THE ASSESSMENT OF THE LEARNING EFFICIENCY OF ASIAN CHILDREN

RAJINDER MOHAN GUPTA

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF ASTON IN BIRMINGHAM

OCTOBER 1983

The University of Aston in Birmingham

The Assessment of the Learning Efficiency of Asian Children

Rajinder Mohan Gupta

PhD 1983

Two aims were central to the present research: i) to develop a test battery for assessing learning efficiency for Asian children brought up in England; ii) to test the generalisability of Jensen's theory of mental abilities on English and Asian children. The main sample involved 375 children, both English and Asian, in the age range 6-10 years; in addition small groups of ESN-M and Indian children living in India were studied for comparison. The results suggest that the LET Battery is culturally fair for English and Asian children; whereas the Raven's Matrices Test is clearly culturally biased. The LETB is homogeneous in content, has satisfactory internal consistency, test - retest reliability and acceptable face, predictive and construct validity. The LETB is better in predicting reading and reading gains of children in the low ability range compared to the conventional assessment procedures. For children of wider ability range, there is very little difference in the predictive ability of conventional assessment procedures and the LETB: notwithstanding this, the use of conventional assessment procedures on Asian children cannot be recommended because they tend to misclassify a very high percentage of them as being in the "mentally defective" range. Further the outcome from the LETB can be linked with educational objectives. In testing Jensen's hypotheses, out of a total of four, two hypotheses, one concerned with the orthogonality of Level I and Level II ability, and the other, with the mean differences on Level II ability between English and Asian children, can be firmly rejected. The remaining two, concerned with the non-significant differences of the factorial structure and of Level I ability between English and Asian children appear to be in line with Jensen's thinking, although it is argued here that Jensen's theory is too simple to be capable of explaining differences in mental processes across various populations.

Keywords: Assessment, Learning Efficiency, Asian Children, Test.

ACKNOWLEDGEMENTS

It is perhaps a truism that a research project is rarely ever completed without the help, guidance and substantial encouragement from a number of people. This Study has been no exception: during 1977/78 when it was still at its conceptual stage, I benefitted greatly from the experience and constructive criticisms of the following psychologists: Dr. Chapman from Bristol University; Mr. Curr and Dr. Hewitt both from Birmingham Univeristy; Dr. Herbert, formerly Chief Educational Psychologist, Birmingham Psychological Services and Dr. Kline from Exeter University.

I am also particularly thankful to Dr. D.G. Lewis of Manchester University for giving an initial critical evaluation of the whole dissertation; to Dr Alan Slater for commenting on Chapters 1 and 3 and to Professor Howe for commenting on Chapters 2 and 8. I value their help immensely.

I must acknowledge unreservedly that Dr. P. Coxhead has been an exemplary Supervisor. It has been a great pleasure and privilege to carry out this research under his supervision. But for his unfailing guidance and advice, I would never have managed to complete this research. I am sincerely grateful to Dr. Coxhead for the experience I have gained from working under his supervision.

My thanks are also due to several headteachers, teachers and children in the Ward End area of Birmingham for their help and co-operation towards the completion of this project.

Sincere appreciation is also extended to Elaine Turbill (for the earlier draft) and to Annette Saxby and Sarah Smojkis for intelligently and efficiently typing the whole manuscript.

Last but not least, my sincere thanks goes to my Wife, Jessie and my three Daughters, Deepa, Surya and Geeta for their tolerance.

encouragement and understanding while I have been busy with this research for several years.

Rajinder M. Gupta

CONTENTS

LIST OF TABLES AND FIGURES

Tables in Appendix B

Figures

DEFINITIONS

Definitions

English: For the purpose of this study, English children are defined as those whose parents were born in the British Isles. These children are referred to as English, although some may well be Welsh, or Scottish either by parentage or birth.

Indian: This term refers to those children who were born to Indian parents and who had always lived in India. Neither these children nor their parents had ever left India to live abroad.

ESN-M: Children who were receiving their education in special schools designated by the Local Education Authority as Schools for the Educational Subnormal - Mild. (For the criteria employed to recommend children to receive education in special institutions see Chazan et al., 1974; Clarke & Clarke, 1974; Home Office, 1978; Mittler 1970.) With the publication of the Warnock Report (Home Office, 1978) and the new Education Act 1981 (DES 1981, 1983) the terms ESN-M and ESN-S have been replaced by children with moderate and severe learning difficulties. Since the new terminology came into vogue while the present study was more than half way through, it was therefore decided to adhere to the old and well familiar labels. However, wherever it has been possilbe references to the 1981 Act and the Warnock Report have been made.

Learning Efficiency: This term refers to the proficiency or rate of learning of new material. Put differently, learning efficiency refers to whether the child is slow or quick in learning a new concept or task to which he has not been exposed in the past. This should be observable from his behaviour on a learning task or item(s).

Asian: For the purposes of this study, an Asian child is defined as one whose parents or grandparents originally came from India, Pakistan or

Bangladesh. Although of Asian descent, children of parents who have emigrated from East African countries in recent years would not be considered as Asian and would thus not be included in the sample.

The author is inclined to agree with Phillips that this is a "simple and crude" way of defining ethnic membership, but nevertheless is "a term which has currency in the field despite its vagueness" (Phillips, 1979, p.117). Asians as thus defined (irrespective of whether they are Indian, Pakistani or Bangladeshi) all share common linguistic and social difficulties (Schools Council, 1970).

Children from Ethnic Minorities: This classificatory designation is used for all those children whose parents' original country of origin, values, patterns of thought, language, or customs, are different from the dominant culture in which they live. This term is used in preference to labels like "culturally handicapped", "culturally disadvantaged", or "culturally deprived", as this classificatory terminology implies value judgements. It is unacceptable that people should use prejorative terminology because one culture differs from the dominant culture.

Race or Ethnicity: In this work, the word "race" has been used synonymously with "ethnicity" as a form of shorthand.

CHAPTER 1

INTRODUCTION

For nearly two decades now, the performance of children of ethnic minority groups on tests of ability and attainments, has been a subject of considerable interest to many investigators. Several of these studies have found that children from ethnic minority groups do not score as well on attainment and IQ tests as their indigenous counterparts. Ghuman (1975) and Hill (1976) report several studies (e.g. Craig, 1963; Little et al., 1969; Saint, 1963), which have found appreciable differences in ability and attainments in favour of English children. More recent studies (Essen & Ghodsian, 1979; Ghuman, 1980; Mabey, 1981; Philips, 1979; Scar et al., 1983) provide further evidence that children from ethnic minority groups tend to perform less well than their English counterparts (for a further discussion of this topic see also pages 16 to 21).

Jensen, who has studied the performance of different populations in the United States on tests of attainment, ability and learning skills, has added a further dimension to this issue. He has attempted to explain the differences in the performance of various groups that he has studied by his two level theory of mental abilities (Jensen, 1969; see pages 3,4 & 40 for a description of these two levels). Jensen claims that the origin of his theory lies in his observations that:-

> "...low IQ children called 'culturally disadvantaged' appear in certain ways to be considerably brighter than their more advantaged middle-class counterparts of similar IQ." (Jensen, 1969, p. 230)

In a number of studies, Jensen claims to have found that children from disadvantaged backgrounds with low IQ scores (60 - 80) (in this category Jensen seems to include "culturally deprived", low S.E.S. and children from different ethnic minority groups) demonstrated a level of ability on learning tasks which would not normally be expected from children with such low IQs or poor performance on attainment tests. There was, however, some correspondence between their performance on learning tasks and the teacher's estimate of these children's (disadvantaged children with low IQ) ability as observed either in the playground or in social situations. This was not the case with children from upper-middle class homes:

> "...upper middle-class children in the same IQ range (60 - 80) performed on the learning tasks in a way that was consistent with their low IQs and poor scholastic performance - they were consistently slow learners in a wide variety of situations." (Jensen, 1969, p. 232)

In one of his earlier investigations Jensen (1961) compared groups of Mexican, American and Anglo-American fourth and sixth grade children on a number of learning tasks. The subjects were matched for IQ, socio-economic status of the family, age and language spoken in the home. The learning tasks consisted of immediate recall, serial learning, and paired associates. The study showed that Mexican-Americans of low IQ were superior on learning tasks as against Anglo-Americans of the same IQ. No significant differences were found in the learning abilities of high IQ Anglo-Americans and Mexican-Americans. From these findings Jensen suggests that the majority of Mexican-Americans with low IQs have learning potential within the normal range. That these children's attainments are low, Jensen argues, is not due to their inherently low learning ability, but may very well be due to other factors. On the other hand, low IQ among the Anglo-Americans group was an index of poor learning ability.

Jensen seeks to explain this interaction between IQ, learning tasks and different ethnic groups in terms of minimum hypothesis, which is a hierarchical model of mental abilities. This model subsumes two types of abilities, Level I and Level II. Level I ability (associative type) is best measured by tasks such as digit span, serial rote learning, paired associate, free recall and trial and error learning. An important feature of this ability is that little mental manipulation of the information is required prior to its output. Level II ability is akin to Spearman's g, and is best measured by tests such as the Raven's Matrices, tests of general and fluid intelligence, by experimental learning tasks and by Piagetian tests (Jensen, 1973, 1980 b). Unlike Level I, Level II ability involves mental manipulation of sensory inputs.

Jensen (1973) states that although several of his studies support his thesis, he could not yet claim its "generality" and therefore his theory needs to be examined in a variety of different populations (p.264). The other features of his theory and the relevant researches will be described in the next chapter.

Whilst Jensen has attempted to explain the interaction between IQ and learning tasks in different cultural groups in America in terms of his theory of mental abilities, some researchers in Britain (e.g. Little, 1975; Mabey, 1981) have found an association between the different ethnic groups' poor performance on ability and attainment tests and such factors as limited education in this country, cultural differences, "inability" or "unwillingness" of the present educational system to adapt itself to meet the needs of these children, social deprivation, the interference of the dialect, low teacher expectation and negative attitudes of teachers. It would seem that none of the researchers in this country have attempted to examine the ethnic group differences in the light of Jensen's two level theory.

Jensen's theory and the empirical evidence that he claims his studies provide, raise several important issues and have a direct bearing upon the apparent differences in ability and attainment shown between different ethnic groups in Britain, measured by current tests of ability and attainment. Is the "phenomenon" observed by Jensen peculiar to low socio-economic/minority group children in America only? If a similar type of study were carried out in Britain involving different ethnic groups which were assessed on tasks measuring similar functions to those Jensen has used in his investigations, would there be similar results to Jensen's? Would there be a proportion of children from minority groups with low IQ scores who would perform better on the learning tasks compared to their performance on the conventional IQ tests? Would there be any correspondence between their learning ability scores and their performance in academic subjects? Would there be any differences in the relationship between Level I and Level II abilities in different ethnic groups? Would Jensen's model still hold if other measures of Level II ability (e.g. Piagetian conceptual tasks; cf. Jensen, 1973, 1980 b) were employed in assessing the conceptual ability of children from ethnic minority groups?

In the current " social educational milieu"* these seem important questions and deserve answering, particularly as Jensen's views do tend to bias the opinions of many teachers and of other related professions about the intellectual ability of children of ethnic minority groups. In a recent study, Ghuman (1980) has observed that Jensen's claims that blacks have lower genetic potential have been widely publicised in Britain by Eysenck. This bias was confirmed when in a collaborative study (Thomas and Ghuman cited in Ghuman, 1980) the authors found that many teachers regarded low genetic potential as the chief cause of Sikh and West Indian children's poor performance on attainment tests in Cardiff. Such a belief, Ghuman concluded, absolved teachers and others concerned with children from taking any "positive and constructive" steps to reduce the gap between the performance of indigenous and black children.

What are the other practical implications of Jensen's researches in this country insofar as his findings indicate that the low IQ of

---

* See American Psychologist, 36, 1981 - the whole issue is devoted to assessment.

low socio-economic status ethnic minority group children is not a valid indicator of their learning ability? It would seem that Jensen's theory and his findings lead us to question one of the important concepts (i.e. intellectual ability), that is traditionally used when determining the future educational needs of children from ethnic minority groups. Although a large majority of psychologists are aware that the traditional IQ tests are inappropriate for the ethnic minority children (cf. Mackenzie, 1980), they continue to be used as one of the important criteria for the purposes of classifying and placing these children in special schools and units. Commenting on the use and relative importance of IQ tests for the purposes of decision making, Tomlinson (1981, p.294) says:-

> "Although the psychologists regard IQ as only one factor in the assessment process, it is still seen as an important factor and the dubious history of the application of IQ testing...does not seem to worry psychologists unduly".

Since the majority of the existing IQ tests have been mainly designed for and standardised on western-type populations, they should not really be used to make crucial decisions about children whose previous experiences and culture differ substantially from the former group. In testing children from ethnic minority groups, and using the existing IQ tests on them, one needs to bear in mind that there is ample cross-cultural evidence which suggests that different cultures tend to stimulate and favour the development and application of different cognitive skills and strategies for coping with their environments (Anastasi & Foley, 1949; Berry, 1981; Bruner, 1966; Cole et al., 1971; De Vos & Hippler, 1969; Ferguson, 1954; Ghuman, 1975; Vernon, 1969).

On this theme, Telford & Sawrey (1977, p. 351) comment:-

> "No single set of aptitudes is the pre-requisite for survival and adaptability in all environments. Every

> society requires individuals capable of performing the necessary social and economic functions of that society, and inevitably will favour and reward aptitudes, behaviour and values contributing to those capabilities."

Vernon, who has done a considerable amount of cross-cultural research, writes:-

> "We must try to discard the idea that intelligence is a kind of universal faculty, a trait which is the same (apart from variation in amount) in all cultural groups. Clearly, it develops differently in different physical and cultural evironments. It should be regarded as a name for all the various cognitive skills which are developed in, and valued by, the group". (Vernon, 1969, p. 10)

Two studies can be cited in order to support the view that different cultures encourage the development of certain cognitive processes. Davidson & Klich (1980) investigated the role of cultural factors in the development of temporal and spatial ordering in Australian desert aboriginals. Thirty boys and 35 "full-aboriginal" girls, aged 9 years to 16 years 4 months (mean age = 12.7), were administered two free recall tasks; one recall task involved pictures, while the other task involved natural objects. Davidson & Klich found that compared to the "normal" western children, the Australian desert aboriginal children showed preference for spatial over temporal recall order. From their findings the authors suggest that:-

> "...preference for temporal ordering as measured by performance on free recall tasks may vary due to cultural and evironmental influences in children's cognitive development." (Davidson & Klich, 1980, p.571)

Cole et al.(1971) found that on certain tasks the tribal people of Liberia performed better than their American counterparts. For instance, the authors of this investigation discovered that the tribal people were significantly better at estimating various amounts of rice compared to the American sample. Cole et al. attribute this difference to the fact that rice farming is an integral part of the culture of these tribal people and it also involved numerous related activities as well. The truth of the matter is that different cultures tend to make different sorts of intellectual demands and people will be good at doing things that are important to them and they have occasion to do often (see also investigations by Bower (1982), Bruner (1975) and Schaffer (1977) which highlight the role of mother-infant interactions in advancing cognitive development).

It has been noted by Salvia & Ysseldyke (1978) that acculturation is the most important element in judging a child's performance on a test. If the acculturation of the child differs from the acculturation of those children on whom the test was standardised, the test results can be both invalid and biased. Salvia & Ysseldyke, therefore, recommend that when standardised tests are used for determining the future educational needs of children, their level of acculturation should always be born in mind (see also, McShane & Plas, 1982).

Despite the availability of such empirical evidence which highlights the role of culture and experiences in stimulating or retarding certain aspects of cognitive development, in the author's experience, many practising psychologists continue to use IQ and other unsatisfactory and suspect procedures to determine the intellectual level and learning ability of children from ethnic minority groups. One of the IQ tests most commonly used by practising psychologists is the WISC-R (1974) which has both a Verbal and Performance Scale. Some psychologists, a little more aware of the culture bias in IQ tests than others, tend to omit the Verbal Scale of the WISC-R and administer the Performance Scale only. Their rationale for using the Performance Scale seems to be that as the subtests in this scale do not require Verbal responses, do not overtly appear to be confounded by the child's previous experiences compared to the Verbal Scale, and can be administered with relative ease, they are therefore unlikely to penalise children with different cultural backgrounds. In the main, non-verbal scales of many kinds are regarded as being capable of providing a reliable, valid and fair index of ethnic minority children's mental processes. Hence many psychologists see little objection to their use on such children in order to assess their cognitive processes. Unfortunately the acceptance of these non-verbal tests as adequate and fair measures of these children's mental abilities leads to inherent dangers and fallacies. This matter will be discussed at some length in the next chapter. Suffice it to say here that the upshot of depending heavily upon unsatisfactory and biased procedures to determine whether members of minority groups should remain in the mainstream of education or should be recommended for transfer to special schools/units is that serious misclassifications are bound to occur.

Hegarty (1977) in evaluating current assessment practices, wrote that children from non-English speaking homes were at a disadvantage in British schools. Pointing particularly to their linguistic handicap, Hegarty argues that they perform poorly in schools compared with indigenous children, which in turn means that they are more liable to be sent for remedial assessment. Furthermore:-

> "The forms of assessment in common use were likewise biased in favour of the indigenous population so that the disadvantage was compounded.
>
> The practical outcome was that immigrant children came to be over-represented in ESN schools. In the early seventies there were four times as many West Indian children, relative to their number in the population, in these schools as there were indigenous children". (Hegarty, 1977, p.39)

It has been observed not only in Britain but also in the United States (Tucker, 1980) that children from ethnic minority groups are over represented in special units and schools. Is it therefore surprising that leaders from certain minority groups have started to demonstrate concern about the disproportionate number of their children in special schools because of the inaccurate assessment of their needs and potential? (For a detailed review of this issue, see Tomlinson, 1981: Townsend, 1971; Townsend & Brittan, 1972; cf. Larry P.V. Riles, 1979 cited in Bersoff, 1981) In America the indictment of racial, ethnic and cultural bias against the testing procedure, and assessors, has led to the involvement of their judicial system. The findings of the judiciary appear to confirm the existence of discriminatory practices in the procedure followed in the assessment of black children. Consequently, the use of standardised IQ tests for the purposes of identifying and placing these children in educable mentally retarded classes has been prohibited by the judicial order (Duffy et al., 1981). Commenting on the current American attitudes towards the use of intelligence tests, Vernon (1979 pages 2 - 3) writes:-

> "...the testing movement, long regarded as a major achievement of applied psychology, and accepted by most laymen as veridical, is now widely distrusted and criticised, and is even in some danger of abolition in the United States, where it once flourished most luxuriantly. Several States have passed, or at least considered, laws to ban the use of IQ tests in schools, on the grounds that they are culturally biased and do not accurately measure intelligence. Many American parents have successfully challenged in the Courts the allocation of their children to Special Schools or classes on the basis of low IQs. It has also been ruled in some suits that employers cannot refuse to employ blacks or others who obtain low test scores, unless there is clear evidence that suitability for the job depends on what the test measures."

The growing concern about the limitations of IQ tests, especially in relation to their use on minority group children (Cole, 1975; Gallagher, 1976; Haywood et al., 1975; Hegarty, 1977; Lambert et al., 1974 see also McReynolds, 1982; McLoughton & Koh, 1982; Sattler, 1974), has inspired several researchers (Budoff, 1973; Haynes, 1971; Haywood et al., 1975; Hegarty & Lucas, 1978) to examine alternative methods of assessment which attempt to surpass, to a large extent, the confounding influence of previous learning on the present performance. What these researchers have put forward, as an alternative form of assessment, is the notion of the assessment of learning potential. The aim of this method of assessment is to determine how well a child learns a task or a concept to which he has not been exposed in the past. This method of assessment, according to Vernon:-

> "...should overcome the difficulty that different groups of individuals within a group will always be at different stages of familiarisation." (Vernon, 1969, p. 106)

Furthermore, it is also pertinent to find out how well a child is likely to perform/learn in future. It seems irrational, continues Vernon, to administer any performance test to a child from the ethnic minority groups, who has little command of the language, with a view to predicting his capacity to learn English. Vernon further adds:-

> "It is indeed curious that we use intelligence tests mainly to predict capacity for learning and yet none of our subtests involve any learning: instead they give a cross-section of what has been learned." (Vernon, 1969, p.106; see also Cole, 1975; Glaser, 1981; Jensen, 1963.)

Hegarty & Lucas (1978) who have examined the role of learning ability in the assessment procedure, also advance several reasons in favour of using this technique. In summary, Hegarty & Lucas suggest that the assessment of learning ability has a direct bearing on learning and teaching. Unlike conventional assessment where the main emphasis is on performance, in the assessment of learning ability the main focus is on the processess that underlie it. Yet another feature of tests of learning ability is that they enable one to equate the previous experiences by using tasks which, as far as possible, are equally unfamiliar to all children (<u>cf</u>. Vernon, 1969). An additional advantage of assessing the child's learning ability is that, since IQ is not an index of a child's rate of learning, with this measure we can identify children of different learning abilities but in the same IQ range. Hegarty & Lucas observe that in their study:-

> "There were many misclassifications in terms of IQ as measured by simultaneous learning ability scores and by attainment scores obtained after a year's interval. For example, the child with the lowest IQ (64), scored better than average on two of the learning ability subtests. One child with an IQ of 72 scored better than average on all but one of the learning ability subtests and the attainment tests; on two of them his scores were much better than average, i.e. more than two standard deviations above the mean. Of the 28 children with IQs below 75, twenty-five scored better than average on at least one of the learning ability or attainment tests, and fourteen did so on two or more of the eight tests." (Hegarty & Lucas, 1978, p.42; see also Jensen, 1963)

Hegarty & Lucas also claim that for certain groups of children tests of learning ability are more effective than IQ tests in predicting school achievement. Learning tasks, as opposed to IQ tests, offer a better opportunity to the psychologist to observe such attributes as the child's attention span, perseverance, learning style, style of tackling unfamiliar tasks and so on - all of which are of considerable importance to future success or failure (Cattell, 1965; Mackay & Vernon, 1963). Finally, the concept of learning ability is of considerable theoretical importance in its own right.

Notwithstanding the usefulness and importance of assessment of learning ability in the assessment procedure, hardly any well designed tests which measure this function are available in this country. The only test in Britain that has been standardised on ethnic minority children and is based on the notion of learning ability is the NFER's Test of Children's Learning Ability (1978). Although this test departs in its rationale and administration procedure from commonly used IQ tests, it has a few serious limitations. For example, it is designed to meet the needs of a very limited age range (First Year Juniors i.e. 7 - 8 only). Although the sample does not include indigenous children, some of the materials used would appear to have a western-bias (see discussion on pages 56 to 58. The other serious limitation is that the battery provides little information about the child's level of cognitive processes. For the purposes of determining the child's future educational needs and for designing an appropriate curriculum, information about the child's current level of thinking would appear to be of vital importance. Matching the curriculum with the child's level of mental processes is absolutely vital to general cognitive development because even providing rich or varied experiences can result in rote learning only (Kamii, 1974; Wadsworth, 1978). More recently, Toepher (1981) has advanced the view that learning, particularly high level learning, should be matched with the individual's cognitive skills. High IQ and "satisfactory record" alone are not sufficient to cope with high level learning. Toepher came to this conclusion on the basis of his enquiry based on 1,700 (12 to 14) year old children with IQs about 120.

Notwithstanding the shortcomings of the researches of both Haynes and Hegarty & Lucas, they would seem to have clearly established the importance of including the testing of learning ability, especially of children with different cultural and linguistic backgrounds, amongst other assessment procedures. Findings such as theirs not only provide important empirical evidence, they also lead other researchers to design further tests based on the notion of learning ability, which avoid the limitations of their pioneer work.

Thus to summarise, the overwhelming weight of the evidence which has been discussed here has demonstrated that the assessment of learning potential is a far more suitable method of testing children from ethnic minority groups than the traditional forms of assessment (i.e. IQ). As at present hardly any such tests are available, the present study will address itself to developing a test battery which would attempt to meet this need: thus this would be the foremost goal of this enquiry.

In order to design such a battery of tests and standardise it on all the various ethnic groups the support of a fairly large research team would be needed as well as a considerable sum of money. As such resources were not available to this research project, it was limited to Asian children only. This is not to imply that the assessment of learning potential is in any way less relevant or appropriate for children who belong to the dominant culture or other ethnic minorities. It is therefore hoped that although the proposed battery will be designed principally with Asians in mind, it will prove to be useful for other children as well.

In addition to the chief goal of developing a test battery for Asian children, this study will also examine Jensen's hypothesis of Level I and Level II abilities to determine whether his model of mental abilities is generalisable to cultural groups other than those which he and the other workers have investigated so far i.e. English and Asian.

CHAPTER 2

REVIEW OF THE LITERATURE

2.1 British Studies Concerning the Performance of Ethnic Minority Children on Tests of Ability and Attainments

The major interest in this field began in the 1960's. These early investigations, and some carried out in the early part of the seventies, have been well reviewed by Essen & Ghodsian (1979), Ghuman (1975), Haynes (1971), Hill (1976), Little (1975), Mabey (1981) and Phillips (1979). The bulk of these studies seem to demonstrate that, on the whole, the performance on tests of ability and attainment tests of ethnic minority children tend to be inferior compared to the indigenous population. (See also Little, 1982; Mackenzie, 1980; Scar et al., 1983; Resnick & Resnick, 1982 for similar findings in America.) The other main finding which seems to emerge from the review of these studies concerning ethnic minority children, is a positive correlation between the length of education in this country and their performance on attainment and intelligence tests (see Spenser, 1982).

