)
The Use of Psychological Tests in Personnel Work
London: Independent Assessment and Research Centre

MINER, M.G. & MINER, J.B. (1978)
Employee Selection Within the Law
Washington DC: Bureau of National Affairs

MONTAGU, A. (1974)
Man's Most Dangerous Myth
New York: Oxford University Press

NIE, N.H., HULL, C.H., JENKINS, J.G., STEINBRENNER, K., BENT, D.H. (1975)
Statistical Package for the Social Sciences
McGraw-Hill Inc 2nd ed

NOVICK, M.R. & PETERSEN, N.S. (1976)
Toward Equalising Educational and Employment Opportunities
Journal of Educational Measurement, 13, 1, 77-88

NOVICK, M.R. & ELLIS, D.D. (1977)
Equal Opportunities in Educational and Employment Selection
American Psychologist, 306-320

NYFIELD, G., HOLDSWORTH, R. & SAVILLE, P. (1980)
The Advanced Test Battery Manual
Esher: SHL

OFFICE OF FEDERAL CONTRACT COMPLIANCE (OFCC) (1971)
Employee Testing and Other Selection Procedures
Federal Register, 36 (192) 19307-19310

OFFICE OF FEDERAL CONTRACT COMPLIANCE (OFCC) (1974)
Guidelines for Reporting Criterion-related and Content Validity
Federal Register, 39, (12), 2094-2096

OFFICE OF POPULATION CENSUSES AND SURVEYS (1980)
Plans for the 1981 Census
OPCS Monitor, CEN, 80/1, March

OFFICE OF POPULATION CENSUSES AND SURVEYS (1980)
Tests of an Ethnic Question
OPCS Monitor, CEN, 80/2, March

OFFICE OF POPULATION CENSUSES AND SURVEYS (1980)
Government's decision on an ethnic question in the 1981 census
OPCS Monitor, CEN, 80/3, April

OFFICE OF POPULATION CENSUSES AND SURVEYS (1981)
1981 Census - Developing a Question on Ethnic Origins
OPCS Monitor, CEN, 78/4, July

ORTAR, G.R. (1963)
The transfer of psychological diagnostic measures from one culture to another
Acta Psychologica, 21, 218-230

PAUL, S.E., KUPLENS, I. (1977)
Executive Officer Selection Follow-up Study: some tentative findings
Recruitment Research Unit, Civil Service Commission

PEARN, M.A. (1977)
Selecting and Training of Coloured Workers
Training Information Paper No.9
London: HMSO

PEARN, M.A. (1978)
Beyond Tokenism - Equal Employment Opportunities Policies
Runnymede Trust Briefing Paper, June

PEARN, M.A. (1978)
The Fair Use of Tests
Windsor: NFER Publishing Co

PEARN, M.A. (1979)
Exploring the meaning of indirect discrimination in selection
In "A review of the Race Relations Act"
London: Runnymede Trust

PEARN, M.A. (1979)
Selection Fairness: Past and Future Trends in Discrimination
London: Runnymede Trust

PEARN, M.A. (1978)
Employment Testing and the Goal of Equal Opportunities: The American Experience
London: Runnymede Trust

PEARN, M.A. (1979)
Monitoring Equal Opportunity in the Civil Service: A Review of a Report by The Tavistock Institute
Briefing Paper: London: Runnymede Trust

PEDHAZUR, E.J. (1982)
Multiple Regression in Behavioural Research: Explanation and Prediction
New York: CBC College Publishing

PETERSEN, N.S. & NOVICK, M.R. (1976)
An Evaluation of Some Models for Culture Fair Selection
Journal of Educational Measurement, 13, 1, 3-29

POLLACK, L.H. (1974)
Discrimination in Employment: The American Response
London: The Runnymede Trust

PRASHAR, U. & PEARN, M.A. (1979)
Overview. In "A Review of the Race Relations Act"
London: Runnymede Trust

PRICE, C. (1980)
What's in a name? Australia ponders
New Community Vol III Nos 1 and 2, 17-18 Spring-Summer

PSYCHOLOGICAL CORPORATION (1979)
Summaries of the Court Decisions on Employment Testing 1968-1977
New York: The Psychological Corporation

RACE RELATIONS BOARD (1975)
Principles and Methods of Monitoring Equality of Opportunity in Employment
London: Race Relations Board, January

ROBERTSON, I.T. & KANDOLA, R.S. (1982)
Work Sample Tests: Validity, Adverse Impact and Applicant Reaction
Journal of Occupational Psychology, 55,171-183

RUNNYMEDE TRUST (1968)
Runnymede Trust Deeds
London: Runnymede Trust

RUNNYMEDE TRUST (1975)
Racial Discrimination: A Guide to the White Paper
London: Runnymede Trust 1975

RUNNYMEDE TRUST (1979)
Language Testing and Indirect Discrimination: Lessons From the British Steel Case
London: Runnymede Trust

RUNNYMEDE TRUST (1979)
Review of the Race Relations Act 1976
London: Runnymede Trust