There are also a few studies (e.g. Taylor, 1973) which have demonstrated that Indian and Pakistani children who came to Britain when they were quite young when equated with a sample of English children in the same school achieved better educationally. The above summary findings were further confirmed by Monica Taylor from the National Foundation for Educational Research who has very recently carried out a very detailed review of the literature concerned with the attainments of Asian children for the Swann Committee (Personal communication, 1982).

Although some studies have found a positive relationship between the length of stay and academic performance. Ghuman (1975) has argued that despite the exposure of ethnic minority children to the British Educational system:-

> "...all the measured abilities are not equally effected. The Indian children (mostly Punjabi in this case) show marked progress on the range of verbal and educational abilities, but not on the spatial and perceptual abilities and this was true of the West Indian children. This deficiency in immigrant children may be due to lack of perceptual experiences in the home, especially in the early years." (Ghuman, 1975, p. 21)

Studies carried out more recently (Essen & Ghodsian, 1979; Ghuman, 1980; Mabey, 1981; Phillips, 1979), i.e. during the late seventies and in 1980 and 1981, do not reveal any marked change in their findings compared to the earlier studies. For instance, Essen & Ghodsian were concerned in their research with firstly the importance of allowing for differences in home circumstances when comparing the school performance of non-immigrants, and secondly with comparing the performance of over 8,000 sixteen year old West Indian, Asian, Irish, European and English children's performance on Maths and Reading attainment tests. Essen & Ghodsian draw attention to two points about their sample. One, that it was a nationally representative sample; second, that they also made a distinction between children born in this country to foreign born parents and between children born to parents who themselves were born in this country. All the children were administered a Maths and a Reading Test, especially designed by the National Foundation for Educational Research (NFER). The results showed that all the first generation ethnic groups obtained lower scores both on the Maths and Reading Tests compared to the indigenous group. Among the second generation ethnic groups, however, it was only the West Indian children whose mean scores were significantly lower compared to the English sample. All the second generation ethnic groups' Maths and Reading scores (with the exception of Irish children's Maths scores) were higher than the corresponding first generation ethnic minority children. Of all the ethnic groups, Asian children's mean scores were the highest amongst the second generation "immigrants" and the second generation West Indian children's scores remained the lowest, although it was better than the first generations. The data was re-examined after children from the various ethnic groups and the indigenous children were equated for home circumstances, including the language spoken at home. With the re-analysis of the data, two points emerged:- i) Once the differences in the home backgrounds were allowed for, the mean differences between ethnic groups and indigenous children were, in the main, substantially reduced. ii) As a result of equating home circumstances there were no significant differences between some ethnic minority group's performance and English children. The two groups which still had lower mean scores were the two West Indian and the first generation Asian children. Essen & Ghodsian summarise their findings by saying that their results are consistent with previous findings in demonstrating:-

> "...that while many immigrant groups have lower overall mean scores on attainment tests when children of similar financial and housing circumstances are compared only the West Indians have clearly poorer school performance than indigenous children." (p.427)

The next study to be considered is by Mabey (1981). Unlike many other studies concerned with children from ethnic minority groups, this was a longitudinal study initiated by the Inner London Education Authority in 1967. The entire sample (English, West Indians, Indians and Pakistanis) were first tested when they were eight years old; next when they were eleven, then at thirteen and finally at the age of fifteen in 1976. The aim of the survey was threefold: (a) to compare the reading standards of London children nationally: (b) to identify the factors that influenced these children's reading attainments: and (c) to identify children who were likely to be affected or disadvantaged. All the subjects were administered a specially designed and standardised reading test. The results for the third testing (i.e. when children were 13 years of age) for technical reasons not pertinent to the present discussion, were not included. Thus the data analysed was of the tests administered when the children were 8+, 10+ and 15+ years old. Except for 10% of the sample, the remaining children were administered group tests by their teachers. The results revealed that the reading attainments of West Indian, Indian and Pakistani children were lower than the English children. This remained so throughout the three phases of testing, i.e. when these children were 8+, 10+ and 15+ years old. These results, particularly those of the West Indian group as theirs was the lowest mean score compared to the indigenous sample (and has little relationship with the length of education), are discussed in the light of such factors as immigration itself, restricted education in this country, social and environmental deprivation, dialect interference, negative teacher attitudes, low teacher expectations and finally poor self image. The conclusion reached by the author is that possibly with the exception of self image, the other factors mentioned above explain about half the difference between the scores of English and West Indian children.

Phillips' (1979) survey, unlike the researches of Essen & Ghodsian and Mabey, was concerned with younger children - the average age being 7 years 3 months. Phillips' study was aimed at examining underachievement in different ethnic groups and reported the results of a survey carried out in the West Midlands in 1969. Nearly 2,400 English, West Indian and Asian children were administered the English Picture Vocabulary Test (Brimer and Dunn, 1963) and the Southgate Group Reading Test 1, Form 'C' (Southgate, 1959), during their last term of Infant schooling. The results showed that the average vocabulary and reading attainments of Asian and West Indian children were significantly below those of their English counterparts. For instance, on the English Picture Vocabulary Test, English children's mean scores were 99.3 (sd 15.3), West Indians' 79.5 (sd 12.4) and the Asians' 70.3 (sd 12.7). and the reading scores for these groups in the same order were

99.4 (sd 15.3), 89.9 (sd 14.6) and 84.4 (sd 16.1) respectively. Despite these significant differences, the author, by using a rather complex statistical argument, argued that Asians and West Indians were not underachieving and concluded:-

> "...West Indian and Asian children, on average, acquire basic educational skills in British schools at least to the levels of their abilities, when attainment and abilities are assessed on parameters which describe the indigenous population". (Phillips, 1979, p.128)

The next study (Ghuman, 1980) is concerned with the cognitive styles of a total of 136 children (13 - 14 year olds) from three ethnic groups: English, Asian and West Indian. The sample was drawn from a comprehensive school situated in an inner ring area of a city and included children from categories III, IV and V, according to the Registrar General's classification. They were all given the Group Embedded Figures Test, the NFER's Spatial Test EG and also NFER's Mathematics Attainment Test (EF). Three hypotheses were tested: two embedded in the socialisation practices of the three ethnic groups, and the third in the findings of previous researches. The results did not support either of the "major" hypotheses, neither did they show whether there were any ethnic or sex differences on the cognitive style measure. Failure to find the expected differences was explained in terms of the effect of indigenous culture on the cultures of immigrant communities as a result of living in this country and the influence of schooling which Ghuman speculates may be even more important than early socialisation practices. The results on the Maths test were in the expected direction, that is, Asian children scored higher than both the English and West Indian groups, although even their score was below the population norm.

To conclude, the bulk of the studies above confirm the main trend of earlier studies, i.e. the scores of children from ethnic minority groups on scholastic and IQ tests tend to be inferior compared to their

English contemporaries or national norms. Some studies have also demonstrated a positive and significant correlation between the length of education in this country and these children's under-achievement. There are also a handful of studies where the children's performance, particularly Asian children's, equal or is better than their English counterparts; however, by no means can these findings be viewed as representative of the main trend of the majority of studies in this field. Together with these summary findings it is also worth remembering that even if an Asian child has had all his education in this country, because of his early upbringing and cultural differences his pattern of cognitive abilities may still be different from that of the indigenous child (cf. Anastasi & Foley, 1949; Bruner, 1966; Cole et al., 1971; De Vos & Hipler, 1969; Ghuman, 1975; McReynolds, 1982). At the same time it should not be overlooked in view of current thinking and substantial evidence "that test scores are related to acquired abilities, life experiences, and educational opportunities ... we no longer under-emphasise the fact that aptitude tests give weight to educational advantage or disadvantage" (Glaser, 1981, p. 925).

This not only seems to explain why the performance of children of ethnic minority group tends to be inferior compared to that of the indigenous population but also underlines the inappropriateness of using conventional tests on these children.

2.2 American Studies Concerning the Performance of Ethnic Miniority Children on Tests of Ability and Attainments

In the early thirties several workers in the United States recognised that many children from ethnic minority groups had low achievement and IQ scores on traditional tests. Jensen, who addressed himself to the issue of the differences in the achievement of "certain ethnic or national groups" and Anglo-American children wrote:-

> "The apparently poor scholastic aptitude and achievement of a large proportion of the children from certain ethnic or national groups in the United States is a long recognised problem which seems to gain importance over the years." (Jensen, 1961, p. 147)

Jensen, on examining the performance of Mexican American children, a large majority of whom were "slow learners", found their numbers far exceeded what one would expect to find in a normal distribution. The scores of these children on a variety of tests tended to be lower than Anglo-American children. Jensen cited many studies which demonstrated that on average the scores of these children on the Binet cluster around 80. He also added that evidence shows that translation of the Binet into the mother-tongue of these children did not change the pattern of their scores either, i.e. their scores on a conventional IQ test did not come closer to the Anglo-American population.

Jensen discounts such factors as language handicap, or the low socio-economic status of these children as the underlying cause of their inferior performance on tests of attainments and intellectual ability. Instead Jensen advances the view that:-

> "...the low IQ Mexican-American children have not acquired in their environment the kinds of knowledge, habits and skills that provide the basis for school learning and which are tapped by IQ tests. We have evidence that an elaborate matrix of previous learning underlies a person's verbal learning ability. Frequency of past exposure to verbal stimuli is correlated with ease of learning where verbal materials are involved. For example, it has been shown that the speed of learning a list of non-sense syllables is highly correlated with the frequencies of the letters in our language that compose the non-sense syllable." (Jensen, 1961, p. 157)

While advocating measures of learning for ethnic minority children on the grounds that they can enhance the diagnosis of educational disabilities he regards the use of IQ tests on such children as unsuitable. Jensen criticises the use of both verbal and non-verbal IQ tests on ethnic minorities (who have not been exposed to the dominant culture to the same extent as the indigenous population) on the grounds that they are measures of previous learning and are heavily dependent on verbal mediation (see also Resnick & Resnick, 1982).

These studies show that children from ethnic minority groups tend to achieve less on scholastic and IQ tests compared to children from the dominant culture and that this is not peculiar to Britain but also seems to be true in the United States as well. For several years now many practicing educational psychologists have been fully aware of these findings. Their cognizance of these findings has led many of them to abandon the use of verbal tests on ethnic minority children as they obviously seem far more biased. However, many educational psychologists use and regard non-verbal or performance tests (these two terms will be used interchangeably; cf. Butcher, 1968) as quite adequate and fair in determining the intellectual level of children from ethnic minority groups.

2.3 Non Verbal Tests

Performance Scales no doubt are easier to administer to children who have a limited knowledge of the language compared to Verbal Scales and have the appearance of being "culture fair", but are not really free from the environmental influences (Butcher, 1968). For instance, Ortar (1963) in Israel tested the hypothesis that the Performance Scale of the WISC would show smaller differences between high-status Israeli children and oriental immigrants compared to the verbal test scores. Ortar studied five groups of children differing in socio-economic status and level of acculturation. The relevant findings of Ortar's study showed results in an unexpected direction. She found that verbal tests surprisingly placed all five groups in the expected order whilst the scores on the performance test were in favour of children who belonged to the higher socio-econmic status and to children who were longer-acculturated. Ortar, whilst testing children with different levels of acculturation also noted that when, for example, oriental immigrant children were presented with the picture of a head from which the mouth was missing, (an item from a subtest of the Performance Scale of the WISC and WISC-R), when asked what important part was missing they said that the body was missing. These children gave this answer because of their lack of familiarity with the convention of considering the drawing of a head without a mouth, the accepted notion of the drawing as a complete picture. For these children the omission of a body from the picture was more important than just the mouth. The weaknesses of the common assumption that non-verbal tests are "nearly" culture free and measure the same processes as verbal tests fairly across various cultures, have been treated at some length by Anastasi (1976). Butcher (1968) is of the opinion that psychologists who assume that non-verbal or performance tests are "a sure key to the problem" of assessing children of minority groups are "far too simple minded" (p.254).

2.4 Unstandardised Mini Learning Situations

As well as relying on performance scale, the other assessment approach that is currently fashionable with many psychologists is to obtain some information about ethnic minority children's learning ability on a "mini learning situation". The mini learning situations are devised on an ad hoc basis; they range from adapting the Block Design (a subtest from the WISC-R (Weschler, 1974) or employing commercially produced materials, to sometimes presenting a child with a few simple words or letters to learn in the test situation.

Conceptually it is an excellent idea and would seem to suggest that it surmounts the often voiced criticisms against the use of IQ tests on minority groups. After all, what else would inform the psychologist better about the child's learning ability than actually observing the child in a learning situation? Highly commendable though this approach is, the way it is often used has some inherent weaknesses. For example, assessing children's learning this way is neither objective, reliable, valid nor rooted in any learning theory, nor is it based on any empirical evidence which provides even tenuous justification for its use.

In the selection of these mini-learning items, in the author's opinion, some factors that need to be taken into consideration are totally ignored. (For a detailed discussion of the principles that should be taken into consideration in developing items for children with different cultural backgrounds see Section 3.2.in Chapter 3.) One cannot help but regard these mini-learning situations as having an appealing semblance of validity in the assessment procedure employed by an educational psychologist but, in reality, adding very little either for proper decision making about the child's future educational needs or for remediation or for curriculum planning.

There is yet another serious criticism which can be levied against the use of mini-learning situations, particularly when the items are selected from one of the academic subjects (e.g. Reading), with children with learning difficulties. Some of these children referred to the Child Guidance Clinics may have a long history of failure in basic subjects. A detailed assessment of such children (which invariably also includes interviews with parents and teachers concerned) would often reveal that where there has been a long history of failure in a particular school subject, using learning tasks in that subject can be highly anxiety provoking for the child and thus such a mini-learning situation would fail to serve its purpose. Not only would this type of assessment waste the psychologist's time, but more importantly, may provide misleading information about the child's learning ability.

From the foregoing discussion it would seem that the present procedures employed by a large majority of practising psychologists cannot by any standards be regarded as scientific or satisfactory or fair. However, a few psychologists who were faced with a similar problem of assessing children with diverse cultural backgrounds have made serious attempts to resolve this dilemma. Their research efforts, together with other pertinent studies, are presented in the next section.

2.5 Studies Based on the Notion of Learning Ability

In the United Kingdom, Haynes (1971) was among the pioneer workers who provided a lead by developing a battery of tests which attempted to circumvent many of the now well recognised objections to the current practices of assessing minority children (see also Feuerstein, 1980; Haywood et al., 1975; Hegarty & Lucas; 1978; Lambert et al., 1974). The result was an unpublished test battery based on the concept of learning ability.

Hayne's research involved 7 to 8 year old children (125 Sikh and 40 English) from the Southall district of the London Borough of Ealing. The major goal of the study was:-

> "... to devise tests of learning ability which can be used to assess the abilities of children with all degrees of linguistic and other cultural handicaps." (Haynes, 1971, p.29)

As well as attainment and IQ tests, Haynes also administered five self-devised learning tasks to her total sample. The results revealed significant differences on four out of five learning tasks in favour of English children. Significant differences were also found on all the attainment and most of the IQ tests used in the study except for the WISC Coding and Draw-a-Man Tests.

The key finding of this investigation was that Hayne's battery of learning tasks was a better predictor of 125 Punjabi children's academic progress than the Performance Scale of the WISC (Weschler, 1949). Although Haynes made a significant contribution by highlighting the importance of assessing learning ability of children with diverse cultural backgrounds, rather than mainly relying on the use of IQ and attainment tests, her results suggest that her learning tasks seem to suffer from a similar kind of culture bias as is reported in many of the commonly used IQ tests. If Haynes' learning tasks had tapped mental processes, which are reasonably equally favoured in the two cultures, it is reasonable to speculate that there would not have been significant differences in the favour of English children on the four out of five learning tasks (see also pages 75 to 79).

Despite this weakness in Haynes' learning tasks, its "novel approach to testing that offered a way round some of the obstacles associated with assessment in multicultural situations" (Hegarty, 1977, p. 4) led to the NFER's decision to develop and extend Haynes' materials when they were commissioned by the Department of Education and Science to:-

> "...devise and validate tests which would be used effectively to predict the learning potential of children whose linguistic and cultural background precludes the reliable use of more traditional forms of assessment." (Hegarty & Lucas, 1978, p.73)

Adhering to the same rationale as Haynes, the NFER produced two sets of test materials, one for individual use and the other as a group test which could be used as a screening device. The standardisation sample consisted of 386 seven to eight year old Pakistani (n = 175) and West Indian (n = 211) children which were drawn from 68 schools in 14 different education authorities.

In order to compare the predictive validity of the Learning Ability Test battery with the short form of the WISC, Hegarty & Lucas carried out three forms of statistical analysis: correlation, canonical correlation and stepwise regression analysis. On the basis of their analysis, Hegarty & Lucas came to the conclusion:-

> "The consistent picture that emerged is that the Learning Ability Test battery is a better predictor of subsequent attainment." (Hegarty & Lucas, 1978, p.81)

The limitations and strengths of the above two studies have already been commented upon in Chapter 1 and therefore need not be restated. Suffice it to say here that despite the shortcomings of Haynes and Hegarty & Lucas' work, they have provided a significant lead in the field of testing, but unfortunately their work, or for that matter the notion of learning ability as an assessment tool, on the whole, has generated very little interest in the United Kingdom. Between 1960 and 1980, apart from Haynes and Hegarty & Lucas' investigations, just two more studies have been conducted where the notion of learning ability plays a key role.

The first one is by Mackay & Vernon (1963). The main aim of this study was to devise tests of "actual learning" and compare their validity in predicting children's future academic progress in the basic subjects with conventional IQ tests - which hardly ever have items which really test the child's ability to learn. The sample consisted of two groups of 8 - 9 years and 10 - 11 year old children with a good middle class background. All these children were administered nine self devised learning tests and ten reference tests which included tests of intelligence, attainments, memory for digits, spatial and the Bender Gestalt Test. For the purposes of the analysis of the data the authors used terminal scores. The results were in the expected direction with the 10 - 11 year old children, but somewhat "disappointing" with the younger group. With the 8 - 9 year olds, the results tended to be "specific" and less predictive. Despite the disappointing results, the authors seem quite optimistic in that they go on to recommend that it should be possible to devise a battery of learning tasks which should be based on a large sample, and such a battery of learning tasks would not only provide information about children's future academic progress but would also be useful in providing information about their style of working, level of concentration, etc. Mackay & Vernon do not speculate in favour of individual learning tests:-

> "...because these would have to be very lengthy to allow adequate practice, and because of a child's reaction to the learning situation in the clinic setting might differ considerably from his reaction in the classroom." (p. 185)

In this study, although Mackay & Vernon are fully aware of the fact that their results are based on a small sample and any interpretations therefore need to be made with considerable caution, the authors seem to be making claims which a study based on such a small sample does not warrant. For instance, Mackay & Vernon (notwithstanding the small size of the sample) offer quite an unguarded, or unqualified, opinion about the use of gain scores and whether tests of learning should be individual tests, or a group test, and so on.

However, Mackay & Vernon deserve credit for being ahead of their times, at least in Britain, in questioning the commonsense but fallacious thinking of equating intelligence with learning ability.

The other study which also embraces the concept of learning is by Kroeger (1980). Kroeger investigated the role of teaching in the evaluation of learning ability in "migrant" children. Kroeger, in this study was concerned to determine:-

> "...whether "actualisation" takes place or whether there is a true developmental lag between indigenous and immigrant children..." (p.106)

Two hundred and sixty-one boys only, aged 8 years 3 months to 10 years 4 months, took part in the study, of which there were 108 English, 97 West Indian and 56 Indian children. Higher proportions of Indian and English children came from social class I and II: most of the West Indian children came from the lower occupational groups - in fact none of the children's parents held jobs which would be classified as belonging to social class I or II. All the subjects were administered scholastic achievement tests, Piagetian Matrices and the Raven's Coloured Progressive Matrices. In addition, information about the subjects' emotional, social and scholastic adjustment was also obtained. Of the total sample, 49 children were allocated to the Experimental Group, 65 to the "Training Only" Group, 53 to the Control Group, and 94 were classed as "Passers". Two main findings deserve reporting. With regard to the main aim of the study, the author maintains that children from ethnic minority groups did possess competence, but as yet, were unable to perform on the:-

> "...operation demanded, a fact which is corroborated by the finding that West Indian boys improve significantly even after re-testing only - a "triggering" effect can be assumed..." (p.114)

But more important and pertinent to the present discussion is the author's argument in favour of using learning test scores for the purposes of predicting future academic performance as opposed to conventional IQ tests because of the significant correlation between the training score and the improvement score for Indians, English and West Indian children, and a positive correlation with vocabulary score and scholastic achievement - at least for West Indian children. The results also showed that although there were significant differences in the pre-test scores of the three ethnic groups on the Piagetian Multiple Matrices, the West Indian and Indian children benefited more as a result of ten minute training sessions (as was shown by their improved scores on the Piagetian Matrices), than the indigenous group. In the case of the Indian group, the training effect to some extent also transferred to their performance on the Raven's Matrices. Findings such as these should make psychologists extremely sceptical of treating the initial scores of children from ethnic minority groups as valid indicators of their intellectual potential, even though they are shown to be reasonable predictors of children's academic progress.

This would seem to be the sum total of research effort in the domain of learning ability in Great Britain. Although there is clearly a dearth of literature in this field, studies which have been carried out, irrespective of their constraints, shortcomings and disappointments, seem to reiterate the point that in order to measure children's ability to learn, especially that of ethnic minority groups, tests of learning ability should be used rather than conventional IQ tests, which provide little information about the child's ability to learn.

In the United States, unlike Britain, the area of learning ability has been a subject of much more theoretical, as well as of research interest. This may well be due to the fact that American psychologists have a wider experience of assessing the needs of individuals with diverse backgrounds (Anastasi, 1976) and they have also long recognised the limitations of IQ tests (Jensen, 1961, 1963) as measures of abilities of children with different backgrounds and experiences. It is therefore hardly surprising that the justification of using IQ tests on minority groups in the United States these days, is being examined and questioned not just from the psychometric standpoint but also from social, economic, ethical, political and legal standpoints.*

---

* See American Psychologist, 36, 1981: the whole issue has been devoted to the various aspects of assessments including ethnic minority groups: Cole, 1975; Haywood et al., 1975: Jensen, 1961, 1963; Lambert et al., 1974; McLoughton & Koh, 1982.

2.6 Studies Concerning Learning Ability in the United States and Israel

Milton Budoff and his associates at the Cambridge Laboratory, Haywood and his associates at the Nashville Laboratory (both of these groups in the United States) and Feuerstein in Israel are among the leading psychologists who have made significant contributions in the field of learning ability. Budoff has studied "educable mentally retarded" children, Haywood has studied "cultural-familially retarded" children and Feuerstein has studied "immigrants" in Isreal. Although each of these workers has offered different explanations as to why their subjects' performance was considerably lower on the standardised tests, they all came to the same conclusion that conventional IQ tests were not suitable for assessing these children's intellectual ability or drawing inferences about their learning potential: they thought that IQ was an invalid and biased index of these children's learning ability (see also Rohwer et al., 1971). In the succeeding pages the representative work of Budoff, Haywood and Feuerstein which highlights their views on the issue of learning potential will be reviewed (for further details of their work see Haywood et al., 1975; Hegarty & Lucas, 1978; see also Feuerstein, 1980).

2.6.1 Haywood and His Associates

At the Nashville Laboratory, Haywood and his associates discovered that standard intelligence tests under-estimated the ability of "cultural-familially retarded" individuals to form verbal abstractions. These writers claim that such individuals are not necessarily deficient in forming verbal abstractions but suffer from information deficit. Haywood et al. (1975) regard verbal abstractions as being at the heart of social interaction in our day to day living. By verbal abstraction they mean those activities which involve grouping and classifying isolated events and giving them abstract labels on the resultant categories. For instance, various fruits such as banana, grape, plum, apple, etc. can be grouped into a single category and given the abstract label fruit. Haywood et al. further explains:

> "In a sense, the individual performs a factor analysis on the data that impinge upon his senses; that is, he uses a central process to reduce a large number of isolated events to a smaller number of abstract categories to determine whether a new event can be assimilated into an existing category. If so, the new event can be understood more easily." (Haywood et al., 1975, p. 104)

In order to provide empirical evidence to the hypothesis of information deficit in cultural familially retarded persons Gordon & Haywood (1969) administered a twenty two item similarities test (a subtest from the WISC) under two conditions to known organically retarded and cultural-familially retarded institutionalised subjects. The two conditions of administration of the test were: standard and 'enriched'. In the latter condition, there were five exemplars instead of two as in the standard condition. In addition to organically retarded and culturally retarded groups the authors also included a group of non-retarded children who were equated with the retarded children on mental age. The results of this experiment supported the prediction. Under the standard condition the performance of the mentally-age matched, non-retarded group was significantly better than both retarded groups. However, under the enriched condition the scores of the culturally deprived group were significantly better than their own scores under the standard condition; significantly better than the enriched condition scores of the organically retarded group; and there were no significant differences between the non-retarded and culturally deprived children on the conditions which employed five exemplars. From these results Gordon & Haywood conclude that children from culturally deprived backgrounds are not retarded in their ability to form verbal abstractions but suffer from an information input deficit. This deficit is modifiable by enriching the amount of information presented to such children. With regard to non-retarded subjects, the authors maintain since information deficit is not assumed one should not expect that such a group is likely to benefit from enrichment procedure. In the case of the organically retarded group, Gordon & Haywood explain their performance as follows: 1) either brain damage might have actually caused damage to the central abstracting processes; or 2) the enrichment procedure was not sufficient to surmount the deficit that might have been present in these children. Other workers at the Nashville Laboratory (e.g. Call, 1973; Foster, 1970; Tyrnchuk, 1973) have also tested several times the input-deficit hypothesis of Gordon & Haywood involving different samples and found results in line with the earlier study of Gordon & Haywood (1969).

The aforementioned studies provide empirical evidence to the Haywood et al. (1975) hypothesis of input deficit amongst culturally deprived children and their claim from these studies that by enriching the supply of information these children can be helped to perform verbal abstracting tasks at a normal level. However, these authors are careful in their claim in that they do not infer that given enriched input, mildly culturally deprived children would be able to perform as well as their normal peers in any domain. These findings nevertheless have an important bearing on actual classroom teaching practice. For instance, in the light of the input-deficit model, teachers could try enriching their traditional way of presenting teaching materials to culturally deprived or different children, to see, if by so doing, their level of performance in basic subjects can be brought up to par with indigenous children. Although the authors of the input deficiency model present an intriguing hypothesis and support it with well designed empirical evidence, it is difficult to imagine that this alone could be at the heart of the differences that one finds in the performance of ethnic minority children and children from the dominant culture.

2.6.2 Budoff and his Associates

At the Cambridge Laboratory (Massachusetts, America) Budoff and his colleagues have been concerned with determining the learning potential of children who have been diagnosed as educable mentally retarded (EMR) by standardised tests. They found that by assessing these children's learning potential they could identify children who were essentially and intrinsically EMR; whilst others, although identified as EMR by standardised tests, were functioning at that level due to verbal deficiencies (Budoff, 1967, 1969). In order to obtain an index of these children's learning potential Budoff and his associates have adapted some of the existing non-verbal tests: e.g. Koh's Blocks, Feuerstein's Learning Potential Assessment Service and the Raven's Matrices (see Johnson, 1976).

Budoff has carried out several studies (e.g. 1967, 1969) in order to test his hypothesis concerning verbal deficiences in educable mentally retarded children. In one of his earlier studies Budoff (1967) presented Koh's Blocks three times to a group of educable mentally retarded (EMR) adolescents. During the first administration the child's base level of functioning was obtained. This was followed by a period of training with a view to coaching the child in such activities as the analysis of each design into simple elements, the concept of two colour blocks as the components of more complex designs and a systematic comparison with the standard design. After completion of the training period, the children were retested: first a day after the training period and then a month after the training period. The pattern of results that emerged as a result of presenting Koh's Blocks this way, was that there was a small group of children amongst this sample who achieved fairly high scores during the first testing and were able to solve difficult problems; this group showed little evidence of benefiting from the coaching. Budoff called them high scorers. There were some children amongst the EMR group who obtained quite low scores during first testing but made substantial gains as a result of the coaching. They were described as gainers. There emerged yet another group: this group's performance was quite poor on all three occasions (non - gainers). From these findings and on the basis of another study (Budoff, 1969) where he followed the same procedure as in the earlier study (except that in this second study Budoff used the Raven's Matrices, Wechsler's Performance Test and concept shift tasks) Budoff found that, in the main, the performance of gainers was superior to non-gainers both in speed and efficiency of learning. From these two studies Budoff advanced the idea that there are two categories amongst EMR children: some who may be considered as "truly intrinsically mentally retarded" (Budoff, 1969, p. 286), whilst others are not essentially mentally retarded in this sense but their retardation could be associated due to verbal deficiencies.

Furthermore, individuals who demonstrate greater gains on their assessment devices are significantly more competent compared to non-gainers in all aspects of social functioning.

Budoff (1969) maintains that the results obtained as a result of assessing the child's potential has a useful implication for the practitioner. Once the 'gainers' have been identified on the basis of their learning potential score then this information can be used in designing curriculum appropriate for the need of these children. The provision of the appropriate curriculum for these 'gainers' would enable them to benefit from schooling at least as much as their low achieving peers in the mainstream of education. The assessment of learning potential would also help to identify able children who may be likely to be at risk.

2.6.3 Feuerstein

Feuerstein's work developed for more or less the same reasons as the present research: to resolve the problem of assessing "immigrants" with a diverse cultural and educational background. The use of conventional IQ and attainment tests on these adolescent immigrants showed, in the main, the same type of results as have been observed in Britain and the United States, i.e. their performance on scholastic-achievement tests was very low and they also demonstrated a developmental lag ranging from nearly four to six years in several cognitive areas (Feuerstein, 1980). Disenchanted with conventional assessment procedures, Feuerstein developed his own theory and an assessment device (Learning Potential Assessment Device, Feuerstein, 1968) rooted in "a theory of the nature and development of intelligent functioning" (Haywood et al., 1975). One of the key features of this assessment device is that assessment is not carried out in a void. It is linked with intervention and the remediation of areas found deficient as a result of assessment. The Learning Potential Assessment Device (LPAD) is based on the test-teach-test model (cf. the model used by Hegarty & Lucas, 1978). It is designed to provide information about the general learning "modifiability" and the extent of "modifiability". By "modifiability" Feuerstein means the alteration of cognitive structures in a more or less permanent way. In addition the LPAD also provides information about the teaching effort an individual would need in order to effect any changes, his ability to generalise and his preferred cognitive modality for learning and responding (for further details of this instrument see Feuerstein, 1980; Johnson, 1976).

In addition to the concept of modifiability, "mediated learning" is also central to the understanding of Feuerstein's theory of assessment of learning potential. Feuerstein (1970) defines "mediated learning" as:-

> "...the interactional process between the developing human organism and an experienced adult who, by interposing himself between the child and external sources of stimulation, mediates the world to him by framing, selecting, focusing and feeding back environmental experience in such a way as to create appropriate learning sets." (Pages 358 - 359)

Thus the role of an adult is crucial in a child's cognitive development. Feuerstein distinguishes mediated learning from 'direct exposure learning' in that, in the latter case learning takes place as a result of chance encounters between the organism and the environment, a process similar to assimilation and accommodation (see Piaget & Inhelder, 1969). Although non-mediated learning does help in the modification of cognitive structures, this form of learning is not as effective and efficient as mediated learning. However, direct exposure learning and mediated learning are the two fundamental modes of learning and the processes can occur simultaneously. These two types of learning, as distinguished by Feuerstein, not only differ in their degree of importance but their importance may also differ at different levels of developmental stages. Lack of sufficient exposure to mediated learning often results in cultural deprivation and arrests cognitive growth. Feuerstein (1979) has postulated twenty-eight different skill areas which in his opinion may be absent or insufficiently developed in those who have not had the adequate benefit of mediated learning. (For details of these skill areas see Feuerstein et al., 1972.) These deficits can be modified and the appropriate learning sets can be developed.

Feuerstein's approach, which embraces both assessment and remediation has:-

> "...improved the functioning of thousands of immigrants who arrived in Israel functioning academically at a retarded level." (Haywood et al., 1975, p. 114; see also Feuerstein, 1980; Hegarty & Lucas, 1978)

Essentially, what this review of the literature so far - and the Introduction - have unequivocally demonstrated is that assessment practices which are currently in vogue for testing children from ethnic minority groups, are less than satisfactory and should not be employed. There is a danger that continued reliance on these practices can result in these children's over-representation in special schools which is socially and ethically indefensible. It has also been highlighted that it is erroneous to use IQ tests as measures of learning potential. In order to measure learning potential instruments designed to measure this process are needed. There has been some modest effort in this direction in Britain, but more research in required. The present study would appear well placed to make some contribution towards fulfilling that need; since one of its main aims is to develop a test battery which would be able to assess learning efficiency and cognitive functioning of Asian (and other) children.

2.7 Jensen's Theory Concerning Population Differences

Jensen's (1961) earlier views about IQ tests and their inappropriateness for children from ethnic minority groups were not too dissimilar from some of the views that have been described above. At that time, in common with those views, Jensen too believed:-

> "...that nearly all standard IQ tests were grossly biased against virtually everyone but the white middle class."
> (Jensen, 1980 b, p.xi)

However, now, Jensen (1980 b) regards those views as "a bit soft headed". His present views about population differences - possibly as a result of becoming more "dispassionate" and "more incisively critical" - appear to be embodied in his two level theory of mental abilities (see 1969; 1973; 1974a;Jensen & Inouye, 1980) which shows a marked shift from his earlier thinking. Jensen's theory of mental abilities, like some of the earlier theories, also seeks to explain why the performance of certain ethnic groups tends to be inferior in relation to the majority of the "white" American population on certain types of mental processes (which are typically measured by the commonly used IQ and attainment tests).

Some key features of Jensen's two level theory of mental abilities have already been described in the previous chapter. In summary, it was said that Jensen postulates two types of mental abilities, Level I and Level II. Level I ability requires little mental manipulation of the input prior to its output and is best measured by tasks such as digit span, serial rote learning, paired associate and so on. Unlike Level I ability, Level II ability involves mental manipulation of sensory inputs. It is akin to Spearman's g and can be measured by such tests as the Raven's Matrices, tests of general and fluid intelligence, Piagetian tests and so on.

Level I and Level II abilities, as well as having a distinct genetic basis, according to Jensen, also have a hierarchical relationship. Level II ability has some degree of functional dependence on Level I ability, but the reverse is not true, i.e. for the development and operation of the former ability, the latter is necessary but not sufficient. Because of the hierarchical nature of Level I and Level II, it would be extremely rare to find anybody with a very high level of level II ability, but a very low level of Level I ability, the obverse of this, however, should not be uncommon, i.e. individuals will have high Level I ability, but poor Level II ability. Jensen (1973) postulates a correlation between Level I and Level II abilities in any population which he accounts:-

> "...more a result of there existing a genetic correlation between the abilities which have come about through selection and assortive mating... Population groups that have developed under different selective pressures for different abilities and through historic, geographic and relative social isolation from one another, might therefore

> be expected to differ in Level I and Level II abilities and to show differences in the degree of correlation between Levels I and II." (Jensen Op. Cit. , p. 264)

Yet another key feature of Jensen's theory is that different ethnic groups differ far less - or hardly at all - on Level I ability compared to Level II ability. On Level II ability, amongst various ethnic groups, marked differences (one standard deviation or more) have been noted (Jensen & Inouye, 1980).

2.7.1 Empirical Studies Concerning Jensen's Level I & Level II Theory

In the last decade or so Jensen has carried out several studies involving different populations to test his theory of mental abilities. For instance, in a study conducted in 1973 Jensen was concerned to determine whether the relationship between Level I and Level II abilities and race and socio-economic status (SES) would essentially be the same in an agricultural community in the central valley of California as he had found in his other studies which involve mainly highly urbanised white and Negro children in the greater San Francisco Bay area. Some 2,000 white, Negro and Mexican children from grades 4, 5 and 6 took part in this study. The majority of the white children represented middle and low middle-class; the Negro children came from lower middle; and the Mexican-Amercians were mostly socio-economically disadvantaged with the exception of a small proportion from middle and upper middle-class. Tests administered included tests of intelligence (Lorge-Thorndike Intelligence Test: verbal and non-verbal; Raven's Coloured Progressive Matrices), scholastic achievement and a short term memory test. The raw scores from all the tests were factor analysed and each grade was treated separately, i.e. as if they were three independent replications of the study. Three factors emerged in the case of each ethnic group and for each grade. Jensen described the first two factors as "types of Level II ability": they correspond to Cattel's (1971) crystalised and fluid intelligence which are two aspects of the general factor, g. The third factor, on which the tests like Memory-immediate, Memory-delayed and Memory-repeated loaded significantly as a measure of Level I ability. The results showed significant ethnic differences (more so in the case of white-Negro) on Level II ability than on Level I. On Level I ability there was a greater difference between the white and Mexican groups than between the Negro and white groups. The difference between the Mexican and white groups was greater on crystalised intelligence than on fluid intelligence and memory. Between the Mexican and Negro groups the differences on crystalised intelligence and Memory mean factor scores were marginal compared to their differences on fluid intelligence, the differences being in favour of the Mexicans. The correlation between Level I and Level II abilities was higher for white children than for Negro children; the difference however was not "impressive". In the case of Negro and Mexican children, the correlations between the measures of SES and the three ability factors (crystalised intelligence, fluid intelligence and memory) were low, whilst the correlations between Level I and Level II and SES for the white populations were different.

Jensen claims that the above results are in accord with his theory. As the results of the Mexican group are not quite consistent with his theory, Jensen therefore advances the view that his hypothesis:-

> "...applies more to the white-Negro racial difference rather than to their SES difference per se."
> (Jensen, 1973, p. 269)

In another study, Jensen (1974a) was concerned to test his theory of mental abilities in a total school population. The major difference between his previous studies and this study is that in earlier studies Jensen used specially selected samples only from the population: in the present study the entire school population was involved.

There were three hypotheses that this study set out to test: (i) Level II ability has a positive correlation with socio-economic status (SES) while Level I ability has not; stated differently, there will be no differences between the different SES children on Level I ability but there will be significant differences on Level II ability. (ii) The regression of Level I on Level II is not as great in a low SES population compared to upper and middle SES children. The inference from this is that one would expect a lower correlation between Level I and Level II abilities in low SES populations compared to upper and middle SES populations. (iii) The third hypothesis was concerned with the hierarchical relationship between (which is necessary but not sufficient) Level I and Level II abilities.

For testing Level I ability the subjects were administered the digit span memory test designed by Jensen himself; for assessing Level II ability, they were given the Lorge-Thorndike IQ Test. The results of this study failed to provide support to Jensen's first hypothesis, namely, that children from low and middle socio-economic groups differ in Level II ability but not in Level I. Despite 'highly significant' differences between the 'Black' and 'White' children on both Level I and Level II ability, Jensen seems somewhat reluctant to admit that his own research has failed to provide evidence in support of one of the key aspects of his theory. This is evident from Jensen's comment:-

> "We see that although the white-black difference is highly significant, both on the memory and on the intelligence tests, the difference on the intelligence tests is more than twice the difference on the memory test. It is thus unclear whether this finding disproves or supports the hypothesis." (Jensen, 1974a, p. 104)

The data was consistent with his second hypothesis. The third hypothesis received only partial support in that in the case of the non-verbal tests, only the hypothesis of hierarchical dependence of level II upon Level I was confirmed, but not in the case of the verbal test. Clearly, these results cannot be considered as providing full support to Jensen's hypothesis of mental abilities. From these findings it seems quite reasonable to predict that if a similar study was conducted it would seem unlikely that the results would completely be in line with Jensen's theory of mental abilities.

In yet another study Jensen (1975) was concerned to test his hypothesis of two levels on black and white children, aged 5 to 12 years. The sample consisted of 669 white and 662 black children randomly selected from 98 school districts in California. All the subjects were administered the Wechsler Intelligence Scale for Children (R) including the Forward and Backward Digit span subtests. Jensen claims that the results were in line with his predicitions. Jensen found that the Backward Digit span subtest, which he regards as a measure of Level II ability (because more mental manipulation is required than in the Forward Digit span subtest which is considered as a test of Level I ability, because little transformation of the input is required), was more highly correlated with IQ than the Forward Digit span subtest. Jensen also found, as predicted, that the differences in performance on Level I and Level II ability decreased with age. The other hypothesis that was confirmed was that there would be greater differences between whites and blacks in the Backward Digit span subtest compared to the Forward Digit span subtest. No significant interaction of socio-economic status (independent of race) both with Forward and Backward Digit span was found. Commenting on these findings Jensen says:-

> "Thus the theoretical prediction of a larger group difference in BDS than in FDS is substantiated for race, but not for SES. It is beginning to appear that the Level I - Level II theory may hold more strongly for race (i.e. white-black) than for SES." (Jensen, 1975, p. 888)

This shows that in recent years, Jensen's views with regards to the application of his theory to SES have been modified, because originally Jensen's hypotheses were cast in terms of socio-economic status and left unclear as:-

> "...to what degree the theory applied to SES and to race (i.e. white vs. black) independently of one another." (Jensen, 1975, p. 886)

More recently Jensen & Inouye (1980) tried to extend the hypothesis of two level mental abilities to Oriental* Americans. In this study, the authors compared Level I and Level II abilities of Oriental (Chinese and Japanese), Black and White Americans. The sample consisted of all the children from Grade 2 through to Grade 6 in a California school, who were available at the time of testing. Tests given included measures of intelligence, scholastic achievement and short-term memory. These tests have been used by Jensen in his previous studies as well where he had been concerned in determining Level I and Level II abilities amongst different ethnic groups. The data was factor analysed and for each grade, and within each ethnic group, two factors were identified which Jensen & Inouye regard as representing Level I and Level II abilities. As expected no significant differences were found between Whites and Orientals on Level II ability, but they differed on Level I (Whites greater than Orientals). Orientals and Blacks did not differ on Level I but on Level II (Orientals greater than Blacks). Whites and Blacks differ both on Level I and Level II abilities (Whites greater than Blacks). The difference between the Whites and Blacks on Level II ability was slightly more than four times the difference on Level I ability.

---

* Note that Jensen uses the term 'Asian' rather than 'Oriental' which has been substituted here to avoid confusion with the definition given on page xvi.

Jensen & Inouye also found that the overall mean loadings of the various achievement tests used was only 0.22 on Level I ability whilst 0.85 on Level II ability. These results, the authors feel, provide support to Jensen's hypotheses of mental abilities. As these findings are consistent at all grade levels, Jensen & Inouye consider that these results should be generalisable to other Oriental-Americans, Whites and Blacks. At this stage, the authors acknowledge their inability to explain the obtained Oriental-American and Blacks' and Whites' differences on Level I ability (American-Orientals less than Whites; American-Orientals greater than Blacks). These findings cannot really be considered as fully consistent with Jensen's formulations about Level I and Level II ability. It is therefore doubtful whether Jensen's theory is sufficiently robust and, if tried on other populations, would receive partial or full support.

Jensen's theory receives additional support from Longstreth (1978) who attempted to extend it to college students with the hope that it might explain the:-

> "...achievement differences between races in that age group."
> (Longstreth, 1978, p. 290)

Like Jensen (1969), Longstreth too had observed that with the disadvantaged college students there was little correspondence between their IQ and achievement scores and their performance "in certain situations". (Longstreth does not state under what circumstances this has been observed.) In order to examine Jensen's hypothesis for college students, Longstreth compared the performance of black (n = 20), Mexican American (n = 31), "white" American (n = 85) and Oriental (n = 42) on multiple choice exam, essay scores, true-false tests and on the Forward Digit Span test (a subtest from the WISC-R (Wechsler, 1974)). Longstreth found that black and Mexican-American students' scores were significantly lower compared to white and Oriental students on multiple choice exams, on essay, and on cognitive abilities test but there were no significant differences on the Forward Digit Span Test. On the true-false test the means for Oriental and white Americans were higher than the black students but not significantly different. Interestingly, there were also statistically significant differences in favour of white males as opposed to white females ($F(1, 85) = 6.9$, $p < .02$). These differences are explained not in terms of Jensen's theory (i.e. clearly it does not help to explain these differences) but in terms of "selection", motivational and cognitive factors. It could be argued that the other differences amongst the four racial groups that have been observed and interpreted as consistent with Jensen's thesis could have occurred due to those three factors as well. Further, the inference drawn from the correlations amongst the test forms that true-false type of testing is more akin to Level I ability than to Level II ability is not convincing. It would seem that when Longstreth's results are examined critically they do not quite support his conclusion that:-

> "These results successfully extend Jensen's theory to the college classroom." (Longstreth, 1978, p. 289)

In addition to Jensen's (and Longstreth's) studies other workers too have tested his hypotheses of mental abilities. Green & Rohwer (1971) examined Jensen's claim that Level I ability is equally distributed across various SES groups while Level II ability is not. In order to test this hypothesis Green & Rohwer randomly selected 60 fourth grade black children of low SES, middle SES and low - middle SES; all were administered a paired associate task, a digit span test and the Raven's Coloured Progressive Matrices. In addition information was also available with regard to their ability in reading, arithmetic, total reading score on the Stanford Achievement Test and the Lorge-Thorndike Intelligence Test. Of the three tasks administered (paired associate, digit span, Raven's Matrices) SES related differences emerged on the latter two, whilst on the paired associate tasks the performance of all SES groups was equivalent (only "neglibible" amount of SES related variance was found). Inter-correlations among various variables showed that paired associate tasks may be tapping an ability which is quite different from other measures. The results obtained for paired associate and the Raven's Matrices provide support to Jensen's model (1969) provided one accepts that paired associate tasks measure Level I abilities (cf. Rohwer et al., 1971). Peformance on the digit span tasks was in direct contradiction to the model offered by Jensen. The authors conclude that Jensen's model may hold for SES differences in white populations but not in black populations; or it is also possible that Jensen's model is not satisfactory.

Like Green & Rohwer, Rohwer et al. (1971) who tested the hypothesis that the low achievement of low SES black children was on account of their corresponding defficiencies in their ability to learn (cf. Kee & Rohwer, 1973), came to the conclusion that Jensen's hypothesis did not offer a "satisfactory explanation of population differences in school achievement" (p. 12). In this study Rohwer et al. (1971) also argue against Jensen's thesis that paired associate tasks are measures of Level I and not Level II ability and in support of their contention they report several studies which have demonstrated that these tasks involve a considerable amount of manipulation of sensory input. The evidence advanced by Stevenson (1972) would also suggest that Level I and Level II tasks are not as independent of each other as Jensen argues. Stevenson found correlations ranging from .38 to .56 between measures of Level I and Level II abilities.

Clarke et al. (1967) and Clarke et al. (1969) have carried out a series of studies involving pre-school children who were mainly from low social class. These children were asked to perform sorting tasks and in order to be able to do them correctly they had to employ super ordinate concepts. Their initial attempts (i.e. prior to any training) showed that these children used no systematic strategy. However, after these children had been trained on different materials where they had to use a conceptual strategy, they were retested on the initial tasks. The results showed that there was a marked improvement in these children's later performance (i.e. after training). From these findings the authors suggest that the biologically normal child is essentially capable of using Level I and Level II strategies; however of the two approaches the child is likely to use one which is habitually preferred in his culture. This conclusion is in line with much of the cross-cultural studies which have been referred to in the Introduction and in this Chapter. The author's personal opinion based on having experienced two cultures, one for nearly 26 years and the other for nearly 17 years, is in accord with Clarke <u>et al</u>. (1967) and Clarke <u>et al</u>. (1969). However, there are some other studies which did not set out to test Jensen's hypotheses as such (their main concern was to examine the relationship between learning and intelligence), but which nonetheless throw some light on them. Consider the work of Woodrow (e.g. 1938 a, b, 1939 a, b, c,); although it was done prior to Jensen propounding his theory, it may be considered as inconsistent with Jensen's theory. Contrary to what would be expected from Jensen's theory, Woodrow found that the learning tasks he used did not load on one factor only, although they were factorially distinct from intelligence tests. Mackay & Vernon (1963) too, found that:-

> "...learning measures tend to group under a number of rather ill-defined factors, according to the type of material and the conditions of learning and recall. Some of these group factors do seem to relate to ability factors, for example, verbal learning correlating more highly with verbal ability; non verbal learning with spatial ability etc." (p. 179)

Stake (1961) also came to the conclusion that the various tasks designed to measure learning ability did not form a general factor like Spearman's g (or in other words did not load on factor which according to Jensen's theory could be described as a measure of Level I ability) but tended to be specific to a particular type of task (cf. Mackay & Vernon, 1963). For instance, Stake used 12 learning tasks on 240 children and when these tasks were factor analysed there emerged four factors: two were memory task factors; one numerical task learning factor and a concentration factor. Duncanson (1964) and Malmi et al. (1979 and a study cited therein by Underwood et al.,1978) too found that learning tasks do not load on one factor. Guilford (1967) who, after a review of the literature, came to the conclusion that there are several types of learning ability and that these cannot be encompassed under the umbrella of one general learning ability which is common to all types of learning. Like Guilford's conclusion, Hegarty & Lucas (1978) postulate that there are several types of learning tasks and:-

> "...competence at one does not guarantee competence at another. The ability to master algebra does not entail the ability to comprehend historical processes or to play the piano." (p. 46)

Like Guilford (1967) and Hegarty & Lucas (1978) Cronbach (1970), after a survey of a considerable amount of literature pertaining to the relationship between learning and intelligence, came to the conclusion that learning is not unidimensional; some individuals could excel in one type of learning task but could be poor on others, and vice versa.

Thus, the majority of work that has been considered which has either direct or indirect bearing on the Levels theory, provide little support to it. As a matter of fact some inconsistencies are present even in Jensen's own findings as well. Despite all this, Jensen's views seem quite influential: in line with his thinking, many teachers in Britain appear to think that children from ethnic minority groups have low genetic potential notwithstanding that Jensen himself has not studied them. Since Jensen's theory has very serious economic, social and ethical implications, it is important that its generalisability is tested on populations other than those he has studied. Of all the ethnic groups, for the present study, only Asian children settled in Britain have been selected.

2.8 Hypotheses Related to Jensen's Theory of Mental Abilities

Consistent with Jensen's theory regarding Level I and Level II ability (Jensen, 1969; 1973; 1974 a; 1980) the present study would test the following hypotheses.

Hypothesis 1: Tests of associative type of learning and short term memory (i.e. Level I ability according to Jensen) would be orthogonal to IQ tests (e.g. Raven's Matrices and Draw-a-Man) and Piaget based tests which are also regarded by Jensen as measures of g (i.e. Level II ability according to Jensen). (But see page 53)

Hypothesis 2: There would be no real differences in the factorial structure of English and Asian children on the subtests of the proposed test battery and the two conventional IQ tests.

Hypothesis 3: There would be real differences between English and Asian children on tests of Level II ability as measured by the Raven's Matrices, Draw-a-Man Test and Piaget based tests.

Hypothesis 4: There would be no real differences between English and Asian children on tests of Level I ability as measured by the associative type of learning tests and a short term memory test.

Hypotheses 2, 3 and 4 would have a bearing on the "culture fairness" of the proposed test battery for English and Asian children (see Jensen 1974 a; 1980 a; 1980 b).