RUNNYMEDE TRUST AND RADICAL STATISTICS GROUP (1980)
Britain's Black Population
London: Heinemann Educational

SAWYER, R.L., COLE, N.S., & COLE, J.W.L. (1976)
Utilities and the Issue of Fairness in a Decision Theoretic Model For Selection
Journal of Educational Measurement, 13, 1

SCHMIDT, F.L., BERNER, J.G., HUNTER, J.E. (1973)
Racial Differences in Validity of Employment Tests: Reality or Illusion
Journal of Applied Psychology, 58, 1, 5-9

SCHMIDT, F.L., GREENTHAL, A.C., HUNTER, J.E., BERNER, J.G. & SEATON, F.W. (1977)
Job samples vs paper and pencil trades and technical tests: Adverse impace and examinee attitudes personnel psychology, 30, 187-197

SCHMIDT, F.L., PEARLMAN, K., HUNTER, J.E. (1980)
The Validity and Fairness of Employment and Educational Tests for Hispanic Americans: A Review and Analysis
Personnel Psychology, 33, 705-724

SEEAR, B.N. (1979)
The Management of Equal Opportunity
Briefing Paper, London: The Runnymede Trust

SIEGEL, S. (1956)
Non-parametric Statistics
New York: John Wiley and Sons Inc.

SMITH, D.J. (1974)
Racial Disadvantage in Employment
PEP Vol XL, Broadsheet, 544,

SMITH, D.J. (1976)
The Facts of Racial Disadvantage: A National Survey
London: PEP Vol XLII

SMITH, D.J. (1974)
The Extent of Racial Discrimination
London: Political and Economic Planning

SMITH, D.J. (1980)
Unemployment and Racial Minority Groups
Dept of Employment Gazette, June, 602-607

SNEATH, F., THAKUR, M., MEDJUK, B. (1976)
Testing People at Work
London: IPM

STEVENS, J. (1973)
The Case for Keeping Ethnic Records
Camden Committee for Community Relations

TAYLOR, M.J. (1983)
Caught Between: A Review of Research into the Education of Pupils of West Indian Origin
Windsor: NFER-NELSON

THORNDIKE, R.L. (1971)
Concepts of Culture fairness
Journal of Educational Measurement, 8, 63-70

TRATTNER, M.H., O'LEARY, B.S. (1980)
Sample Sizes for Specifed Statistical Power in Testing for Differential Validity
Journal of Applied Psychology, 65, 2, 127-134

TUKEY, J.W. (1949)
Comparing Individual Means in the Analysis of Variance
Biometrics, 5:99-114

U.S. GENERAL ACCOUNTING OFFICE (1979)
Federal Employment Examinations: Do they achieve Equal Opportunity and Merit Principle Goals?
Comptroller General's Report to the Congress of the United States
Washington D.C. : U.S. General Accounting Office

VERNON, P.E. (1965)
Environmental handicaps and intellectual development Part 1
British Journal of Educational Psychology, XXXV, 1, 9-20

VERNON, P.E. (1965 b)
Environmental handicaps and Intellectual Development Part II
British Journal Educational Psychology, XXXV, 2, 117-26

VERNON, P.E. (1969)
Intelligence and Cultural Environment
London: Methuen

WAINWRIGHT, D. (1979)
Discrimination in Employment
London: Associated Business Press

WAINWRIGHT, D. (1980)
Learning from Uncle Sam: Equal Employment Opportunities Programme
Runnymede Trust Industrial Unit Briefing Paper No. 3, May

WALLIS, D. (1980)
Sex Discrimination and the Law
Bulletin of the British Psychological Society, 33, 1-5

WALLMAN, S. (1978)
Race Relations or Ethnic Relations
New Community Vol VI, No. 3, Summer, 306-309

WATSON, P. (1973)
Stability of IQ of immigrant and non-immigrant slow learning pupils
British Journal of Education Psychology, 43, 80-2

WHEATON, G.R. (1968)
Development of a Taxonomy of Human Performance: A Review of Classificatory Systems Relating to Tasks and Performance
American Institute for Research, Washington Office, Skills Research Programme, December

WHITE, R.M. (1979)
What's in a name? Problems in official and legal uses of race
New Community, VIII, No.3, Winter, 333-349

WILDT, A.R., AHTOLA, O.T. (1978)
Analysis of Covariance
Sage Publications USA

WINER, B.J. (1971)
Statistical Principles in Experimental Design
2nd Ed. New York McGraw-Hill

WRIGHT, J., HEWITT, P., SEDLEY, A. (1978)
Race Relations Guide
London: NCCL

YULE, W., BERGER, M., RUTTER, M & YULE, B. (1975)
Children of West Indian Immigrants - II, intellectual performance and reading attainment
Journal of Child Psychology and Psychiatry, 16, 1-17

ZEDECK, S. (1971)
Problems with the Use of 'Moderator' Variables
Psychological Bulletin, 76, 4, 295-310

ZEDECK, S. & TENOPY, R.(1975)
Issues in Selection Testing and the Law
Industrial Relations Research Association Series, Equal Rights and Industrial Relations, 167-196