The second aim of the study was to contruct a test battery for Asian children. More specifically:

1. To design a battery of tests suitable for use with 6 to $9\frac{1}{2}$ year old Asian children (hopefully it should be equally useful for other children), with the following main characteristics:-

(a) That it is capable of discriminating the bottom 5% of the population in order to enable psychologists to assign children, particularly those referred for mild and severe learning difficulties, to appropriate educational facilities, e.g. ESN-M, ESN-S, or for children with learning difficulties or for additional remedial help.

(b) That it is minimally dependent upon language.

(c) That it can be used for assessing the child's learning efficiency so that this information can be used for advising the teacher about the teaching effort the child is likely to require in order to grasp new concepts or skills.

(d) That it can provide information about the child's level of cognitive functioning within the Piagetian framework so that the information obtained can be used for the planning of curriculum which is appropriate to the child's level of cognitive development (the age range covered for this purpose would be approximately three and a half years to about eight years.

2. To study the proposed battery's predictive validity empirically by instituting a small longitudinal study with a view to comparing it with conventional assessment procedures.

3. To include a small number of children in the present study who are already in ESN-M schools so that their performance can be compared

with the main sample in order to provide evidence on the construct related validity of the proposed battery.

4. To compare the predictive validity of the proposed test battery with IQ tests (e.g. Raven's Matrices and Draw-a-Man Test) and the teacher's rating of children's academic ability and potential.

5. To compare the performance of indigenous, Asian and comparable Indian children (a small sample to be tested in India) on the proposed battery and IQ tests.

<u>Note continued from P.51</u>

However, although Jensen appears to state that Level I and Level II abilities should be orthogonal (i.e. uncorrelated) he also claims a hierarchical relationship between Levels I and II - (see page 40) which implies that the levels cannot be independent but must be correlated. The interpretation of Jensen taken here is thus open to question, but is consistent with that made by other reseachers e.g. Stevenson (1972).

CHAPTER 3

RATIONALE BEHIND THE PILOT VERSION OF THE LEARNING EFFICIENCY TEST BATTERY

3.1. Introduction

It was decided to include the following subtests in a pilot version of the Learning Efficiency Test Battery (LETB):-

A Piaget based

(i) Seriation A

(ii) Seriation B

(iii) Serial Correspondence

(iv) Ordinal Correspondence

B Based on association learning principles

(v) Paired Association

C Based on short term memory

(vi) Visual Sequential Short Term Memory Test (VSMT)

D Based on Haynes (1971), Hegarty (1977) and Hegarty & Lucas (1978)

(vii) Logic Blocks

(viii) Number Tests

The theoretical basis and rationale for selecting each subtest or group of subtests for inclusion in the pilot version will now be discussed in turn.

In the development of the LETB a number of principles were taken into consideration, which will be described shortly. The genesis of these principles is from personal experience of assessing both indigenous and ethnic minority children for several years employing a wide variety of psychological tests, first hand knowledge of the child rearing and raising practices in Asian homes, and also the insights gained by several experienced workers in the field of assessing children from non-Western cultures (Berry, 1966; Hudson, 1967; Irvine, 1966; Jahoda, 1980; Schwarz, 1961; Vernon, 1969; Winter, 1963; Wober, 1967).

3.2 Principles

1. The presentation method should be such that the administration of the LETB should require only a minimal use of the English language either on the part of the Examiner or the child. This is because sometimes one has to assess a child who has little or no knowledge of the language.

2. The battery should include a variety of tests in order to sample a diverse range of mental processes, similar to the ones required in a classroom learning situation. It is often recommended that when testing children from minority groups, one should attempt to assess a wide range of cognitive functioning rather than relying on a single test or an extremely small number of tests (Anastasi, 1961; Davis, 1971; Jensen, 1970). The sampling of cognitive abilities is especially important in order to discover the strengths and weaknesses of children of educationally disadvantaged backgrounds (Jensen, 1970).

3. As far as it is possible to determine, the Battery should not attempt to assess any functions or processes which might be dependent upon or confounded by the child's previous learning. On the other hand, the Battery should be able to provide some objective evidence about the child's ability to learn new tasks which, it is assumed, have not been encountered before, this information would be an indication of the child's learning efficiency. The learning tasks should be based on those learning principles which can account for a substantial amount of learning, possibly across various cultures.

4. The Battery should be able to yield some information about the child's level of mental processes embedded in some theory so as to permit interpretation of his performance within that model.

5. Qualitative and quantitive information obtained from the administration of the Learning Efficiency Test Battery should be able to assist in determining future educational needs, in predicting performance in basic subjects and in curriculum planning.

6. The LETB should be able to discriminate approximately the bottom 5% of the population. Therefore the Battery should not include tasks which assess higher order mental processes, e.g. rule learning. The assessment of such processes would be appropriate if the Battery were designed to identify "high flyers" of the population and not the bottom 5%.

7. The test items should be such that they least resemble tasks which children often encounter in the classroom. It was thought that if the test items were similar to tasks performed in the classroom this might adversely affect the performance of children with a history of failure.

8. The other factor that received due importance was that the Battery should not be heavily dependent on items whose presentation rests on visually complex two-dimensional material, because many Asian children during their early developmental years are not sufficiently exposed to such materials, as compared to their Western counterparts. Hebb has marshalled impressive evidence which clearly emphasises that "perception is affected by past experience" (Hebb, 1949, p. 111). In view of this Ferguson (1954) maintains that in a normal child "the limits of learning in many perceptual tasks may be reached at a fairly early age"(p. 137). The implication is not that lack of sufficient experience with two dimensional materials is peculiar to Asian children only; this would apply to some English children too, particularly those whose parents belong to the lower-socio-economic group and maybe indifferent to their child's proper development or do not know how

best to help him. Whatever the case may be, such children end up being at a disadvantage when they are required to pick out, abstract, identify, integrate or analyse information from a two dimensional visually complex material. (cf. Ghuman, 1975; Hilgard & Atkinson, 1967: Oleron, 1977; Price-Williams, 1969; Vernon, 1969; Wadsworth, 1978).

9. As far as possible objects, materials or toys should not be used which may be favoured in one culture.

10. Yet another factor that was taken into consideration was that the test items should not be too much dependent on colours. Many Asian families, as far as experience goes, place little emphasis in teaching colours to their young children. In the author's professional contact with many schools he has often heard from headteachers that many Asian children when they come to school do not know the names of colours.

11. No subtest in the battery should be timed. At the same time it should not take an excessive time to administer because a test battery which took an inordinate amount of time to give would not really be practical.

12. In the light of Haynes' (1971) experience, in the development of the proposed LETB every attempt will be made to ensure that its administration is based on a model which encourages the active participation of the child, in order to maintain his interest and motivation in a situation where verbal communication with a non-English speaking child is likely to be extremely limited.

13. Ease of administration, portability and the cost of materials were not overlooked either.

This obviously is a formidable and very constraining list of principles to be taken into consideration in designing a new test battery. Nevertheless, they were deemed important. As far as the author's knowledge of existing tests goes, there would appear hardly any tests which can stand up to these principles. It was assumed that if a Battery could be developed which did not infringe any of the above principles, it would be highly likely that the developed Battery would not discriminate unfavourably either the English or Asian child. By bearing the above principles in mind, it can be reasonably argued that any of the well-recognised factors which might be a crucial source of variance in the performance of Asian and English children on the available tests has been controlled.

3.3. Search for the Content for the LET Battery

In the search for suitable content for the LETB it was deemed essential that any items or subtests designed or adapted for the Battery, should not be incongruent with the principles which have been stated above, the key aims of the study (the details of which have been given on pages 52-53 and the test administration model which has been specially developed (see pages 86 to 88) to meet the needs of the present investigation. First of all many of the commonly used IQ tests were critically examined to assess their suitability for the Battery.

3.4 Evaluation of the IQ Tests for Adaptation

In the previous Chapters, references have been made to the fact that the existing IQ tests, for various reasons, are not appropriate measures for children from ethnic groups. Criticisms have been voiced against tests like the WISC-R in that they assess the products of previous learning and thus penalise children from different cultural backgrounds and experiences (Cole, 1975; Haywood et al., 1975; see also Sattler, 1974). Lesser et al.(1973) also criticize IQ tests on the grounds that they do not assess a wide range of mental functioning: and are too heavily saturated with verbal items, and thus cannot be regarded as fair to certain sections in our populations. Lesser et al. state:-

> "A score based on a test that is heavily loaded with one factor can tell little about the quality and quantity of the various talents that an individual has."
> (Lesser et al., 1973, p. 281)

At a more general level IQ tests have also come under attack on the grounds that information about the child's IQ is hardly an aid in designing a curriculum appropriate to the child's developmental level. This limitation of IQ tests has received considerable attention these days as the demands from test results have changed. Tests must now be useful in planning educational programmes rather than just being tools of prediction or identifying which children needed special education and which would benefit from receiving education in the mainstream of education. Despite the fact that one needs tests to perform different functions, it seems that they still:-

> "...continue the tradition of identifying particular symptoms of learning disability or inexperience and then using this information for classification and differential assignment. There is little empirically derived and conceptually understood relationship between test score information and specific instructional activity."
> (Glaser, 1981, pages 924-25)

Glaser does not make it clear in his article whether he recommends that future tests should abandon the role of prediction and should mainly serve as agents of curriculum planning. It seems that both are important functions and tests are needed which can fulfil both these roles and therefore meet the changing demands.

Carroll & Horn (1981) (see also McReynolds, 1982) among several other limitations of IQ tests also point out that:-

> "The IQ for one individual can be based on a combination of high and low scores for component abilities that is different from the combination for the same IQ for another individual, and there is insufficient evidence that such combinations have the same implications in particular applications." (p. 1017)

If the available tests are not suitable, can they not be adapted to meet the assessment needs of ethnic minority children? Later in this Section, references will be made to some evidence which suggests that items or subtests from existing IQ tests cannot satisfactorily be adapted to meet the assessment and other needs of these children. This will be followed by a discussion of the reasons for turning to Piagetian tasks for inclusion in the LET Battery.

Haynes (1971) prior to devising her own test battery of The Learning Ability Tests, tried to adapt some of the existing IQ tests (e.g. NFER Non-Verbal Test, the Moray House Picture Test, Raven's Matrices, items from the Nebraska and Snijders-Ooman tests and so on), but abandoned the idea for several reasons which included: difficulty in conveying to the child the requirements of the test; some children were either unable to write or draw their responses; children found some of the test items boring where they had passively to observe demonstrations and so on.

Jensen (1961) also indicated that the translation of an IQ test into the language of a minority group does not necessarily improve the general performance of that group on that test. For instance, Jensen reports several studies which found that the Mexican-American's average score on the Stanford Binet tended to cluster around 80. If bilingual Mexican-Americans were administered the Stanford Binet translated into Spanish, it still did not make significant differences in their performance.

Moreover, if most IQ tests are examined in the light of the principles and aims of the study it would seem that the majority of IQ tests tend to violate them. For illustration, take one of the most commonly used IQ tests, the WISC-R (1974). Both the Verbal and Performance Scales are totally dependent upon language. All the Performance subtests are timed: they require complex manipulation of the stimuli presented visually. Some of these criticisms would be valid for the recently devised British Ability Scales (Elliot et al., 1977) as well.

As many of these shortcomings would apply to most IQ tests, and some to "culture free" or "culture fair" tests (for details see Anastasi, 1976; Jensen, 1980 b; Thorndike & Hagen, 1969) and to the learning ability tests (described by Hegarty & Lucas, 1978), it was, therefore, decided that for the development of the proposed LET Battery the author should look further afield rather than trying to model it on some of the existing tests. By so doing, it was assumed, the LET Battery would not suffer from the weaknesses often found in traditional IQ and culture fair or free tests when using them on children from ethnic minority groups. Furthermore, it would not be at logger-heads with any of the key aims, principles and the model for administration of the LET Battery.

3.5 Piaget Based Tests

3.5.1 Critical Examination of Piagetian Cognitive Tasks

For nearly two decades now many notable writers have employed Piaget's cognitive tasks for designing standardised tests. It was therefore decided to review some of his experiments also, for possible inclusion in the LET Battery especially as there seemed little evidence which suggested that, like IQ tests, they had been tried on Asian children and found unsuitable.

Although the literature on comparative studies of Asian and European children's performance on Piagetian tasks is rather sparse, Piagetian theory has inspired a vast amount of cross-cultural research (Berry & Dasen, 1974; Dasen & Heron, 1981 who in their chapter on "Cross-Cultural Tests of Piaget's Theory" refer to nearly 20 reviews and books of readings concerned with the cross-cultural aspects of Piaget's theory; Modgil & Modgil, 1976, Vol. 8). A reasonably non-controversial conclusion from the literature on Piaget based cross-cultural research would seem to be that children from societies quite different from highly technological Western societies go through the same sequence of stages of intellectual development as postulated by Piaget. There is, however, less unanimity in opinion with regards to the ages at which children from cultures other than Western attain these stages (Bruner et al., 1966; Vernon, 1969; Wadsworth, 1978 see also Dasen & Heron, 1981). Piaget (1971) would seem to be aware of this as is evident from his following statement:-

> "In some social environments the stages are accelerated, whereas in others they are more or less systematically retarded. This differential development shows that stages are not purely a question of the maturation of the nervous system but are dependent upon interaction with the social environment and with experience in general. The order, however, remains constant." (p. 7)

According to current literature, although Piaget's tasks cannot be considered as completely culture-free, they are less sensitive to cultural differences than traditional tests. In this connection Wadsworth observes Piaget's cognitive tasks:-

> "...are less affected by cultural differences than others... Comparisons of children in schools is less likely to be influenced by cultural or class differences using Piagetian measures than by using standardised tests." (Wadsworth, 1978 p. 230)

Yet another advantage of Piagetian tasks is that performance on them is not influenced either by the socio-economic status or amount of schooling so long as the tasks are not too dependent upon the quality of verbal responding (Wadsworth, 1978)

Flavell (1963) in the early sixties commenting upon the potential of using Piaget's experiments for using test purposes wrote:-

> "Piaget himself has never bent his talents towards making standardised tests out of the innumerable cognitive tasks he has created during his professional lifetime. But such an endeavour would surely be a logical extension of his work, and there is some activity in this direction, both in Geneva and elsewhere." (p. 361)

Flavell then describes the test-construction effort of several research workers, including Vinh-Bang, Barbel Inhelder, Pinard, Woodward, Lovell & Slater, Lunzer, Bibace & Reiff and Scheerer, and reviews the bulk of the literature concerning Piaget's theory and test development between 1950 and 1960. Flavell (1963) concludes the section on the "test development" by saying:-

> "...positive findings here make one optimistic about the test construction potentialities of his [Piaget's] research." (p. 364)

More recent literature concerned with the standardisation of Piaget based tests have been reviewed by Modgil & Modgil (1976,Vol 4,see also Ceri, 1977) which lends enough support to the fact that standardised tests could be developed from Piaget's theory of cognitive development. Piaget's theory of mental development is essentially based on the basic biological processes of accommodation and assimilation which can be witnessed throughout animal evolution and it seeks to explain the child's understanding of the world around him. Piaget has argued that all children go through a set of qualitatively different but invariant sequence of stages, the attainment of one stage being dependent upon the completion of the prior stages. The main stages of mental development advanced by Piaget are: sensorimotor, pre-operational and ego-centric, concrete and finally formal operations. Piaget's biological:-

> "...interpretation, in terms of action schemes, assimilation and accommodation, is radically different from theories which account for development by associative or conditioning mechanisms. The latter are essentially based on links between events and objects, imposed on the subject by his external environment, whereas Piaget regards the subject

as interacting with his environment, integrating new elements into already existing structures and gradually elaborating new structures." (Inhelder et al., 1974, p. 3)

This clearly is not as detailed a sketch of Piaget's theory as it deserves, but perhaps detailed enough for the purposes of the present enquiry. Piaget's theory since the late fifties has inspired a tremendous amount of research activity, and in recent years excellent reviews of these researches have become available. (See Dasen & Heron, 1981 who refer to numerous reviews; Lunzer, 1973; Modgil, 1974; Modgil & Modgil, 1976, Vol. I to VIII, 1980.)

A review of some of the representative literature supported the idea that many of the Piaget based tests had the potential to predict the child's future academic success in the basic subjects; were well suited for clinical work; shared important features with criterion referenced testing; could assist in determining future educational needs; could aid in curriculum planning (see Elliot, 1981 who refers to several projects where Piaget's theory has been the basis for curriculum design; Lunzer, 1973; Lunzer & Dolan, 1977; for a comprehensive review of the literature see Modgil, 1974; Modgil & Modgil, 1976, Vol 4; Modgil & Modgil, 1980; Schwebel & Raph, 1974). According to Hathaway & Hathaway - Theunissen (1974), Piagetian tasks assess areas of mental functioning which are of great importance to educators and clinicians - and the children they serve - but more importantly, they are areas which are not satisfactorily tested by prevalent psychometric procedures. The foregoing discussion would clearly suggest that the potential of Piaget based tests appeared in line with some of the key aims of the present research, particularly those concerned with the development of the LET Battery.

3.5.2 Need to Adapt Piagetian Tasks

However, a large majority of the tests as they have been developed or used in their author's researches could not be incorporated in the LET Battery without any adaptations. The reason for this is that most of them are dependent on the use of language, both on the part of the child and the examiner. If any Piagetian cognitive tasks selected for inclusion in the LETB had relied heavily upon the use of language, this would then have defeated one of the crucial aims of the proposed LET Battery; i.e. the Battery should be only minimally dependent upon language so that it can be used for children with little or no familiarity with the English language.

Yet another reason for departing from the Piagetian clinical interview technique which relies so heavily on language was that some studies have demonstrated that when a child fails a Piagetian task, it is difficult to decide whether the child's failure to perform the task is due to the relative contributions of cognitive or linguistic deficiences (Clarke, 1973 a, b; Donaldson & Wales, 1970; Lawson et al., 1974; Townsend, 1974).

The main thrust of the argument in these studies is that children between the ages of three to six years experience difficulty in understanding such relational concepts as more, longer, less, same, big, bigger, small and little. Furthermore, these children as well as betraying poor comprehension of the relational terminology, also show confusion about words which are related to length (e.g. longer) and number (e.g. more), (Donaldson & Balfour, 1968; Lawson et al., 1974). Siegel (1977) found that young children, 3 - 4 years old, were capable of finding solutions to a class inclusion problem when the language employed by these authors was less confusing compared to the type of language used in traditional tasks. For instance, when presented with an array of two kinds of candy, smarties and jelly beans (more smarties than jelly beans), instead of posing the question in the more traditional way, "are there more smarties or more jellies", these 3 - 4 year old children were asked whether they would rather eat the candy or the smarties. Siegal claimsthat it was as a result of removing the linguistic confusion from the way the questions were formulated that these young children were capable of solving the class inclusion problem; a cognitive task which according to Inhelder & Piaget (1964), the majority of children are unable to solve until they reach the age of around 7 years (cited in Siegel & Brainerd, 1978).

Thus these studies would seem to indicate that any observed differences in children's level of cognitive functioning may well be a reflection of differences in their linguistic ability. Commenting upon this issue, Siegel concludes:-

> "...this correlation between language use and problem solution makes it difficult to draw conclusions about cognitive processes when they are assessed by linguistic methods". (Siegel, 1978, pages 45 - 46)

As a result of the counfounding effect of language on Piagetian tasks, some researchers (e.g. Braine, 1959, 1962; Siegel 1978) have argued:-

> "...for a non-verbal approach, both on the basis of the assumptions of the Piagetian theory and on empirical demonstration of the lack of relationship between language and thought in the young child". (Siegel op.cit. p. 55)

Although the evidence is by no means fully consistent or clearly conclusive (Siegel, 1978; see also Elliot & Donaldson, 1981; Sinclair,. 1981), there is, however, sufficient evidence available which provides some justification for not depending upon language in the administration and in eliciting answers from the child. For instance, Miller (1976) in a recent review of the literature has concluded that performance on adapted Piagetian tasks which are only minimally depended upon language (i.e. non-verbal) is not significantly different from performance on traditional Piagetian tasks.

Some critics may be tempted to call the non-verbal Piagetian tasks merely "perceptual", i.e. successful completion of these tasks is mainly dependent upon perceptual cues and not on cognitive structures. Yet Gibson (1969) has argued, quite rightly, that there is an intimate relationship between perceptual processes and the ability which requires extracting information from the environment, perceiving distinctive features, concept formation and transfer, and that "higher-order structures" represent an extension, perhaps not clearly distinct, of the development of these perceptual processes.

Barnette & Adams in an unpublished study (no date, cited in Copeland, 1979) which involved 40 first grade children found that:-

> "The test of ordering the sticks might be criticised on the grounds that the child's failure could be caused by a visual perception problem. I thought the same thing at first because of the small difference between the sticks. In conducting the test, however, it was obvious that the children who were successful were using the operations as described by Piaget. They paid attention to both the stick before and stick after the one they were working with. The ones who failed only made sure that the stick was taller than the preceding one or shorter than the following stick. It was obvious that they were only concerned with one aspect of the operation. The problem was not visual. They could see a difference in height between the sticks. What they didn't do was to make sure that the stick was taller than the preceding one and, at the same time, shorter than the following one." (pages 93 - 94)

Also,non-verbal Piagetian tasks would seem quite appropriate to assess cognitive processes because Piaget himself has advanced the view that mental processes emerge and develop independently of language and the latter is not instrumental in the emergence of cognitive structures (see Sinclair, 1981). In addition to being consistent with the Piagetian theory, non-verbal tasks not only would circumvent the confounding effect of the language, but would also provide, perhaps, more objective information about the child's level of competence in that concept. Perhaps the over-riding point to bear in mind in presenting Piagetian tasks non-verbally is that the critical features of these tasks should be retained.

3.5.3 Validity of Piaget's Claims

In the previous section while discussing the effects of language in performing Piagetian tasks, there was some implication that possibly some of his tasks could be performed by children younger than Piaget had propounded if some of the parameters of the tasks were changed. Because this issue is quite central to Piaget's theory - and has relevance to the tasks selected for the present enquiry - this will be dealt with in the succeeding pages at some length. Some salient studies will be described which have attempted to challenge the validity of Piaget's findings followed by an evaluation of the whole issue.

During the last decade or so several researchers (e.g. Bryant, 1974; Bryant & Trabasso, 1971; Donaldson, 1978; Borke, Ennis, Cornell, Odom, Moore, Harris & Siegel cited in Siegel & Brainerd, 1978) have started to question the validity of some of Piaget's findings. For instance, Piaget et al. (1960) claim that five to seven year old children do not seem to understand the concept of transitivity, i.e. A>B and B>C, therefore A>C. Piaget and his associates (Piaget, Inhelder & Szeminska, 1960) have tried to provide support to their claim by what they describe as 'transitivity' experiments. These transitivity experiments involve either different weights or different sizes and are based on three separate stages. During stage one, the child observes the examiner comparing quantities A and B. During the second stage, usually B is then compared to C. In the final stage, the child is asked the relation between A and C, a comparison which he had not directly observed, and thus the examiner assesses the child's ability to draw a logical inference. Piaget et al. (1960; see also Youniss & Furth, 1973) have shown that it was after the age of 7 years that children were able to make deductive transitive inferences. From these findings Piaget concludes that first of all children have to understand the logical principle (this usually happens when a child is between 7 - 9 years old) followed by complete familiarity with this principle prior to allowing it to influence their perceptual judgements.

Bryant (1974) has challenged Piaget's position on transitivity (although he concedes that "...not withstanding this kind of disagreement, Piaget's idea of linking inferences and external frameworks as a vehicle for percpetual judgements is an important one..." p. 8) and claims that quite young children can make transitive inferences and are far more influenced by the external frameworks than the older children and adults.

Bryant & Trabasso (1971) carried out an experiment to determine whether four, five and six year old children were capable of making deductive transitive inferences. The experiment has two phases: training and assessment. At the time of training Bryant & Trabasso ensured that the children learnt or remembered the four initial direct comparisons: A>B, B>C, C>D, D>E. At the time of testing, the focus was to ensure whether the child remembered these initial comparisons and to determine if he could also combine them inferentially. The results showed that the four, five and six year old children were able to make transitive inferences, provided they could remember the information which they had to combine. The authors claim that they have obtained consistent results when these experiments have been replicated several times "...in many other versions of the same experimental design." (Bryant, 1974, p. 45).

In the above experiment there is a significant departure from the Piagetian way of presenting the task to the child: the original experiments did not include any training period. Bryant & Trabasso incorporated this training phase because they argued that the reason young children failed to make the deductive inference was not because of their lack of ability to do so; it could very well be due to the memory factor. In the words of Bryant (1974):-

> "The trouble with this assumption is that failures may well be caused by other factors than an inability to make inferences. An alternative possibility is that they could be due to lapses in memory. The transivity problem is a successive one, and this involves memory." (p. 42)

Other workers who have also tried to demonstrate that some Piagetian tasks can be performed by younger children than Piaget claims are, Borke, Ennis, Cornell, Odom, Moore, Harris & Siegel (cited in Siegel & Brainerd, 1978) and Donaldson, 1978. Is Piaget's theory seriously challeneged by studies which claim to demonstrate that many cognitive tasks are within the competence of much younger children than he claims?

In evaluating these studies, first of all, it should be remembered as Bryant (1982) points out, that there are as many studies which support Piaget's findings as contradict (e.g. for a comprehensive review see Modgil, 1974, 1980). More importantly if a child can perform a particular cognitive tasks at a different age level than Piaget's theory propounds, it is quite possible, that it may represent a different skill or different mental processes (cf. Anastasi, 1976; Winer , 1980). Along with this, it is also worth remembering, as Lunzer (1981) quite rightly points out, that these critical studies are contradictory to the Genevan findings only when the original conditions have been changed; but when the original experiments are replicated without introducing any changes, more often than not, they are in line with Piaget's findings. On this issue Sinclair's (1981) view is also quite instructive:-

> "Results of the kind discussed by Elliot & Donaldson (referring to their critical findings) may clarify certain points, bring to light aspects of child development not discussed by Piaget and his co-workers, and useful information (generally of a more psychological than epistemological nature), but they are not in contradiction with the theory." (p. 180)

It would therefore seem that although attempts have been made to challenge Piaget's theory it is still robust. Despite some evidence which contradicts some of Piaget's claims, a large body of evidence suggests that there was sufficient justification for adapting some of Piaget's experiments to include in the LET Battery. However, one problem still remained. Of the innumerable experiments which Piaget had devised, which experiments should be incorporated in the Battery? This issue is taken up in the next section.

3.5.4 Selection of Piagetian Tasks

As is well known Piaget has devised very ingeneous experiments many of which are quite suitable for the age range under consideration. Of the several tasks that Piaget has devised, it was decided to favour Seriation, Serial Correspondence and Ordinal Correspondence (Piaget, 1952; Inhelder & Piaget, 1964).

(i) Seriation: According to Inhelder & Piaget (1964), it is one of the fundamental developments in mental operations and consists of a chain of "asymmetrical", "connex" relations. The ability to seriate "operationally" does not develop before the age of 7 or 8 years. Essentially the task assesses the child's ability to order 10 elements according to increasing or decreasing size. The test material should be presented in such a way that the child must engage in active exploration of the material to enable him to find differences in length and to order the elements according to their size (Piaget, 1952; Inhelder & Piaget, 1964).

There are three main stages in the acquisition of seriation: ia) No attempt at seriation; ib) Small unco-ordinated series; ii) Success by trial and error; iii) Success with operational method (Piaget, 1952; Inhelder & Piaget, 1964).

ii) Serial and Ordinal Correspondence: Serial correspondence refers to double seriation. The requirement of the task is to seriate two sets of elements from shortest to longest in such a way that the shortest element of one set corresponds with the shortest element of the second set and so on. Piaget (1952; Inhelder & Piaget, 1964) maintains that the difficulty level of seriation and serial correspondence is more or less the same. Piaget found that children who could do seriation could also do serial correspondence.

In the development of this concept, children go through stages of understanding which are analogous to those for seriation. In other words between the ages of 4 - 5 years (stage 1) the serial ordering is haphazard. Around 5 - 6 years (stage 2) the arrangement is mainly by trial and error (no percpetual method). It is during stage 3 (around 7 years) that children employ the operational method to solve the problem. At this stage, rather than attempting to correspond two or three elements at a time (Stage 2) children can make the direct correspondence; i.e. they can select the biggest element from the one set of materials and can match with the biggest element of the other set and so on.

The chief difference between the serial and ordinal correspondence is that in the latter in order to demonstrate full understanding of this concept the child should be able to match the ordinal position of one element of one set with the similar ordinal position of the element in the second set of objects and so on. The objects in the first set are not arranged in their order of magnitude.

This would appear to be a slightly more difficult task than seriation and serial correspondence. The child passes through the same stages (little or no understanding, trial and error, and full understanding) as with seriation and serial correspondence, prior to demonstrating operational understanding of this concept. It would be quite pertinent to question why these tasks were favoured against experiments like conservation of weight or volume; or why, for that matter, these experiments were not included? By the inclusion of additional tasks it would have been possible to obtain more information about the child's level of cognitive functioning and since it is often recommended that when testing children from minority groups, one should attempt to assess a wide range of cognitive functioning rather than relying on a single test or an extremely small number of tests (Anastasi, 1961; Davis, 1971; Jensen, 1970).

On the one hand these are pertinent objections; on the other hand, if these objections were to be examined in the light of some of the principles that have been mentioned above it becomes clear that it was not possible to include these and some other tasks on the following grounds. For instance, it is difficult to imagine how the conservation of volume or weight can be adapted and presented to the child without using some language. Resorting to using language would have meant infringing one of the key aims of the Battery, namely, it should only be minimally dependent upon language. Also, if more tasks had been added to the Battery, this would also mean that it would take more time to test. (As it stands, it takes nearly an hour and a half to give the whole LET Battery.) As has already been noted - and personal experience confirms - practising psychologists tend to shy away from test batteries which take too much time to administer. Since the present project was conceived to meet a real need, making the Battery too time consuming to administer would have meant that it would have fewer chances of being used - thus it would have failed to meet one of the essential purposes of this study. Above all, many Piagetian tasks (e.g. Conservation of length, number, seriation, multiple seriation, linear and circular order and so on), do tend to correlate with each other and when factor analysed tend to have significant loadings on the same factor (Lunzer & Dolan, 1977).

Therefore it is reasonable to speculate that besides the reasons mentioned above, further addition of Piagetian based measures would not have been of any substantial value in providing extra information about the child's mental processes.

In summary, then, the reasons for the inclusion of Seriation and Serial Correspondence and Ordinal Correspondence only and not any other Piaget based cognitive tasks in the LET Battery, are that they are compatible with the aims of the study, do not infringe the principles, serve practical needs, and do not make the Battery too time consuming to administer.

3.6 Test Based on Associative Learning Principles

In addition to Piaget based tests, it was also decided to develop one learning test and a short term memory test. With the addition of these two types of assessment instruments, the expectation was that the LET Battery should be able to yield information about a reasonably wide range of mental functioning within the given amount of time (see Anastasi, 1961; Davis, 1971; Jensen, 1970). Above all, there are hardly any batteries available for this age range which provide information, particularly about the child's learning efficiency.

It has already been noted in the Introduction and elsewhere, that on account of current awareness about the limitations of IQ tests, especially in relation to their use on ethnic minority group children (Cole, 1975; Haywood et al.,1975; Hegarty, 1977; Lambert et al., 1974; McReynolds, 1982; McLoughton & Koh, 1982), several workers (Budoff, 1973; Haynes, 1971; Haywood et al., 1975; Hegarty & Lucas, 1978) have advanced the notion of the assessment of learning potential. By assessing the child's learning ability, as has been observed earlier, it should be possible to overcome some of the commonly associated criticisms of IQ.

besides being able to provide information about the child's learning efficiency - an important variable related to the child's academic attainments in school (Haynes, 1971; Hegarty & Lucas, 1978).

There is little need to go into any further details of justifying the need for assessing the child's learning efficiency because this has already been dealt with previously. However, it is apposite to reiterate here that in designing the learning tests, as in the case of Piaget based tests, due consideration was given to the fact that the tests developed should be consistent with some of the aims of the study, be congruent with the principles, the test administration model, and based in some well established theory.

With these factors in mind, a survey of the major learning theories (Hilgard & Bower, 1975; Hill, 1980) and of the basic forms of learning model propounded by Gagne (1977) was undertaken. As a consequence of this review, it was decided to develop a learning test rooted in the principles of associative learning rather than in any other learning model. The rationale behind this choice will become clear in the ensuing pages.

Associative learning is a fundamental kind of learning and it involves establishing new connections or associations between events in the environment (Hilgard <u>et al</u>., 1979). Historically speaking, the interest in associative learning is not of recent origin - in fact, this line of thinking dates back to the ancient Greeks, especially to Aristotle. Aristotle attempted to explain the psychological processes which were the basis of associative learning. They were: similarity, contrast and contiguity. Aristotle theorized how:-

> "One idea would be followed by another which was similar or contrasting, or which had been present together with it in one's past experience."
> (Woodworth & Schlosberg, 1954, p. 43)

Later on, these principles of association - similarity, contrast and contiguity - were adopted by the British School of Association Psychologists, and described as "Laws of Association", and the subject received considerable attention from them. The British Empiricists, then, used the "Laws" to explain the basis of all our knowledge. The knowledge is acquired through the principles of similarity, contrast and contiguity. It is ironical that the root of the bulk of present day "theory of conditioned responses" should be in the "arm chair psychology" which was not supported by any empirical evidence but developed as a result of:-

> "Careful scrutiny by each philosopher of his own experience." (Woodworth & Schlosberg, 1954, p. 49)

The credit of bestowing scientific credibility to "arm chair psychology" goes to such eminent workers as Galton, Wundt and Trautcholdt. They carried out the first experiments which were concerned with the role of associations in thinking and memory (Woodworth & Schlosberg, 1954).

Although the experimental work of Galton, Wundt and Trautcholdt is deemed original and significant, it suffered from two serious weaknesses. One, their experiments were confounded by the subjects' prior learnt associations. For instance, in free association experiments when the subject was presented with the stimulus word, he was free to give a response that he felt went with it, e.g. knife - fork; cup - saucer. The second confounding effect - that is, if we were to use the experimental model as it was originally used without any adaptation - would be the reliance on language. Invariably, both the stimulus and the response tended to be of a verbal nature. In earlier experiments, the subject might have been asked to give rapidly a string of words (response) that went with the stimulus word say, shoe or cake.

Ebbinghaus, however, overcame the first limitation (i.e. testing previously learnt assoications) by presenting subjects with unfamiliar materials to learn and then asking them to recall them. As in Ebbinghaus' experiments, in the Paired Association Test, the child has to learn new associations which, hopefully, he has not encountered before. As the child has to demonstrate an evidence of learning something new, this means the influence of his previous learning on the performance of this subtest has been minimised. Moreover, this subtest cannot be a subject of one of the often levied criticisms against IQ tests, that they test the product of previous learning (e.g. Haywood et al., 1975; Hegarty & Lucas, 1978; McReynolds, 1982). It would therefore be reasonable to expect no real differences in the performance of Asian, English and Indian children on this test.

In evaluating the child's ability to learn new associations the examiner's interest would reside in determining the number of trials the child is likely to take in order to master these associations. This should be a useful way of guaging the child's learning efficiency and thereby estimating the teaching input the child is likely to require. In turn this information could form a useful basis in helping an educational psychologist in resolving the future educational needs of a child as well as advising the teacher of the possible amount of teaching effort the child is likely to require in order to understand new concepts. This rationale is in line with current thinking, that the assessment should be linked with prescription (for the evaluation of currently popular assessment approaches see Eaves & McLaughlin, 1977).

Furthermore, the underlying principles on which the Paired Association Test is based is such that they are likely to be employed across various cultures. This claim is purely hypothetical and is not based on any empirical evidence because such evidence is hardly available. However, if one looks critically at the processes involved in learning by association one must conclude that this type of learning is unlikely to be peculiar to just one or two cultures alone.

The fact that this test should be based on a learning principle which has the semblance of being used and receiving almost equal emphasis across various cultures was very important. Theoretically, this would suggest that if it were possible to identify such a learning principle, embedding test items in such a principle should eliminate significant differences across cultures.

However, all the preceding claims about learning by association do not negate the fact that some psychologists do not subscribe to the view that all learning can be explained in terms of acquiring associations between stimulus and response. They stress that at least in some forms of learning understanding plays a crucial role. Even in simple S-R learning cognitive psychologists claim that the "reorganisation of one's perceptions, knowledge and ideas" take place (Hilgard et al., 1979, p. 592). They regard organisms as information processing systems: our minds reorganise the sensory input, "code it, store it in memory and is retrieved for later use" (Hilgard et al., 1979, p. 6).

Despite the controversy between cognitive and behaviourist psychologists, it seemed that the tasks rooted in simple associative type of learning would best meet at least two of the major aims of the present study, namely, that the Battery should be able to discriminate approximately the bottom 5% of the population and should be only minimally dependent upon language.

It is difficult to conceive how learning tasks assessing abilities such as problem solving, rule learning, concept learning, discriminations (see Gagne, 1977) can be presented without the use of language, besides which they are likely to discriminate a much higher proportion of children than this study is designed to do. Moreover, they will violate many of the principles (see Section 3:2) too.

3.7 Rationale Underlying the Visual Sequential Short-Term Memory Test

This test attempts to assess short term memory by presenting the child with visual stimuli. In order to be successful on this task, the child must pay attention to the visually presented stimuli, retain them for a short period of time, and then reproduce the stored information by sequencing the stimuli in the order in which it was presented to him or her.

Abstract and non meaningful stimuli were selected for the test items of this test as opposed to the use of pictures or numbers. This was done because the use of pictures or any meaningful stimuli carries the inherent danger that some children may be familiar, or more familiar than other children, with such stimuli. It is hoped that by employing non-meaningful stimuli the previous experiences of both Asian and English children have been equated. The other advantage of basing items on such stimuli is that they "counteracted the tendency to label the figures and therefore recall them through auditory and kinesthetic rehearsal" (Paraskevopoulos & Kirk, 1969, p. 47).

This test has been designed because short term memory would appear to be an important school related ability and would also seem to have relevance to the written English language. In the written English language "words...occur in horizontal orientation, in simultaneous presentation, and in close succession..." (Paraskevopoulos & Kirk, 1969). The method of presentation of the stimuli of this test would appear to correspond quite closely to this process. Furthermore, the importance of sequential memory can also be witnessed in such activities as in saying the days of the week, or in the act of counting. Likewise, the word order in a sentence is crucial too. A child who constructs his sentences "in the order of object-verb-subject 'milk want baby' may be exhibiting a disturbance in sequential memory" (Lerner, 1976, p. 186).

In summary, memory is at the heart of almost all kinds of learning (Lerner, 1976).

The reason that the visual modality was preferred to auditory modality was not due to the fact that the former is considered more important than the latter or vice versa. It is acknowledged that both modalities are important and perhaps have bearings on school related activities. Ideally tests for both modalities should have been designed. This would have meant an increase in the administration time of the Battery. As it has already been observed in the previous pages that if a test battery takes too much time to give, many psychologists, personal experience shows, tend not to use it. It seems illogical to devise a test battery knowing full well that it is unlikely to be used because it is too time consuming.

As well as the consideration of time, the other important factor was deciding for which modality a test could be designed which would be least dependent upon language in its administration. It seemed far harder to design a test using the auditory modality without depending upon language.

As the exposure is timed, it may therefore seem that in the design of this test one of the key factors (namely, the test should not be timed) has been violated. Strictly speaking that is not the case. Since the test is designed to measure short-term memory it was therefore essential to impose a time limit on the exposure of the stimuli. Not restricting the time would have meant defeating the main aim of the test. However, the recall period is not timed.

CHAPTER 4

PILOT STUDY

4.1 Introduction

The major aim of the pilot study was to try the proposed Test Battery (Learning Efficiency Test Battery) on a small sample in order to decide if it would need any modification prior to using it on a larger sample. The chief factors that were taken into consideration in designing this Battery and the theoretical model(s) on which these tests are based are discussed in Chapter 3 (see especially Section 3.2).

Studies subsequently referred to are already described in the review of literature.

4.2 The Procedure Adopted to Select Children From "Normal" Schools

The method adopted to draw Asian children from two normal schools was essentially similar to Haynes (1971). However, in order to meet the needs of the Pilot Study several modifications were also introduced. The method that was used in the selection of the sample is briefly outlined below.

Each headteacher of a normal school was requested to prepare a list of one of his second year junior classes. They were also asked to provide information about these children's ethnic origin, educational experience (i.e. whether Asian children have had a complete education in the U.K.) and in the majority of cases, father's occupation as well. In addition, each child was also rated according to his attainments and potential by his class teacher on an unambiguously defined five point scale. The five points represent:-

Grade	Description
1	Definitely bright; certainly well above average
2	Definitely above average
3	Average
4	Definitely below average; dull

5 Definitely well below average; in need of special school provision

It was assumed that this procedure would permit inclusion of children in the sample of various abilities which should correspond as closely to the normal distribution as possible. It seemed essential to include children of a wide range of abilities notwithstanding the fact that the LET Battery was largely designed for the bottom 5% of the population. This was necessary because the only way to find out whether the LETB would discriminate the bottom 5% from the rest was to try it on children of wide ranging abilities rather than on children of a limited ability only. Unless a vast majority of children of differing abilities reach the criterion, it is difficult to be sure about the LET Battery's intended discriminatory power. Hence the rationale for attempting to include children of almost all ranges of abilities except those who were severely retarded.

There were also some practical considerations. It was quite obvious that these considerations were of relatively less importance at the time of the Pilot Study; however it was anticipated that they would be at the time of the main study. Consider that in order to obtain children of a very narrow range of abilities a very large number of schools would be needed. It is almost a truism that it is not always convenient to enlist assistance from many willing schools for research. Further, if many more schools were to be involved they would be geographically much more scattered. In order to collect data from geographically scattered schools, not only would far more travelling time be required, but more importantly, it would not be possible to obtain a fairly homogeneous sample. This would be a considerable price to pay.

It is anticipated that by trying to meet one of the major aims of the study (i.e. the LETB should discriminate the bottom 5% of children) and by including children of almost all categories of ability the data may not yield a normal distribution of scores but a negatively skewed distribution. Regrettably this limitation will have to be accepted.

4.3 Subjects

With the foregoing considerations in mind, a total of 38 children, aged 8 : 10 to 10 : 4 years were randomly selected from three schools with whom the author had professional contacts and who were willing to provide the necessary facilities. Of the 38, 34 were Asian children (aged 8 :10 to 9 : 8 years) randomly drawn from two normal junior and infant schools with a high (nearly 80%) Asian population. The total sample also included four English children aged 8 : 6 to 10 : 4 years who came from a special school which mainly caters for the needs of ESN-M children. The ESN-M group was included to determine if the test battery discriminated between special school children and children in ordinary schools.

Table 4.1 Breakdown of the Sample According to Sex, Race and Type of School (n = 38)

	School 1 (Normal J and I)		School 2 (Normal J and I)		School 3 ESN (M)	
	Boys	Girls	Boys	Girls	Boys	Girls
English	0	0	0	0	3	1
Asian	3	6	12	13	0	0

The distribution of the Asian sample according to sex and the teacher's rating is summarised in Table 4.2. This Table does not include the four ESN-M children who fall into category 5 according to the assessment of pyschologists, the school medical officer and other professionals, rather than according to their teachers.

Table 4.2 Distribution of the Asian Sample According to Teacher's Rating and Sex (n = 34)

Teachers's Rating	Boys		Girls	
	n	%	n	%
1	3	20	5	26
2	2	13	4	21
3	5	33	5	26
4	4	27	1	5
5	1	7	4	21
Total	15	100	19	100

At this stage it did not seem crucial that this distribution should correspond closely to the normal distribution. What did seem important at this phase of the test development was that as many children in each category as possible should be tested because this way it would be possible to determine the suitability of the items for each category of children. The majority of the Asian children had a very homogeneous socio-cultural background. Most of the parents of these children worked in factories as labourers or as semi-skilled workers, and some worked on buses, either as drivers or conductors. None of the children in this sample came from a professional home.

4.4 Model for the Administration of the LETB

For the purposes of administering the Learning Efficiency Test Battery, a special model was developed. In the development of this model, the experience gained by several other workers who have carried out researches with non-Western populations played an important part. (Haynes, 1971; Hegarty & Lucas, 1978; Jahoda, 1956; Lloyd & Pidgeon, 1961; Ortar, 1960; Scott, 1950; Silvey, 1963; Vernon, 1972.)

The model on which the LETB is based is as follows: Demonstration, Demonstration and Practice, and Testing.

4.4.1 Demonstration

The chief function of this phase is to give the child a foretaste of the task. Here, the examiner performs the task himself, and the child is not expected to participate. The child's function is just to observe what the examiner does in front of him. However, the child is not discouraged if he offers to take part in the demonstration. This is an important part of the model as demonstration is considered a vital link in our understanding and learning (Williams, 1958 cited in Lunzer, 1973).

4.4.2 Demonstration and Practice

This phase is designed to incorporate teaching and practice of the same task. In addition, this intermediary stage, further ensures that the child really understands the task prior to coming to the next phase, Testing. Neither during the first (demonstration) nor the second stage are the child's responses scored. Although his way of responding or learning does provide qualitative information about his learning efficiency.

4.4.3 Testing

This is the final phase of the model. Here the testing proper is carried out but not in a fail-pass fashion. Even this phase involves teaching and provides feedback to the child whether his response is correct or incorrect. As a result of this teaching and feedback, he has an opportunity to demonstrate that by additional help, he is capable of learning new material. Here the child's responses are scored according to the teaching effort that he required in order to reach criterion.

This model enables the LETB to be administered requiring little use of language either on the part of the examiner or the child. Our experience has shown that this model ensures that the child really understands the requirements of the test, and it can be claimed with a reasonable amount of confidence, that if a child has not been able to do a task it is mainly because he cannot do it and is not being penalized for not fully grasping what he is required to do. In addition this model also provides a warming up period prior to the actual assessment and also helps to alleviate the child's anxiety (cf.Scott, 1950).

This model was adhered to even with those children who had satisfactory familiarity with the English language. However, this does not imply that the examiner was not permitted to talk to the child to build rapport with him and prevent long silences.

4.5 Materials

The subtests are described in detail in the following sections. For the four subtests (Seriation A and Seriation B, Serial Correspondence and Ordinal Correspondence) specially designed materials were used for the purposes of Demonstration and Demonstration and Practice phases of the administration; whilst for the purposes of Testing, the third phase of the model, additional materials were designed. For the other subtests, there were no separate materials for the different phases of the administration model. For all the subtests of the LETB three dimensional materials were used.

4.6. Method of Scoring

The scoring procedure adopted for the LETB was mainly right/wrong and it seemed satisfactory. To avoid undue repetition here, the details of the scoring method used for each subtest which was finally retained in the Test Battery are given in the Test Manual (see Appendix A).

4.7 Procedure

All the tests were given individually with the assistance of an experienced Senior Remedial Teacher, a Psychology Undergraduate from Cambridge and myself. Both the Remedial Teacher and the Student were fully trained in the administration procedure of the Learning Efficiency Test Battery. It took nearly two hours to test each child. The testing of each child was divided into two sessions of approximately one hour each. Due importance was given to the child's interest, motivation and co-operation at the time of testing. The testing was carried out in the following order:-

1. Seriation A
2. Seriation B
3. Serial Correspondence
4. Ordinal Correspondence
5. Classification of Logic Blocks according to colour, size and thickness
6. Number concepts
7. Paired Association Learning Test
8. Visual Sequential Short Term Memory
9. Raven's Matrices (Raven, 1962)
10. Reading Test (Schonell, 1951)

4.8 Tests Used, Preliminary Results, Amendments and Conclusions About Their Suitability

In the following paragraphs a description of each subtest will be presented together with the relevant preliminary results, any amendments that had to be made and conclusions reached with regards to their suitability.

i) Seriation A: For the purposes of assessment of this concept the child is required to seriate 10 Wooden Blocks (in natural wood) presented to him in a random order either without help ("spontaneously") or with help ("cues") and feedback. Irrespective of the strategy used, the correct solution is 10 blocks seriated starting from the smallest and progressively increasing to the largest. The length of these blocks range from 2.9 cms to 7.4 cms each differing in size by approximately .5 cms.

ii) Seriation B: This is virtually the same as Seriation A and the same testing materials are used as in Seriation A. The main difference, however, is that in order to produce a correct response the 10 Wooden Blocks need to be arranged in the descending order of their magnitude. In order to arrive at the correct response the child is free to use whatever strategy he wishes.

Table 4.3 outlines the number and percentage of Asian children classified by the level of understanding of the concept suggested by their performance on Seriation (A and B), in other words, their different levels of operational thinking according to the Piagetian model. This classification is based on the children's combined scores on Seriation A and B. Scores on Seriation A and B were combined in view of the apparent similarities between the two tests.

Table 4.3 Classification of Asian Children According to Their Performance on Seriation (A and B) n = 33

Raw Scores (Maximum 10)	Stage	Description	n	%
8 - 10	3	The child demonstrates full understanding of the concept	27	82
5 - 7	2	Trial and error - not full understanding	2	6
1 - 4	la	Limited understanding	3	9
0	1	Hardly any understanding	1	3

(Adapted from Inhelder & Piaget, 1964)

A significantly high proportion of the children (82%) reach the concrete operations level as measured by Seriation A and B, whilst only a small proportion (3%) demonstrate a complete lack of understanding of the concept. Between these two extremes are the other two categories: "trial and error" and "limited understanding". The "trial and error" category refers to those children whose grasp of the concept of Seriation may be slightly shaky; there are 6% in this category. In the "limited understanding" classification came those children whose performance would suggest that they do not quite understand the process involved in Seriation; of the total sample, there are 9% such children.

iii) Serial Correspondence: While the child is watching the examiner arranges 10 Wooden Blocks in front of the child. These are arranged in their serial order, starting with the smallest to the largest. The child also has in front of him 10 Wooden Blocks placed randomly. The requirement of this task is that the child has to place each Wooden Block underneath the ordinally matching Wooden Block either without help (spontaneously) or with help (cues) and feedback.

The wooden Blocks are the same as are used for Seriation A and B. The 10 Wooden Rods are in natural wood, round with a flattened edge; 2cms in diameter; length varying from 2.7 cms to 7.2 cms each, differing in size by approximately .5 cms.

iv) Ordinal Correspondence: The materials used for this test are the same as for Serial Correspondence. The major difference between this test and Serial Correspondence is that for this test the 10 Test Blocks are presented in a randomly pre-arranged fashion. The child has the 10 Wooden Rods randomly placed in a sort of pile in front of him - however the child should be able to see each element. The child has to match each Wooden Rod with its corresponding Wooden Block so that the ordinal position of the two matches with each other. He can do this either without help (spontaneously) or with help (cues) and feedback.

In Table 4.4 are shown the results for the same children, as above, on the Serial Correspondence and Ordinal Correspondence tasks. For the purpose of analysis, children's scores on both Serial and Ordinal Correspondence were combined for the same reason as for Seriation A and B.

Table 4.4 Classification of Asian Children According to Their Performance on Serial Correspondence and Ordinal Correspondence (n = 33)

Raw Scores (Maximum 10)	Stage	Description	n	%
8 - 10	3	The child demonstrates full understanding of the concept	27	82
5 - 7	2	Trial and error - not full understanding	4	12
1 - 4		Limited understanding	1	3
0	1	Hardly any understanding	1	3

It is interesting to compare the results shown in Table 4.3 with the results shown in Table 4.4. On both these concepts, the percentage of children who show full understanding of the concept is indentical. Likewise, the number of children who demonstrate little understanding of the concept of Ordinal Correspondence is identical with the number of children who fail to demonstrate any grasp of the concept of Seriation (i.e. 3%), although the former would appear to be a more difficult task than the latter.

These four tests rooted in the Piagetian theory and experiments seemed satisfactory with regards to their administration procedure and test materials. The distribution of scores and the face validity seemed satisfactory: in addition children enjoyed doing these tests. Therefore it was decided to retain these tests for the main study with no alteration.

v) Visual Sequential Short Term Memory Test: In certain ways this test resembles the Visual Sequential Memory Test from the Illinois Test of Psycholinguistic Abilities (Kirk et al., 1976) in the mode of presentation. In the Visual Sequential Short Term Memory Test the child is shown a predetermined sequence of abstract designs, each drawn on a 3.5 x 3.5 cms card for a set number of seconds. After the child has seen this sequence the cards are jumbled up and placed in front of him. The child has to reproduce this sequence.

A fixed number of trials is permitted and the items gradually become of increasing difficulty. For instance, the first item consists of 3 cards only and the last item consists of 6 cards.

Like the Piaget based tests this too appeared satisfactory and needed little alteration. Thus it was retained and tried on the main sample.

vi) Paired Association Test: In its original form the child was required to learn to associate six pictures each drawn on a card (approximately 1.5" x 1.5") with their appropriate lables. In selecting these lables, particular attention was paid to ensure that these words should not be familiar - at least not to the large majority of children of the age range selected for the research.

This was first achieved by ensuring that the words selected for this test were not amongst those lists of words which are commonly used, or read, by children between the ages of 6 to 10 years (Edwards & Gibbon, 1973; Thorndike & Lorge, 1944). From these word lists it became quite apparent that these six words selected were not used at all by children in the 6 to 10 years age range. The second method used was to actually try them out in a school which was situated in a middleclass residential area. The class teacher was asked to select ten children who were nearly 10 years of age and were the best readers and about whom the teacher felt that they had a wide vocabulary. Each child was seen individually and asked if they could define these six words. None of the children were either able to define or say that they had heard any of the six words selected for inclusion in this test. From this it was reasonable to deduce that it was highly unlikely that any child between the age of 6 to 10 years would have come across any of the six words selected for inclusion in the Paired Association Learning Test.

The experience of the Pilot Study suggested that this test needed some alterations. The chief reason for these amendments was that from the child's verbal responses it was difficult to be sure whether they deserved credit or not; this was particularly so during the first few trials when the child was still not <u>au fait</u> with the sounds of the totally unfamiliar words. One of the shortcomings of verbal responses is that, especially in the case of those children whose mother tongue is not English, certain phonemes in their own language tend to interfere with the acceptable production of certain sounds in English (e.g. in the case of Asians confusion with 'v' and 'w').

For the main study therefore, it was decided that the Test should be amended so that instead of making a verbal response the child made a non verbal response to ensure objectivity in scoring. The amended version of the Paired Association Test was renamed as Word Object Picture Association Test (WOPAT) in order to match the changed requirements of the new test. The details of the WOPAT can be found in Section 5.8.3 (vi) in Chapter 5.

vii) Classification and Number Tests: Unfortunately the remaining two subtests, the Classification and Number Tests turned out to be unsatisfactory and somewhat incongrous with some of the principles (see Section 3.2 in Chapter 3).

Essentially the Classification Test required the child to learn to classify the Logic Blocks (produced by Invicta) according to their colour, size and thickness. Whilst the purpose of the Number Test was to assess the child's ability to learn simple addition and subtraction by using three dimensional materials (Unifix Interlocking Plastic Cubes, produced by Philip and Tracey Limited).

These two subtests were adapted from the researches of Haynes (1971), Hegarty (1978) and Hegarty & Lucas (1978). It was very tempting to adapt and include the Classification and Number Tests in the LETB particularly since they had shown to be satisfactory by these researches. However, when these tests were tried out, their weaknesses became quite apparent. The Number Test to some extent appeared to be dependent upon the child's previous exposure to number manipulation (see Section 3.2 in Chapter 3). This was a serious limitation which was not acceptable despite the fact that in other respects (e.g. administration, discriminatory power) the Test was quite adequate.

The problem with the Classification Test was that it was difficult to be certain from the child's response whether it was as a result of understanding the task or through sheer trial and error. This weakness could not be envisaged without actually trying it out. In order to differentiate between the two types of responses, need for verbal instructions and timing of the Test seemed essential. Clearly, if this Test had been retained with these two modifications, then, it would not have been quite consistent with one of the aims and some of the principles, i.e. the Test Battery should be minimally dependent upon language and should not be timed. Because of these weaknesses it was decided neither to include them in the LETB in order to try them on the main sample nor to analyse the data.

Having thus found the Classification Test and the Number Test unsatisfactory, the author was still left with the possibility that the main study might still throw up one or more tests from the remaining six which might also turn out to be unsatisfactory when they were tried on the main sample. This would have meant that the LETB could have ended up as being a measure of a very limited range of intellectual abilities. This was not acceptable either (cf. Anastasi, 1961; Davis, 1971; Jensen, 1970 who recommend that when testing children from ethnic minority groups one should endeavour to test a fairly wide range of cognitive functioning rather than relying on a single test or an extremely small number of tests). It was therefore decided that in order to replace the two unsatisfactory Tests, two more new tests should be developed - needless to add which should be consistent with the aims of the study and with the principles (see Section 3.2 in Chapter 3).

It was these considerations which led to the development of the Object Picture Association Test (OPAT) and the Symbol Manipulation Test (SM). The exigencies of the time did not permit the piloting of these two subtests in the same way as the other six Tests of the LETB were tried out. Piloting these two subtests, like the other six, would have delayed embarking upon the main study by several months. However, prior to using the OPAT and the SM on the main sample, they were tried on a few children to determine if there were any obvious weaknesses in them. In addition, their face validity was also checked by consulting Dr.J. Hewitt (Psychology Department, Birmingham University), Mr. W. Kerr (Faculty of Education, Birmingham University) and Dr. M. Thompson (Department of Educational Enquiry, University of Aston in Birmingham): in view of the circumstances, they all agreed that the OPAT and SM should be incorporated in the LETB without any further trials. The details of these two subtests are given in Section 5.8.3 (vii and viii) in Chapter 5.

4.9 Other Tests Used

Except for the ESN-M sub-sample, all the Asian children were also tested on the Raven's Coloured Progressive Matrices (1962) and the Schonell Graded Word Recognition Test (1951, norms revised by the publisher in 1972), according to the standardised procedure set out in their respective manuals. At the pilot stage, there appeared little need to obtain an index of the ESN-M sample's intellectual ability because this information, in a way, was already known. The fact that they were in an ESN-M school would suggest that their IQ should be in the 55-70 range - a practice commonly followed by many practising educational psychologists (see Chazen et al., 1974; Mittler, 1970). The Reading Test was administered twice: first in June 1978 and then after a year in 1979. The details of the Raven's Matrices and the Schonell Graded Word Recognition Test are as follows.

i) Coloured Progressive Matrices: This test covers an age-range of 5 - 11 years and was designed as a measure of Spearman's g. Raven claims that it can be used with different types of populations, such as children with little knowledge of the language, people suffering from aphasia, cerebral palsy, deafness and with individuals of varying degrees of mental retardation.

As the Coloured Progressive Matrices are so well known (see Raven et al., 1978) , there seems little need to give any further information about this test expcept to add that the test was administered as a group test without any time limit - numbers in the group never exceeding 10 - in strict accordance with the administration procedure as laid down in the Manual. The raw scores were used for the purposes of the analysis of the data.

Test-retest reliability for children under 7 years of age is reported to be 0.65; it increases to at least 0.80 for 9 year olds. The author claims that the Coloured Progressive Matrices' correlation with the Crichton Vocabulary Scale and with the revised Stanford Binet Scale for children aged 7 years is about 0.5 and for children aged 9 years is 0.65. Some authorities (e.g. Thorndike & Hagen, 1969) consider that the evidence provided for the test's reliability and validity are "inconsistent" and the "normative data at best are sketchy" (Thorndike & Hagen, 1969).

ii) Schonell Graded Word Recognition Test: This test has a range from "6.0 minus" to 12.6+ years. Children who did not score at all on this test were given the reading age equivalent to 5.0 years. For the purposes of analysis, instead of raw scores, reading ages in years and months were computed. The test was administered individually according to the instructions laid down in the Handbook of Instructions (Schonell, 1951). The Handbook provides no information about the test's reliability or validity.

4.10 Inter-Correlations

In addition to the above analyses of individual subtests, inter-correlations were obtained among the tests which included: LETB subtests, Raven's Matrices, Teacher's Ratings and the Reading Ages for 1978 and 1979 on the Schonell Graded Word Recognition Test. These inter-correlations are presented in the following Table 4.5. It needs to be recognised that these correlations are based on a small sample therefore only limited reliance can be placed in the results because of the large standard error associated with small samples.

Except for the Teacher's Ratings (which more often than not tend to be based on the child's attainments), the Paired Association subtest has the highest correlation of all the predictors, r = 0.67, significant at a 1% level. As against the Paired Association Test, the correlation between the Raven's Matrices and the criterion (Reading 1979) is 0.58, significant at a 1% level.

For ease of inspection, Table 4.6 presents the inter-correlations between the LETB subtests only.

Inter-correlations between the Learning Efficiency Test Battery subtests show a high degree of homogeneity. All the inter-correlations between the LETB subtests (except Visual Sequential Memory and Paired Association which is significant at a 4% level), are significant at a 1% level. As one would expect, the highest correlations are between the Piaget-based tests, r = 0.89, significant at 1% level. However, the correlations do suggest that the LETB measures a number of distinct mental abilities rather than one underlying 'g': thus some of the correlations are relatively low e.g. VSMT and Paired Association is 0.43; Serial and Ordinal Correspondence and VSMT is 0.50.

Table 4.5 Correlation Matrix for Variables : LETB, Raven's Matrices Teacher's Ratings, Reading Ages for 1978 and 1979 (n = range from 30 to 34)

	1	2	3	4	5	6	7	8	9
	Teacher's Rating on the basis of attainments	Teacher's Rating on the basis of potential	Raven's Matrices	Seriation A and B	Serial and Ordinal Correspondence	Visual Sequential Memory	Paired Association	Reading 1978	Reading 1979
*1 Teacher's Rating on the basis of attainments	X	.86	.50	.56	.50	.08	.66	.89	.74
*2 Rating on the basis of potential		X	.48	.55	.52	.20	.53	.87	.66
3 Raven's Matrices			X	.24	.31	.26	.36	.51	.58
4 Seriation A and B				X	.89	.55	.62	.49	.53
5 Ordinal Correspondence					X	.50	.54	.42	.43
6 Visual Sequential Memory						X	.43	.06	.24
7 Paired Association							X	.66	.67
8 Reading 1978								X	.73

*Rescaled to avoid negative signs

Table 4.6 Correlation Matrix Showing Correlations Between the LETB Subtests

	Seriation A and B	Serial and Ordinal Correspondence	Visual Sequential Memory	Paired Association
Seriation A and B	X	.89	.55	.62
Serial and Ordinal Correspondence		X	.50	.54
Visual Sequential Memory			X	.43

4.11 Discussion and Conclusions

Several important and interesting points emerge from this small study, although it is worth reminding ourselves that only limited confidence can be placed in the results, nor can we make any generalisations.

After having eliminated the unsatisfactory tests from the Battery, it became quite clear:

(1) That the model (Demonstration $\longrightarrow$ Demonstration and Practice $\longrightarrow$ with cues and feedback), designed for this study was useful and was worth retaining for the major study.

(2) That a range of mental functioning (e.g. cognitive processes within the Piagetian framework, short-term memory, and associative type learning) related to classroom learning can be sampled by using the Learning

Efficiency Test Battery (cf. Anastasi, 1961; Davis, 1971; Jensen, 1970).

(3) That the Paired Association Test significantly predicts children's future academic performance (at this stage, we cannot claim any other subject except Reading) better than the Raven's Matrices and the teacher's estimate of children's potential .

(4) That the low performance on the Raven's Matrices should not necessarily be taken as an indication of the child's inability to profit from instruction as is commonly believed; nor should it be regarded as a final word about Asian children's level of cognitive functioning. Some children who were amongst the bottom five per cent of the population on the Raven's Matrices present an entirely different picture when their level of cognitive process is assessed by alternative means, i.e. tests based on Piaget's theory. Furthermore, not only did these children perform well on some of the Learning Efficiency Test Battery subtests (total score 50% or more), but they also made satisfactory gains on the criterion (Schonell: Graded Word Recognition Test).

(5) That the assessment carried out by employing Piaget-based items may possibly be less culture biased compared to the Raven's Matrices. If we examine the performance of these children on the Serial and Ordinal Correspondence Tests, only 3% of the total sample demonstrate little or no understanding of the concept that is, have a level of mental functioning comparable to a typical four year old child (Inhelder & Piaget, 1964). Compare this with 24% of the Asian children who according to their performance on the Raven's Matrices, fall at or below the 5th percentile level (based on U.K. norms) and would thus be viewed as in need of special education. It is perhaps this type of misclassification which may eventually result in the child's transfer to a special educational provision. If that does not happen low scores may affect the teacher's expectation.

(6) That none of the four ESN-M children demonstrated full understanding of the concepts of Serial and Ordinal Correspondence, despite the fact that one child was nearly 10 years 6 months. Equally their scores on the Paired Association Test and the Visual Sequential Short Term Memory Test are considerably lower than the children in normal schools. This finding would appear to provide some evidence of construct validity of the Learning Efficiency Test Battery.

(7) There is significantly higher correlations between the Piaget based tests (measures of cognitive functioning) and the learning tests (associative type and memory), than one would expect in the light of Jensen's theory. Jensen states:-

> "...in the low SES groups, correlations between the learning tests and IQ are in the range from 0.01 to 0.20. The correlations for middle-class children for various tests range between .60 and .80." (Jensen, 1969, pages 232-233)

None of the children in the present investigation came from a higher SES. In addition, nor do the correlations between the LETB subtests and the Raven's Matrices, which range from .24 to .36, provide support to Jensen's above hypothesis.

(8) That the inter-correlations between the LETB subtests do not seem to provide support to Jensen's "Minimum Hypothesis". Jensen states:-

> "The biological or structural basis of Levels I and 2 are thought of as independent(although they are functionally related, since the rate of asymptote of phenotypic development of Level 2 performance depends upon the individual's status on Level I processes)." (Jensen, 1969, p. 237)

In conclusion, the results of the Pilot Study were sufficiently encouraging to mount a major investigation. These results also helped in the formulation of several aims and hypothesis of the major study.

CHAPTER 5

MAIN STUDY : METHOD

5.1 Sample

A total of 455 children took part in this study which included English, Asian and Indian children (for definitions of each category see pages xvi - xvii). Their ages ranged from approximately 6 years to 10 years. Of the 455 children, 385 were drawn from normal Junior and Infant schools and 40 children were selected from special schools (designated as schools for the Educationally Subnormal - Mild, by the Local Education Authority; see Home Office, 1978 for the new terminology in place of ESN-M) in England. The total sample of 455 children also included a small sub-sample of 30 children who were tested in India. In the following pages the details of how the schools were selected, a brief description of schools and the details and breakdown of the total sample are presented.

Prior to proceeding to give fuller details about the schools and the sample, it should be observed that it would have been interesting and desirable to include children from other cultural backgrounds as well, e.g. West Indian children, as this would have permitted cross-cultural comparisons on the proposed Battery. However, this could not be achieved in this project as it would have taken an inordinate amount of time to test a large sample of children belonging to several different cultural groups (note that the LETB takes approximately one hour and a half of individual testing time). It was therefore decided to restrict the present investigation to English, Asian and Indian children only.

Footnote: Although at present the sample includes Asian, English and Indian children only, there is a strong possibility that the LETB may be tried out on other ethnic groups in the United States. If this were to happen this would be principally due to the efforts of Professor Gredler from South Carolina University and Professor Sethi from California State College.

5.2 Selection of Schools in Britain

The headteachers of several Infant and Junior schools in the East District of Birmingham, and of special schools (ESN-M only) with whom the author has professional relationships were asked if they would offer research facilities in their schools. Those headteachers who showed keenness to co-operate in this project were asked to supply specific details about each child. Although this method of selecting schools may have introduced some bias in the sample, this procedure was preferred to selecting schools at random from the area. More often than not, schools selected randomly show an inability and unwillingness to meet the needs of the research worker (e.g. a quiet room for testing, information on children, filling in of rating scales and so on), and tend to withdraw their support in the middle of the project - in fact, this happened with the present study as well, despite the fact that assurance was given that they would continue to offer co-operation until the termination of the study. While in the midstream of testing one normal school and one special school decided to withdraw their support completely, the reason given being that the project made much more demand on their time than they had anticipated. In fact, many heads began to betray signs of impatience and lack of co-operation when it came to testing the first year Junior and Infant children because of the demands the project made on them. It was anticipated that these schools would need to be used again for the purpose of collecting data on the criterion and other measures. It therefore seemed important to enlist the help of only those schools which were willing to extend their co-operation until the end of the project, rather than selecting randomly and running the possible risk of losing their support half way through the completion of the project. In view of these considerations nearly 20 schools were approached with a request for co-operation and 15 schools, including both 'normal and 'special', offered to assist.

5.3 Selection of a School in India

The school in India was selected through a personal contact. The headteacher of this school was approached by a friend in India to enquire if she would extend research facilities during the author's field trip to India. Then the headteacher of this school was informed about the relevant details of the project, the type of sample and the information that would be needed about the sample. The headteacher agreed to extend co-operation.

5.4 Brief Description of the 'Normal' Schools in Britain

Of the 12 'normal' Junior and Infant schools, seven schools had approximately 80% to 90% children from ethnic minority groups, the majority of whom were Asian. Three schools' ethnic minority population was between 20% to 40%. Two schools were situated in a predominantly 'white area' and in those schools there were no more than 5% to 10% children from ethnic minority groups.

As the author worked in the area from where the schools were drawn, it is possible to comment that all these schools appeared to have at least one thing in common, that is, a very high proportion of children came from homes where the parents either did not know how, or, did not make much effort to provide a stimulating home background. There were few children who attended these schools who came from professional homes. It would be quite appropriate to label the districts where these schools were situated as "disadvantaged".

5.5 Brief Description of the Special Schools

A common feature of all three special schools is that they all take children who have been assessed by Educational Psychologists and School Medical Officers as ESN-M. Two schools are designed to take ESN-M children only while the third school took a few ESN-S children as well.

5.6 Brief Description of the School in India

The school was situated in a large industrial city in Punjab. It was a fee paying school and children represented a cross-section of society. The academic standard of this school seemed higher than the average State school. The school accepted children from the age of three and a half to approximately 16 years. All children spoke the same language - Punjabi.

5.7 Details of Obtained Sample

5.7.1 Selection of Children from 'Normal' Schools

In the main the method adopted to select children from 'normal' schools was quite similar to the procedure adopted by Haynes (1971). For the details of this procedure see Section 4.2 in the Pilot Study.

i) Second Year Junior Sample: From the second year Junior sample available, 210 children were randomly selected. In Table 5.1 is shown the distribution of second year Junior English and Asian children according to teacher ratings and sex in the available and selected samples. It will be noticed that there were only a few "bright" children in the available sample, whilst there were a considerably high prcportion of children who were rated 5 by their teachers.

Table 5.1 Distribution of Second Year Junior English and Asian Samples According to Teacher Ratings (Attainments) and Sex in the Available and Selected Samples

Grade	English Sample				Asian Sample			
	Children Available		Children Selected		Children Available		Children Selected	
	Boys (n=116)	Girls (n=101)	Boys (n=42)	Girls (n=45)	Boys (n=215)	Girls (n=165)	Boys (n=70)	Girls (n=53)
1	0	2	0	2	2	1	2	0
2	15	13	10	8	21	17	13	10
3	38	47	18	23	56	54	29	21
4	51	33	11	9	101	55	15	14
5	12	6	3	3	35	38	11	8

Table 5.2 shows the breakdown of education experience for the second year Junior Asian Sample.

Table 5.2 Second Year Junior Asian Sample - Length of Previous Educational Experience in England

Educational Experience	Children Available		Children Selected	
	Boys (n=215)	Girls (n=165)	Boys	Girls
1 year or less (Juniors only)	29	32	13	7
Full education (including nursery)	28	35	10	4
Full education (no nursery experience)	158	98	47	42

ii) First Year Junior Sample: Of nearly 312 available children, 109 children of both races and of both sexes were randomly selected. Table 5.3 presentes the breakdowns of First Year Junior sample by ability.

Table 5.3 Breakdown of First Year Junior English and Asian Samples According to Teacher Ratings (Attainments) and Sex in the Available and Selected Samples

Grade	English Sample				Asian Sample			
	Children Available		Children Selected		Children Available		Children Selected	
	Boys (n=81)	Girls (n=97)	Boys (n=31)	Girls (n=24)	Boys (n=70)	Girls (n=64)	Boys (n=18)	Girls (n=36)
1	-	-	-	-	-	-	-	-
2	10	18	5	6	11	6	4	7
3	33	50	13	11	18	18	7	16
4	31	25	9	4	26	26	6	6
5	7	4	4	3	15	14	1	7

Table 5.4 First Year Junior Asian Sample - Length of Previous Educational Experience in England

Educational Experience	Children Available		Children Selected	
	Boys	Girls	Boys	Girls
One year or less	None		None	
Full education (including nursery)	10	12	0	6
Full education (no nursery experience)	70	64	18	30

Table 5.4 shows that unlike the second year Junior children, among the first year Junior children there was no Asian child who was a recent arrival to this country or had a restricted educational experience.

iii) Top Infant Sample: For this age range, 97 children of both sexes and of both races were available. From this available sample 66 children were randomly selected following the same procedure as was adopted for the second and first year Junior children. The details of this sample are presented in Table 5.5 and Table 5.6. Table 5.5 shows the breakdown of the top Infant sample according to teacher's ratings based on attainments, race and sex in the sample that was available and actually selected. In Table 5.6 is shown the length of previous educational experience of the Asian sample in Britain. It is interesting to note that in this sample there was hardly any child who went to a nursery school or had a limited educational experience.

Table 5.5 Breakdown of Top Infants English and Asian Samples According to Teacher's Ratings (Attainment), Race and Sex in the Available and Selected Samples

	English Sample				Asian Sample			
Grade	Children Available		Children Selected		Children Available		Children Selected	
	Boys (n=26)	Girls (n=35)	Boys (n=15)	Girls (n=15)	Boys (n=48)	Girls (n=49)	Boys (n=16)	Girls (n=20)
1	-	-	-	-	-	-	-	-
2	7	5	2	2	9	10	2	5
3	11	21	8	8	12	10	9	8
4	7	6	4	4	17	17	3	3
5	1	2	1	1	10	12	2	4

Table 5.6 Top Infants Asian Sample - Length of Previous Educational Experience in England

Educational Experience	Children Available		Children Selected	
	Boys	Girls	Boys	Girls
One year or less	None	None	None	None
Full Education (including nursery)	None	None	None	None
Full education (no nursery experience)	48	49	16	20

5.7.2 Details of the ESN-M Sample

Barring multiply handicapped children, all the children of both sexes and of both races in the age range 6+ to 9+ available in the three schools were selected. The breakdown of this sample is shown in Table 5.7.

Table 5.7 Breakdown of English and Asian ESN-M Sample According to Race and Sex

English		Asian	
Boys	Girls	Boys	Girls
23	6	8	3

Although not strictly pertinent to comment here, it is worth observing that in the case of English as well as Asian children, far more boys seem to be in the ESN-M category than girls.

5.7.3 Details of the Indian Sample

All children who were between 8 to 9 years were selected. There were 16 boys and 14 girls.

In Table 5.8 the mean ages and SD of the total sample according to race, year and sex are presented. Table 5.8 also shows that in each year group the mean Chronological Ages and the Standard Deviations of both races and of both sexes are very similar. The Mean Chronological Age of both the Indian and ESN-M children are very close to the second year Junior children.

Table 5.8 Mean and SD of Chronological Ages of 455 Children According to Race, Year and Sex

	Boys		Girls	
	M	SD	M	SD
Top Infants				
English	6.83	.48	6.83	.34
Asian	7.14	.44	7.25	.33
First Year Junior				
English	8.10	.29	8.05	.32
Asian	8.06	.29	8.11	.28
Second Year Junior				
English	8.86	.23	8.92	.25
Asian	8.94	.29	8.95	.27
ESN-M				
English	8.55	.88	8.89	.52
Asian	8.82	1.05	8.67	1.10
Indian	8.43	.38	8.41	.34

5.8 Tests Used in the Main Study

Apart from the proposed Battery (Learning Efficiency Test Battery - LETB) the other tests used fall approximately into two categories: i) measures of cognitive areas; ii) measures of academic areas.

Each of the measures used is described in the following pages. For tests other than the LETB, an effort was made to ensure :-

i) That their inclusion in the total Test Battery was compatible with the aims of the study.

ii) That the tests covered a wide range of mental functioning.

iii) That the measures of academic ability and intelligence tests had the desired psychometric attribute, and had a satisfactory upper and lower limit to cover both 'bright' and 'retarded' children.

iv) That their administration and scoring should not take inordinate amounts of time since in the collection of the data, time is always a key factor.

5.8.1 Measures of Intelligence

It will be recalled that for the Pilot Study only one test of intelligence, Raven's Matrices, was used. For the main study it was decided to use the Goodenough Draw-a-Man Test (Harris, 1963) as well. At the time of the Pilot Study the Draw-a-Man could not be used because the author was given the impression, by a friend who runs a course in multicultural education, that Islam prohibits the drawing of human figures. Since the sample for the Pilot Study did include some Muslim children the author did not wish to offend their parents.

However, later enquiries revealed that although Islam does forbid the drawing of human figures, Muslim parents should not object to the type of task required in the Draw-a-Man Test. When the author found this information it was too late to include the Draw-a-Man Test in the Pilot Study. However it was decided that it should be included in the main study.

The Raven's Coloured Progressive Matrices and the Draw-a-Man Tests were preferred to the WISC-R (Weschler, 1974) and the Stanford Binet (Terman & Merrill, 1960) Tests for the reasons outlined above and also for the following reason. The literature seems to indicate that the former two tests have been far more widely used in the cross-cultural research than the latter two (see Anastasi, 1976; Raven et al., 1978). Since the present research involves children from Britain, Bangladesh, India and Pakistan, i.e. has a "cross-cultural" dimension to it, there seemed more justification in including the Raven's Coloured Matrices and Draw-a-Man Tests than the WISC-R and/or Stanford Binet Tests. It is hoped that the inclusion of these two tests would throw some light on their usefulness, or lack of it, as predictors of academic performance of both Asian and English children.

i) Raven's Matrices: The details of this test have already been described in Section 4.9 (i) in Chapter 4.

ii) Draw-a-Man: This test was one of the earliest attempts at devising a "culture-free" test. The Draw-a-Man Test has been used since 1926 and been widely used in clinics and across diverse culture groups. The test was included because the author's personal experience shows that it is extensively used by School Medical Officers, Remedial Teachers and Psychologists for obtaining an index of children's ability and for the purposes of determining the future educational needs of children - particularly with children with little knowledge of English language. The additional appeal of this test to a research worker is its ease of administration and the fact that it does not take up too much time. The children in the sample were given this test in small groups - never exceeding 10 - without any time limit. It was the revised version of the Draw-a-Man Test that was used for this research (Harris, 1963). Since it is a well-known test, suffice to say here that it is claimed that the Draw-a-Man Test is a measure of the child's accuracy of observation and the development of conceptual thinking. Its various types of reliabilities (e.g. test-retest, split half, inter-scorer and so on) are reported to be in the region of .69 to .89 (Dunn, 1967; Harris, 1963). Information about the test's validity is provided by correlating it with other intelligence tests. There is a fair variation in these correlations but the majority of them are in the region of 0.50 (Anastasi, 1976).

5.8.2 Attainment Tests

For the purpose of obtaining the child's level of reading and mathematics standard, all the children except the Indian and ESN-M subjects were given the Schonell Graded Word Reading Test (Schonell, 1951) and the Graded Arithmetic-Mathematics Test (Vernon & Miller, 1976) respectively. At the outset of this section, the reasons which dictated the selection of these and other tests have already been given and therefore there is little need to repeat them here. In view of this, only minimum but necessary, details of each attainment test are provided.

i) Schonell Graded Word Reading Test: For details of this test see Section 4.9 (ii) in Chapter 4.

ii) Graded Arithmetic-Mathematics Test: This is a very well known test and has been in use for nearly twenty-five years. For the purposes of the present research, the first form of the revised version which covers the age range from 5 to 12 years was used. Vernon & Miller claim that the Junior version has items of sufficient difficulty which would "tax" even fairly bright 12 year old children. The Test Manual provides norms for English, Scottish and Canadian children. For the purposes of the analysis of the data, raw scores were converted according to the English norms. Children who failed to score were given a score of 4.0 years. Vernon & Miller do not provide any information about the validity, but report the reliability coefficient, after correction, to be "0.90 or over". The test was administered orally in groups of 6 - 10 children, according to the instructions set out by the authors.

5.8.3 Learning Efficiency Test Battery

Five of the eight subtests of the LETB were used exactly as in the Pilot Study. They were:-

i) Seriation A

ii) Seriation B

iii) Serial Correspondence

iv) Ordinal Correspondence

v) Visual Sequential Short Term Memory (VSMT)

For details of these subtests see Section 4.8 (i) to (v) in the Pilot Study. However, the description of the subtests which were not properly tried out or had to be amended in the light of the Pilot Study are as follows. There are three such tests: Word Object Picture Association Test (WOPAT), Object Picture Association Test (OPAT) and Symbol Manipulation Test (SM). All three subtests are a derivation from the Paired Association Test (see Section 4.8 (vi) in the Pilot Study for description and the reasons for changing it). The details of each of these three subtests and the way they were used in the main study are given below.

vi) Word Object Picture Association Test: In this test the child has to learn to associate the appropriate label with the predetermined wooden Block and both of these with the picture of the label which was drawn on a 5 x 5 cm card. There are six such associations to be learnt. The number of trials to reach the criterion level was fixed (six). In selecting the labels (Femur, Murex, Libra, Cacti, Gimlet, Brazier) particular attention was paid to ensure that a large majority of children of the age range on whom this test was designed to be standardised should not have met these words. (This was done by asking 20 English children, 9.0 - 9.6 years old, - randomly selected from two schools situated in mainly residential areas - if they knew the meaning or were familiar with these six words. None of the children asked knew either the meaning or were familiar with these six words.) Even if somebody may be familiar with one or more of these labels, hardly any child would come to the test situation having already come across an association between the label and the wooden Block and the two together with the picture of the object.

vii) Object Picture Association Test: There are some similarities between this test and the Word Object Picture Association Test. This test, too, required the child to learn to associate an irregular shaped wooden Block with a very familiar object (e.g. Tree, Cup, Face). There are six such associations to be learnt. As the shape of each wooden Block is completely irregular it is therefore highly unlikely - almost impossible - for any child to come to the test situation having already learnt one or more association.

viii) Symbol Manipulation: Unlike the rest of the tests, this is a pen and paper test. Although it is mainly a two dimensional test, the manipulation of stimuli is very simple. The essential requirement in this test is to learn to associate one abstract symbol with another abstract symbol. Throughout the test the stimulus model is always in front of the child. The test starts first having to learning to associate one symbol with another, and gradually progresses to four. In other words, the child has to select - in fact draw - four symbols from the stimulus model which correspond with the test items. Although part of the test is based on two dimensional representation it does not require complicated manipulation of the stimulus.

For administration of the LETB and scoring see Appendix A. On average it took approximately one hour and a half to test each child on the LETB. The LETB and the rest of the tests were presented in the following order:-

Seriation A

Seriation B

Serial Correspondence

Ordinal Correspondence

Visual Sequential Short Term Memory

Word Object Picture Association Test

Symbol Manipulation Test

Raven's Matrices

Draw-a-Man Test

Schonell Graded Word Recognition Test

Graded Arithmetic-Mathematics Test

Invariably the testing was completed in two or three sessions depending upon the child's interest, motivation and co-operation.

5.9 Testing

First Phase: During the first phase of the main enquiry, the follwing children were tested:

Table 5.9 Number, Description and the Time When Most of the Children During the First Phase Were Tested

Description of Children	n	When Tested
2nd Year	210	Jan - April 1980
1st Year	109	Jan - April 1980
Top Infants	66	Jan - April 1980 and Sept - Dec 1980
ESN-M	40	Feb - March 1981
Indian	30	Jan - Feb 1981

Second Phase: During the second phase of the enquiry, after one year, the second year Junior, first year Junior and Top Infants were retested on criterion tests. On retesting, some subjects from all the year groups were lost because some were out of the country and some had moved out of the area and could not be conveniently approached. The number of subjects lost were :- nine from second year Junior, five from first year Junior, and three from Top Infants.

5.10 LET Battery and Jensen's Theory

From the standpoint of Jensen's theory (1969, see also, 1980 b), the face validity of the tests of the LET Battery, and the theories on which they are based (Associative Learning and Piagetian), it would appear that the Piagetian tests would be subsumed as measures of Level II ability, whilst the learning tasks and the Visual Sequential Short Term Memory Test as measures of Level I ability.

Discussing the psychometric properties of Piagetian tests, Jensen (1980 b) considers them as measures of g. Jensen states:-

> "...that [Piagetian items] get at the most fundamental aspects of intellectual development...[and] that the general factor of the Piagetian battery is almost pure g in the Spearman sense." (1980 b, p. 675)

And g according to Jensen measures Level II ability. Thus there are four Piaget based subtests which under the rubrics of Jensen's hypothesis of mental abilities would be subsumed as measures of Level II ability.

CHAPTER 6

RESULTS: I PSYCHOMETRIC PROPERTIES OF THE LET BATTERY

6.1 Mastery Learning Tests and the LET Battery

When devising most of the items of the LET Battery, attention was paid to the fact that a large majority of children with a predetermined amount of teaching input should be able to master a high proportion of them. It was necessary to construct such items otherwise it would not have been possible for the LET Battery to perform some of its intended functions which have been stated in the previous pages. (How the LET Battery would be able to execute those functions is discussed in the succeeding pages, see Section 8.5 to 8.5.4 in Chapter 8.)

The consequence of designing such items is that the scores are not as 'Normally' distributed or spread out as is often expected and found in conventional tests designed to assess the whole range of abilities. Although sometimes the 'Normal' or near 'Normal' distributions of scores is regarded as a sine qua non in the construction of tests, at the same time, it needs to be borne in mind that employing the 'Normal' curve in the scaling of tests is simply a matter of convenience and is not based on "Normal distribution of behaviour" in nature (Cronbach, 1970, p. 100; see also Anastasi, 1976; Nunnally, 1978 who maintain that it is permissable to devise tests for restricted ranges only).

With tests such as the LET Battery, which are very close to the mastery learning tests, it is commonly acknowledged that the key point in their design is their:-

> "...appeal to the appropriateness of the content rather than in terms of any experimentation or statistical results as would be required for predictive validity or construct validity". (Nunnally, 1978, p. 310)

Furthermore, if mastery learning tests show only a minimal difference in the level of performance of children, Nunnally maintains that one should not be unduly concerned because this would result in almost negligible variance, reliability and correlation with any other test.

Lack of desirable levels of variance, reliability coefficient and correlation coefficient, however, does not negate the value of the test in terms of its content validity:-

> "...and at least in principle, there is nothing strange about this circumstance occurring." (Nunnally, 1978, p. 310)

New procedures to analyse data yielded by mastery learning tests are underway (e.g. Ferguson & Novick, 1973; Millman, 1974, cited in Anastasi, 1976) but such data analysis techniques are still at their exploratory stage (Anastasi, 1976). Because of the uncertainty surrounding the value of new techniques, it was therefore decided to adhere to the long established methods of determining internal consistency, test-re-test reliability, validity, and for carrying out other analyses. However, it is important to remember that because these analyses are based on skewed distributions therefore any statistically significant results should be interpreted with some caution.

6.2 Factorial Analysis of the LETB Items

It seems important to explain at the outset, prior to proceeding to give details of the factorial analysis of the LETB items, that most of the items of the Battery may not be, strictly speaking, considered as items in the conventional sense of the term; they appear to fall between a subtest and an item. Consider for example, Piaget based tests. For the purposes of the present analysis each Piaget subtest is treated as consisting of one item only. It could be argued that in fact each subtest consists of three items. On the other hand, these three items did not progressively become difficult as is commonly found in conventional tests owing to the special administration model that was developed for the LET Battery (the details of the model are provided on pages 86 to 87). It therefore seemed preferable to treat each Piaget based subtest as consisting of one item only. It seemed more important that rather than trying to impose a typical character on an item and by so doing make them inconsistent with the aims of the study, that the presentation of these items should remain congruent with the general aims of the study. Since it was not possible to achieve both goals the choice fell to the latter assuming that by so doing it would not in any serious way effect the psychometric properties of the LET Battery.

Turning to the details of factor analysis of the LETB items, it is important to ascertain, first of all, whether all the items claimed to belong to a subtest load substantially on a factor representing that subtest. It is often recommended that the items of a subtest should load on one major factor and:-

> "The best measures of each factor will be those that correlate highly with one factor and have low correlations with other factors." (Nunnally, 1978, p. 274)

If factor analysis of the data reveals such results as implied in Nunally's statement, a test is then said to be homogeneous in content.

Thus in order to determine whether the subtests of the LETB were homogeneous in content, firstly, a Principal Component analysis of all the LETB items was carried out to decide how many factors should be retained. For the purposes of this analysis, Top Infants, First Year Junior and Second Year Junior children were combined; the total number of children was 385. To reduce the size of standard error generally associated with small samples it was considered preferable to combine all the children for factor analysis rather than carrying out separately for different races and age groups. Some of the subsequent analyses of results demonstrate that by so doing no important statistical information about the groups and races has been lost.

Both the "eigenvalues greater than one" rule (Child, 1970) and the "Scree test" (Cattell, 1966) suggested that six factors should be extracted. Having thus decided on the number of factors to be used, this was followed by subjecting the data to Principal Factor extraction followed by two methods of rotation, Varimax and Oblique (Delta = 0.000). (All programmes were as implemented in SPSS, Version 5 - see Nie et al., 1975) Initially, Varimax rotation was used to determine an orthogonal simple structure for the LET Battery and the details are presented in Table 6.1.

Table 6.1 Varimax Rotated Factors Matrix of the LETB Items (Salients Underlined)

	FACTORS					
	1	2	3	4	5	6
RSSA	14	08	69	05	15	08
RSSB	25	12	73	04	08	13
RSSC	17	09	66	20	10	02
RSOC	16	22	41	27	00	04
VSMT 1	08	40	10	-06	19	07
VSMT 2	12	44	17	-07	18	07
VSMT 3	08	59	05	-01	05	02
VSMT 4	11	57	11	-05	09	03
VSMT 5	14	56	00	15	-02	12
VSMT 6	06	63	07	14	04	03
VSMT 7	11	60	01	20	-02	11
VSMT 8	10	65	08	18	-03	14
OPAT 1	26	25	14	13	14	89
OPAT 2	29	27	15	12	15	77
WOPAT 1	11	11	17	87	07	10
WOPAT 2	16	16	19	86	03	10
SM 1	26	17	18	08	86	12
SM 2	36	13	17	06	77	14
SM 3	76	17	22	16	20	10
SM 4	79	18	17	17	22	12
SM 5	90	13	17	09	15	12
SM 6	90	16	17	02	13	09
SM 7	86	20	13	04	10	15
SM 8	83	15	15	07	06	13

Decimal Points Omitted

Table 6.1 shows that all but the first one or two items (criterion for meaningful loading = .3 or above; Child, 1970) of the Symbol Manipulation Test (SM) clearly load on Factor 1. The second item of this test loads on Factor 1 as well as on Factor 5; item 1 also loads on this factor. Thus items 1 and 2 form a cluster on a separate factor suggesting that they are not quite a part of the same subtest, Symbol Manipulation (this issue will be taken up again when it will be discussed as to why the first two items of the Symbol Manipulation Test form a separate cluster).

On Factor 2 load all the items of the Visual Sequential Short Term Memory subtest (VSMT). On the third, fourth and sixth Factors cluster items from the Piaget based subtests, Word Object Picture Association Test and the Object Picture Association Test respectively.

Although the Varimax rotation yielded a reasonably clear picture of the underlying factors of the Learning Efficiency Test Battery subject to the constraint that they be orthogonal, it was decided to perform an Oblique rotation as well to determine if, allowing for correlated factors, it would produce a clearer solution. The outcome of an Oblique rotation (Direct oblimin; Delta = 0) are presented in Tables 6.2 and 6.3.

Examination of the Oblique and Varimax Pattern Matrices suggests that the major difference between the two rotations is that in the former, as expected, the loadings are somewhat clearer, i.e. close to 1 and 0. For instance the loading of RSSB on Factor 1 in Table 6.1 (Varimax rotation) was 0.25 whilst in Table 6.2 (Oblique solution) it has been pushed down to a mere 0.09. The outcome of the Oblique rotation meets far better Thurstone's (1947) simple solution criterion than the Varimax rotation.

Table 6.2 Oblique Factor Pattern Matrix of the LET Items*
(Direct oblimin, Delta = 0.000)

	FACTORS 1	2	3	4	5	6
RSSA	-04	-03	-03	05	05	72
RSSB	09	-02	-07	-05	08	76
RSSC	02	-02	12	02	-03	68
RSOC	06	15	20	-06	-02	39
VSMT 1	-02	39	-09	15	04	04
VSMT 2	01	42	-12	13	03	12
VSMT 3	02	61	-06	01	-04	-00
VSMT 4	03	58	-10	04	-02	-06
VSMT 5	07	55	10	-07	08	-08
VSMT 6	-03	64	09	01	-03	00
VSMT 7	04	59	15	-07	06	-07
VSMT 8	-00	64	11	-09	10	00
OPAT 1	-02	-02	-00	01	99	01
OPAT 2	04	03	-00	02	85	03
WOPAT 1	-00	-00	87	10	04	06
WOPAT 2	04	04	85	05	04	09
SM 1	02	04	10	90	05	00
SM 2	16	-01	06	79	06	-01
SM 3	75	01	09	10	-01	07
SM 4	79	01	10	12	01	00
SM 5	95	-04	00	02	-00	00
SM 6	97	08	-07	-00	-03	02
SM 7	91	04	-06	-03	04	-03
SM 8	89	-00	-03	-07	03	00

* Salients underlined
Decimal points omitted

Table 6.3 Factor Correlations of the LETB Items

	FACTORS					
	1	2	3	4	5	6
Factor 1	1.00	34	21	41	48	42
Factor 2		1.00	19	23	41	28
Factor 3			1.00	-02	23	25
Factor 4				1.00	32	35
Factor 5					1.00	31
Factor 6						1.00

Decimal Points Omitted

Table 6.2 shows that all the items of each subtest form separate clusters (except the first two items of the Symbol Manipulation Test which load on Factor 4). Barring the first two items of the Symbol Manipulation Test, it can be seen that the LET Battery has five underlying factors and the Battery, on the whole, is homogeneous in content. As all the Piaget based subtests have substantial loadings on the same factor (Factor 6) it was decided to treat all these subtests as one test (in the subsequent analysis the scores from these four subtests have been combined together and labelled - 'Piaget Test').

However, Oblique rotation did not resolve the problem of the Symbol Manipulation Test's first two items forming a separate cluster (alluded to earlier as well) indicating that they do not really belong to the same test. The fact that the first two items of the Symbol Manipulation Test load on a separate factor (although this factor is correlated at the 0.41 level with Factor 1 where the remaining six items load) suggest that they do not really belong to the same test. This was particularly puzzling as the face validity of these two items beyond doubt suggested that they should belong with the rest of the six items of this test. The explanation must lie in the fact that these two items are so easy that nearly 93% of children pass each item. Thus there is hardly any variance in the score distributions of these two items which must reduce the correlations of these two items with the rest of the items of the subtest and which in turn must influence the factor analysis outcome. This becomes clear from the following correlation matrix (Table 6.4) of the Symbol Manipulation Test (SM) items. Although all these correlations are significant at the 1% level, the correlations of items 1 and 2 are much smaller with the rest of the six items as opposed to the inter-correlations amongst items 3, 4, 5, 6, 7 and 8.

Table 6.4 Correlations of Symbol Manipulation Test

	SM1	SM2	SM3	SM4	SM5	SM6	SM7	SM8
SM1	1.00							
SM2	.84	1.00						
SM3	.47	.50	1.00					
SM4	.49	.52	.90	1.00				
SM5	.43	.50	.81	.81	1.00			
SM6	.41	.49	.77	.80	.92	1.00		
SM7	.40	.46	.69	.74	.84	.86	1.00	
SM8	.36	.41	.67	.72	.81	.82	.90	1.00

The second explanation must reside in an artefact due to the test administration procedure. Except for one subtest, Word Object Picture Association Test, on the remaining seven subtests of the Battery children were not allowed to proceed with subsequent items if they failed to reach a predetermined criterion on the first few items. The procedure, therefore, precludes the possibility of getting any items right from amongst the later items of a subtest through sheer chance alone. This artefact is inevitably bound to have some effect on the correlations and thus on the factor analysis results as well. (For further details of the administration of the LETB see Appendix A.)

Notwithstanding the above explanation, strictly speaking, on the grounds of factor analysis results, these two items cannot justifiably be retained with the rest of the items. However, it would seem that their satisfactory face validity and their usefulness for the purposes of screening would over-ride the statistical reasoning. Besides employing the LET Battery for collecting data for the research, the author has also used this Battery in his professional work, particularly, when children had been referred to the Centre for determining their future educational needs or their rate of learning or their level of thinking (within the Piagetian framework only). At that time, the author found these two items of the Symbol Manipulation Test quite useful. Children who flounder with these two items invariably failed on the other items of this as well as the other subtests. In a sense, therefore, these two items would appear to be a good predictor of the child's performance on the rest of the Battery. In view of these considerations it is suggested that the first two items should not be scored but kept in the Symbol Manipulation Test for the purposes of screening only. (See Cronbach, 1970, who argues that by dropping undesirable items it may make the test pure but by so doing the test "no longer represents the intended universe" (p.148).) This practice was therefore adopted in all the subsequent analysis.

6.3 The Reliability of the Learning Efficiency Test Battery

Reliability is one of the key aspects of a test. It may be defined as:-

> "...the consistency of scores obtained by the same persons when re-examined with the same test on different occasions, or with different sets of equivalent items, or under other variable examining conditions. (Anastasi, 1976, p. 103)

The reliability of a test is usually expressed either by a reliability coefficient or by the standard error of measurement. The most well known and frequently used methods of estimating reliability are; test re-test, parallel-forms, split-half and internal consistency. The details of these various methods are available in several standard books dealing with measurement and evaluation in psychology and education (e.g. Anastais, 1976; Lewis, 1974; Nunnally, 1978; Thorndike & Hagan, 1969).

To determine the reliability of the Learning Efficiency Test Battery, internal consistency and test re-test methods were employed. Although a useful method of estimating reliability, alternate and parallel form method could not be used in estimating the reliability of the proposed Battery as it would have taken an inordinate amount of time to develop a parallel test; because such a test has to meet the criteria of parellelism, e.g. test content, type of items, instructions for administering and so on.

The split-half method, although useful in many situations and easy to compute, did not seem suitable, particularly for the Piagetian based, Visual Sequential Short-Term Memory and the Symbol Manipulation Tests, because these tend to be hierarchially ordered rather than equivalent. Further the split-half method would have yielded rather inflated estimates of reliability, because of the inherent property of the method when used with tests such as those included in the LETB where subjects are prevented from proceeding to later items based on performance on earlier items (Anastasi, 1976; Nunnally, 1978). Also as the LETB is designed to discriminate only the bottom 5% of the population, any estimates of reliability are likely to be somewhat exaggerated. In view of this it was decided that estimates of reliability should be obtained by means of alpha coefficient and test re-test methods. These two methods, for our purposes, should be quite adequate for providing a reasonable estimate of the LETB's reliability.

Prior to proceeding to providing details about the reliability coefficients, it seems apposite to mention here that ideally reliability coefficients should have been determined on the bottom 5% of the children rather than on the whole samples. Ideal it might have been, but it was not practical. In order to have sufficient numbers who fell in the bottom 5% of the population for the purposes of analysis large numbers of children would need to have been tested. It has already been mentioned that the administration of the LETB takes about one hour and a half to test one child. Thus the time was the prime constraint in obtaining the ideal sample.

As it has been indicated earlier (and the issue is taken up again in the following pages) on the whole scores on the subtests of the LET Battery have negatively skewed distributions. Clearly such distributions are likely to underestimate the correlations since the maximum values of the correlations are usually attained when the two variables are normally distributed (Cronbach, 1970; Jensen 1980 b).

The artefact of the testing procedure which has been alluded to earlier (see pages 129-130) needs to be taken up here again, albeit briefly, because of its effect on the reliability coefficient. In all but one subtest of the LET Battery if children were unable to attain a predetermined standard on a subtest they were not allowed to proceed with the remainder of the subtest. For instance, with Seriation A, if the children were unable to seriate with the 6 Demonstration and Practice Blocks after they had had four demonstration trials, they were not allowed to proceed with the actual subtest. Thus the testing procedure did not permit them any opportunity to get any of the later items right by sheer chance alone. (For details of the testing procedure for the rest of the subtests see Appendix A.) This artefact is likely to inflate the obtained correlations.

It is, therefore, important in considering the reliability coefficients to bear in mind the facts that they are based on the whole sample and that there is an artefact of the administration procedure.

6.3.1 Internal Consistency

The determination of the internal consistency of a test is based upon single administration of the test and it refers to the degree of relationship amongst the test items (Nunnally, 1978). Nunnally considers coefficient alpha (Cronbach, 1951) as the fundamental formula for obtaining the reliability which is based on internal consistency.

Coefficient alpha was chosen because it can provide a satisfactory measure of reliability in most situations. Furthermore, Nunnally recommends that coefficient alpha should always be obtained prior to estimation of any other types of reliability coefficient. Coefficient alpha was obtained by using the following formula:-

$$\alpha = \frac{k\,\overline{r}}{1 + \overline{r}\ (k - 1)}$$

Where

α = Coefficient alpha

k = Number of items

$\overline{r}$ = Average correlation between the subtests.

Coefficient alphas, for all subtests of the LETB for Top Infants, First Year Junior and Second Year Combined (n=385) may be found in Table 6.5. In the same table, for the purposes of comparison, test re-test reliability coefficients have also been included.

Table 6.5 Test Re-Test Reliability Coefficients and Coefficient Alphas of the LET Battery (n=385)

	Tests	Coefficient Alphas	Test Re-Test Reliability Coefficients
1.	Piaget based test	.77**	0.67**
2.	Visual Sequential Short Term Memory	.80**	0.83**
3.	Object Picture Association	.94**	0.46*
4.	Word Object Picture Association	.93**	0.89*
5.	Symbol Manipulation	.96**	0.98**

* Significant at the 2 percent level

** Significant at the 1 percent level

6.3.2 Test Re-Test

The major source of error variance for test re-test reliability is due to "time sampling", but nevertheless it is regarded as "the most obvious method" of estimating a test's reliability (Anastasi, 1976). In order to obtain test re-test measures of reliability from nearly 120 available children, 23 were randomly selected to represent children of both sexes and of both races. Ideally, the number selected should have been considerably higher but due to the constraints of the time involved in testing and re-testing, the author had to be content with a small group. It should be made clear that for the purposes of estimating test re-test reliability, a new sample was identified rather than re-testing a proportion of the same children who were originally tested. The age range of the sample was between approximately 7 years to 8 years 6 months.

The gap between the first testing and the second testing was three weeks. The test re-test reliability coefficients are shown in the same Table (6.5) as the coefficient alphas.

It will be noted that both the alpha coefficients and the test re-test reliability coefficients (except for the Object Picture Association Test in the case of test re-test where it was significant at the .02 level of significance) are significant beyond the .01 level of significance.

Not only are the reliability coefficients of the LET Battery of satisfactory magnitude, but they would also appear somewhat better than Haynes (1971) obtained (the pioneer work in this field in the United Kingdom) and the NFER's later adaptation (Hegarty & Lucas, 1978) of Haynes' original test. The reliability coefficients of Haynes' test ranged from -0.02 to 0.77 and of the NFER's ranged from 0.50 to 0.80. The reliability coefficients of the LET Battery are not only better than the tests of the same category but they also compare very favourably with conventional and well established IQ tests. Take, for instance, often used tests such as the WISC-R (1974), Raven's Matrices (Raven <u>et al</u>., 1978), and Draw-a-Man test. The reliability of these three tests range from .63 to .85 (WISC-R, at the chronological age of $7\frac{1}{2}$ years); of the Raven's Matrices from .69 to .93 (age range $6\frac{1}{2}$ to $12\frac{1}{2}$ years); of the Draw-a-Man test around .70 (not reported in Harris, 1963 but in Jensen, 1980 b). None of these reliabilities reach beyond .85. Whilst in the case of the LET Battery 3 out of 5 coefficients alphas are more than .90; 1 out of 5 test re-test reliabilites are well over .90.

A point, however, worth observing is that the reliability coefficients of the LET Battery may be slightly higher on account of the artefact of the test administration procedure and the lack of spread of scores at the top end. At the same time it is also worth observing that the alpha coefficients tend to be conservative estimates of reliability (Nunnally, 1978). This would suggest - though loosely speaking - that the effect of one is being counterbalanced by the other. Therefore it would seem that these reliability coefficients may be safely treated as reasonably stable and dependable.

6.4 Validity

Like reliability, validity is an important aspect of test construction. A test is considered valid if it measures what it is intended to measure (Lewis, 1967; Thorndike & Hagen, 1969). There is a growing awareness and emphasis that the newly developed tests should be both reliable and valid (Kerlinger, 1973).

The bulk of literature on mental measurement (e.g. Anastasi, 1976; Cronbach, 1970; Garrett, 1961; Vernon, 1960) usually refers to three main types of validity : face validity and content validity; construct validity; and criterion-related or predictive validity. Although satisfactory validity is one of the key features of a test, it is not always possible or relevant to determine all types of validity. It is therefore very rare to come across a test where all the four types of validity have been reported. For the purposes of the present Test Battery three types of validity (face/content, construct and predictive) were considered pertinent and adequate.

6.5 Face/Content Validity

At the heart of face validity the key question is : is there any correspondence between what the test looks like and what it is intended to measure? Although it is no more than a "first step" and should not be considered as "the final word" in the test development, face validity nevertheless would seem an important aspect of validity, particularly for a test battery like the LETB (Garrett, 1961). Since some of the subtests of the LET Battery were intended to measure the child's learning efficiency it seemed important that the materials and the administration procedure should look as if they do measure the child's ability. After the items of the test were constructed they were shown to several well qualified educational psychologists and psychologists working in University Departments for their opinion about its face validity. All psychologists consulted agreed that the LET Battery would appear to have satisfactory face validity.

6.6 Construct Validity

This refers to the extent to which a test measures the theoretical traits or constructs which it is intended to measure. Investigating the construct validity of a new test is particularly important. One approach to construct validity would use factor analysis to determine how the factors underlying the LETB correlate with or are independent of other well known constructs such as intelligence (as measured by IQ tests) and academic ability (as measured by attainment tests) and teachers' ratings. The appropriate factor analyses are reported elsewhere (see Section 7.2 in Chapter 7) in relation to testing some of Jensen's hypotheses. It is clear that the LETB subtests do not simply measure intelligence as measured by the Raven's Coloured Matrices and Draw-a-Man tests (see Table 7.4 in Chapter 7) and thus possess a measure of construct validity.

Another approach to construct validity is to examine certain group differences (Anastasi, 1976; Lewis, 1967; Thorndike & Hagen, 1969). The rationale underlying this is that a "theory" will quite frequently indicate that the performances of the contrasted groups on the measure to be validated should differ significantly from each other. Clearly a group of ESN-M children should theoretically possess a lower "learning efficiency" than a group of normal children - thus appropriate contrasted groups seem to be ESN-M versus normal.

For the purposes of the present study the contrasted groups are all ESN-M children (n=40) and the rest of the groups (n=385) which include Second Year Junior, First Year Junior, Top Infants but not the Indian sample. In order to test whether there were significant differences between the performance of these two groups on the LETB, Analysis of Variance was performed. The means, standard deviations, F ratios and eta squared values are shown in table 6.6 below:

Table 6.6 Means, Standard Deviations, F Values and Eta Squared Values of ESN-M and Non ESN-M Children on the LETB Subtests

	NON-ESN-M		ESN-M				
	M	SD	M	SD	F.Value	Eta.sq.	Sig. Level
(Age)*	52.35	(9.24)	55.95	(10.37)	5.36	0.01	.02
Piagetian Tests	9.44	(2.76)	1.70	(3.02)	279.77	0.40	<0.001
VSMT	18.59	(6.08)	3.92	(5.22)	216.34	0.34	<0.001
OPAT	32.04	(7.00)	15.70	(12.51)	164.32	0.28	<0.001
WOPAT	18.53	(10.51)	2.86	(3.57)	87.79	0.17	<0.001
SM	22.58	(5.66)	2.88	(7.43)	411.83	0.49	<0.001

* In months over 4 years

There are statistically significant differences at the .02 level between the ages of non ESN-M and ESN-M children; the latter group is marginally older than the former group. Likewise there are highly statistically significant differences beyond the 1% level between the performances of these two groups on all the subtests of the LET Battery. An examination of the etas suggests that of all the subtests the highest proportion of the variability best explained by group membership are for the Piaget based tests (eta squared = .40) and the Symbol Manipulation Test (eta squared = .49).

These results clearly indicate that even though the ESN-M group is slightly older than the non ESN-M group, the two categories of children have significantly different means on all the subtests of the LET Battery, thus providing evidence of satisfactory construct validity for the LET Battery. It would appear that the LET Battery is the only test of its kind which offers such information. As will be discussed in the subsequent pages, the fact that the LET Battery has a desirable construct validity based on contrasted groups, adds to its usefulness for the purposes of deciding which children should be recommended for special educational facilities.

6.7 Predictive Validity

Also known as empirical validity, the major focus in predictive validity is with criterion prediction and not so much on "what the test measures" (Kerlinger, 1973, p.459; Cronbach, 1970), This type of validity also makes a useful contribution towards arriving at decisions in certain types of applied problems in the domain of education and psychology, e.g. for determining future educational needs of children, i.e. would their needs be best met in the mainstream or in special educational schools or units (Nunnally, 1978).

Central to the issue of predictive validity is the question of what is to be predicted. Sometimes it is quite difficult to find an appropriate measure of the attributes a test is supposed to index. More often than not with intelligence tests, aptitude tests and learning tests (such as those devised by Haynes, 1971 and Hegarty & Lucas, 1978) their predictive validity coefficient is computed by correlating them with academic tests (e.g. Reading, Maths, Spelling). A discussion pertaining to the difficulties associated with choosing an appropriate criterion for predictive validity can be found in most standard text books on psychometrics (e.g. Anastasi, 1976; Nunnally, 1978; Thorndike & Hagen, 1969) and therefore need not be stated here. However, suffice to note here that one of the problems to which predictive validity is sensitive to is "temporal changes" (Anastasi, 1976), This is particularly so if there has been a substantial time lapse between the administrations of the predictors and the criterion. Thus we can speak of concurrent and longitudinal predictive validity.

6.8 Concurrent Predictive Validity

The problem of "temporal changes" associated with predictive validity has been to a large extent surmounted in the present research. This has been achieved as follows. It will be recalled from the Methodology Chapter that the criterion tests, Schonell Graded Word Recognition Test (Schonell, 1951) and the Graded Arithmetic-Mathematics Test (Vernon & Miller, 1976), were given twice: once when the LET Battery and the IQ tests were administered in 1980 and again in 1981 after nearly three academic terms, to all the children except the Indian and ESN-M samples. Correlation coefficients between the reading test given in 1980 and 1981 and likewise between the maths test administered during the same time, were computed. These correlation coefficients computed separately for the two ethnic groups, Asian and English, and according to the age groups (Second Year Junior and First Year and Top Infants combined) are presented below.

Table 6.7 Correlation Coefficients Between Reading Ages and Maths Age Administered at Different Times According to Race and Age

	First Year & Top Infants				Second Year Junior			
	English (n=85)		Asian (n=90)		English (n=87)		Asian (n=123)	
	Reading 1980	Maths 1980	Reading 1980	Maths 1980	Reading 1980	Maths 1980	Reading 1980	Maths 1980
Reading 1981	.93	-	.92	-	.93	-	.89	-
Maths 1980	-	.84	-	.83	-	.83	-	.84

These high correlations between sequential administration of the reading and maths scores clearly suggest that it should not really matter whether the attainment tests administered in 1980 or 1981 were selected for the purposes of criterion. Since the use of the attainment tests administered in 1980 for the purposes of analysis overcame the limitations imposed by "temporal changes" and in no way reduced the information to be yielded by the validity coefficients, it was decided to use them rather than the scores of 1981. Further, the use of 1980 attainment scores offered an additional advantage too. As often happens with longitudinal studies, with the passage of time a certain number of subjects who are part of the original sample move from the area and it often becomes impractical and time consuming to trace them. In fact this happened with the present study. However, since the 1980 attainment scores were used this did not affect the total number as during the first phase of testing hardly any child moved from the area.

6.8.1 Prediction of Maths Age/Reading Age from the LET Battery

For the purposes of determining the predictive values of several independent variables one of the most useful techniques is Multiple

Regression, the details of which can be found in Kerlinger (1973, 1979) and Nie et al, (1975). Essentially what Multiple Regression analysis does is that it provides:-

> "...the magnitude of the relation between, on the one hand, the best possible combination of all independent variables, and on the other hand, the dependent variable." (Kerlinger, 1979, p. 1971)

The dependent variables were the Schonell Word Recognition Test (Schonell, 1951) and the Graded Arithmetic-Mathematics Test(Vernon & Miller, 1976). The independent variables were the subtests of the LETB, the Raven's Coloured Matrices and the Draw-a-Man test. The first question that needed to be answered was: were separate regression coefficients (weights) needed for the four age/race groups; i.e. did the relationships between the LETB and Maths Age and the LETB and Reading Age vary by race or age? The answer to this question, it was thought, would provide some evidence as to the culture fairness of the LET Battery in so far as Asian and English children were concerned as it is maintained that:-

> "An unbiased test should show similar correlations with other variables in the two or more populations. A test's predictive validity ... is the most crucial external indicator of bias. A significant group difference in validity coefficients would indicate bias." (Jensen, 1980 a, p. 328)

The answer to this question as well as the details of predictive validity of the LET Battery, together with the other conventional measures are provided in the following pages. The results for Maths and Reading will be discussed separately.

i) Prediction of Maths Age (Graded Arithmetic-Mathematics Test): Conventionally, "multiple regression" models are estimated by means of least-squares techniques. However, newer maximum-likelihood procedures (Joreskog, 1977), as implemented in such programmes as LISREL-IV (Joreskog & Sorbom, 1978), enable more complex models to be tested in a simpler fashion than is possible with conventional techniques. In particular, models can be set up for different groups of subjects and the coefficients constrained to equality across groups. This is particularly useful in answering the major question posed here, namely are there age/ethnicity differences in the predictive validities?

The Maths Age was predicted by using a "multiple regression" model with the criterion variable being the Graded Arithmetic-Mathematics Test and the predictors being Raven, Draw-a-Man and all LETB Subtests. Two models were estimated using the LISREL-IV programme: the first model allowed separate regression coefficients for all the four subgroups by age/race. Then a second model was tried with the constraint that the (unstandardized) regression weights had to be equal for all four subgroups. A test of this constraint yielded χ^2_{21} = 18.56. $p > 0.05$. Thus it appears that the null hypothesis of equal regression weights in all the four subgroups cannot be rejected.

The common regression weights found by the LISREL-IV programme for the pooled groups are given in Table 6.8; at the bottom of the Table are presented R^2 for the four groups by age/race.

Table 6.8 Standardized Regression Weights for the Pooled Groups and R^2 for the Four Age/Race Groups

Variables	Regression Weights
Raven	33
Draw	19
Piaget	17
VSMT	10
OPAT	(04)
WOPAT	14
SM	11

	1st Year & Top Inf.		Second Year Junior	
	English	Asian	English	Asian
R^2	62	58	52	54

Decimals omitted ( ) indicate a coefficient of less than 2 x its standard error.

Table 6.8 shows that the Raven receives the highest regression weight followed by the Draw-a-Man Test. The regression weight of the Piaget tests is quite close to that of the Draw-a-Man.

A further indication of the similarity of the regressions among the four groups can be seen from proportions of the variance explained. The (unweighted) average R^2 for the four groups only changes from 0.59 when four separate sets of regression weights are used to 0.57 when a single common set is used for the four groups. Thus it may be concluded:-

i) that the regressions are essentially the same for all four groups;

ii) that approximately 57% of the variance of Maths is explained by the combination of the two IQ measures and the LETB subtests.

The question which arises then is of the relative importance of the IQ tests in comparison to the LETB subtests. Using a common set of weights for all four subgroups, the (unweighted) average R^2 dropped from 0.57 to 0.41 using only the LETB subtests as predictors, and to 0.45 using only the IQ tests as predictors. Thus it can be seen that there is little difference in the predictive validity of the combination of the LETB subtests when compared to the conventional IQ measures - the latter perhaps being marginally better predictors.

ii) Schonell Graded Word Recognition Test: In order to determine the prediction of Reading essentially the same steps were followed as for the Maths ability (as measured by the Graded Arithmetic-Mathematics Test).

In the first place, a regression model was set up through LISREL-IV with the Schonell Graded Word Recognition Test as the criterion. The Raven's, Draw-a-Man Test and all the LETB subtests were predictors. As a first step separate regression coefficients for all the four subgroups by age and race were tried. This was followed by constraining all the regression coefficients to equality amongst all the four groups. A test of this constraint produced a χ^2 value of 20.48 with 21 df ($p>0.08$). There thus appears no evidence that different regression weights are needed. Therefore the hypothesis of equality of regression weights across the four subgroups must be accepted. Table 6.9 shows the single set of regression coefficients estimated by using LISREL-IV as being adequate for all four age/race groups. (Note: These have been standardized using the pooled within-groups standard deviations.) At the bottom of the table are given R^2 for the four groups by age/race.

Table 6.9 Regression Weights for the Pooled Groups and R^2 for the Age/Race Groups

Variables	Regression Weights
Raven	23
Draw	18
Piaget	09
VSMT	(08)
OPAT	(04)
WOPAT	15
SM	11

	1st Year & Top Inf.		Second Year Junior	
	English	Asian	English	Asian
R^2	33	54	29	35

Decimals omitted. () indicate coefficient is less than 2 x its standard error.

Further support to the idea that there is similarity of regression weights can be seen from the proportion of the variance explained. When four separate sets of regression weights are used for the four groups, the (unweighted) average R^2 is 0.43. Whilst when a single common set of regression weights is used then the (unweighted) average R^2 is 0.38. From this it may be concluded (and this conclusion is similar to the conclusion arrived at in the previous section while discussing the prediction of Maths ability):-

i) that the regressions are essentially the same for all four groups;

ii) that approximately 43% of the variance of Reading is explained by the two IQ tests and the LETB.

The key question at the heart of the analysis is: between the two types of measures (i.e. the two conventional IQ tests and the LETB subscales) which is more important for predicting reading? From Table 6.9 can be seen that the two IQ tests have more weighting compared to the weights of the subtests of the LET Battery. In fact the contribution of VSMT and OPAT is only minimal. However, from the fact that the two IQ tests contribute more to the prediction compared to the subtests of the LETB, it does not necessarily follow that they are more important. The <u>importance</u> of a variable in a prediction cannot be judged entirely by its relative weight. According to Kerlinger (1979):-

> "Does this mean that X_1 is really more important than X_2 scores [because the regression coefficient of X_1 is considerably greater than X_2] in the prediction? We cannot say clearly. The interpretation of regression weights is usually not simple and easy." (p. 169)

What is really important here is the fact that in predicitng reading ability three out of five LETB subtests make some contribution. Further, by employing a common set of regression weights for all the four sub-groups, the (unweighted) average R^2 drops from 0.38 to 0.30 using IQ tests alone and to 0.28 using the LETB subtests alone. These figures clearly suggest that really there is very little difference in the predictive validity of the LET Battery when compared to the two conventional IQ measures - the latter perhaps being only marginally better predictors.

The above results are based on and are applicable to typical children only. Would these results still hold if the sample on which the predictive powers of conventional IQ measures and the LETB subtests were compared, were mainly children who according to conventional assessment procedures (i.e. IQ measures and or teacher ratings) were in the bottom 5% of the population? This question has been considered in the next section entitled "Longitudinal Predictive Validity"

Before leaving this discussion a comment is warranted on the fact that the preceding regression analysis results show that the LETB alone or in combination with IQ measures is a better predictor of maths ability compared to reading ability. It would appear that this is on account of the relative contribution which the Piaget tests make since these, on the whole, tend to be better predictors of maths ability compared to reading (see Jensen, 1980 b, p. 674 where he reports a few studies which also demonstrate that on the whole Piagetian tests tend to have slightly higher correlations with maths tests compared to reading; Lunzer & Dolan, 1977). It would therefore seem that the variability in the predictive ability of the LETB for maths and reading is in line with some of the evidence in the literature.

Thus studies of predictive validity based on concurrent measures of predictors and criterion show that:-

i) there appear to be no age/ethnicity differences;

ii) the LETB is virtually as effective in predicting Maths and Reading Ages as are conventional IQ tests.

Thus the LETB has been demonstrated to have concurrent predictive validity.

6.9 Longitudinal Predictive Validity

Since Binet it has been commonly accepted that any individual who seems high on an intelligence test is a fast learner because he is good in basic subjects; whilst a child who performs poorly is a slow learner because of his retarded attainments in basic subjects (Guilford, 1967). For many years, therefore, several psychologists have viewed intelligence as an integral part of the "ability to learn" (Jensen, 1979).

Contrary to this commonly held view of equating intelligence with learning ability several workers have found that children with low IQ scores when tested on learning tasks show far more variability than would be expected from their low IQ scores (Feurstein, 1970, 1979; Haywood et al., 1975; Hegarty & Lucas, 1978; Jensen, 1961, 1963, 1967, 1969). For instance Jensen (1963) in one of his studies found that the highest scores on learning tasks were achieved by children whose IQ scores were 147 and 65 respectively. Similar results, perhaps though not as "dramatic", were found by Hegarty & Lucas (1978). The pilot study (details reported on pages 82 - 103) also demonstrated that there were nearly 24% children whose scores on the Piagetian type tasks and other learning tasks were comparable with children with average and high IQ scores, but their own IQ scores were at or below the 5th percentile on the Raven's Matrices (using norms reported in the Manual for the Coloured Progressive Matrices (Raven, 1962)).

Interesting and important though these findings are, none of these workers would appear to have studied these children (i.e. those who score low on an IQ test but high on learning tasks) as a group on a longitudinal basis to determine empirically the validity of their learning tasks. This study was set up to provide such information. It was hypothesised that low IQ scores and/or low rating of these childrens' ability by their teachers, are not necessarily a dependable index of these childrens' rate of learning in school. For this type of sample - the details will follow shortly - the Learning Efficiency Test Battery should be a better predictor of their learning ability compared to the conventional type of assessment. It was also hoped that this study would shed some light on the theoretical issue of the relationship between IQ and learning ability.

6.9.1 Method

Sample: The sample that has been used for this longitudinal study is a small sub-sample of 27 children of both sexes and of both races. They were drawn from the main sample of Second Year Junior children. They had to meet any one or more than one of the following conditions:

i) That the IQ score on the Draw-a-Man Test should be at or below 70.

ii) That the scores on the Raven's Matrices should be at or below the score given in the Manual as corresponding to the 5th percentile - range defined by Ravens as "mentally defective".

iii) That the child, according to his teacher's estimation of his ability, should have been rated as Grade 5 (i.e. in need of special education).

These cut off points were selected because quite often children who are in this range are recommended for special educational provision. It is unlikely that this criterion is going to change as a result of the new Education Act 1981 (DES 1983) not withstanding that it purports to repeal the:-

> "...provisions of the Education Acts relating to special educational treatment and establishes a new framework for the education of children requiring special education provision whether in special or ordinary schools."
> (DES, 1983, p. 1)

The other notable feature of this sample was that a very large majority of these children (81.5% n = 22) had high scores on all the subtests of the LET Battery. There were only a few children

(18.5% n = 5) who did not achieve as well on the LETB subtests as the rest of the children. In other words, their performance on the LET Battery, more or less, corresponded with their low performance on the conventional assessment as well. In the main, this sample consisted of children who had relatively high scores on the LET Battery but were rather low (i.e. in the bottom 5% of the population) on conventional assessment. A breakdown of this sample is as follows:-

Table 6.10 Distribution of Children as a Function of Race and Sex

	Boys	Girls
English (n = 6)	3	3
Asian (n = 21)	10	11

All these children received a variety of remedial help from their schools during the time of the longitudinal study.

Task and Materials: Details of tasks and materials, administration and scoring procedures followed have been provided in Chapters 4 and 5 and Appendix A. Suffice to say here, since this group was part of the main sample therefore like the main sample scores of these children on the following tests were available.

1) Raven's Matrices: instead of the raw scores, estimated IQ scores were used for all the analyses.
2) Draw-a-Man
3) Schonell Graded Word Recognition Test. This test was administered twice. Once in March 1980 and then later in December 1980.*

* The gain scores for reading were obtained by deducting the March 1980 scores from the December 1980 scores for each child.

6.9.2 Results

The mean scores and standard deviations of the two teacher's ratings, two IQ tests, the LETB subscales together with their chronological ages are presented in Table 6.11.

Table 6.11 Means and SDS of the Subgroup and the Parent Population

	Subgroup		Parent Population	
	M	SD	M	SD
Chronological Age in Months	106.78	3.48	107.09	3.20
Teacher's Rating According to Attainment	4.19	0.96	3.23	0.96
Teacher's Rating According to Potential	3.63	1.08	2.96	0.93
Raven	13.85	3.55	22.49	6.46
Draw-a-Man	86.93	12.88	Not Available	
Piaget	8.48	3.80	10.02	2.47
VSMT	16.89	7.17	18.98	5.91
OPAT	29.60	10.99	32.17	7.29
WOPAT	20.63	10.73	23.25	9.56
SM	21.41	8.05	23.35	4.25
% Asian	78	-	59	-

It is interesting to note here that the children in the subgroup have considerably higher mean scores on the Draw-a-Man compared to their Raven's and Teacher's Ratings (it will be recalled from the Method Chapter that the Rating of 4 was described as "Definitely Below Average (DULL)). The scores of the subgroup on the Raven and Teacher's Ratings place these children substantially below a typical group in their intellectual ability as measured by the two conventional measures. It is also interesting to observe that the mean scores of these children on all the subtests of the LETB are fairly high. As a matter of fact their mean scores on the LETB subtests compare very favourably with the mean scores of the parent population which had children of a wide range of abilities. For the purposes of comparison the corresponding mean scores of the parent group have also been shown in Table 6.11. Although the mean differences of the corresponding LETB scores of these two groups would appear only marginal, the percentage difference of Asian children in the two groups is quite substantial. What is really striking is the fact that there are almost four times more Asian children (78%) compared to their English counterparts (22%) in the subgroup whose scores on the conventional type of assessment range in the bottom 5% of the population. From the above mean scores of the subsample it would seem apt to describe them as low on conventional assessment but high on the LETB subtests. From this it follows that it is not essential that a child who scores low on conventional assessment would also score low on the LETB subtests.

6.9.3 Comparison of the Conventional Assessment Measures and the LET Battery as Predictors of Reading

It was thought that this issue would be best handled by correlation and stepwise regression analysis. In order to obtain these statistics the appropriate SPSS procedures were used (for details of these programmes see Nie et al., 1975). The consistent picture that has emerged from these two approaches is that the Learning Efficiency Test Battery is a better predictor of the subsequent learning ability (i.e. reading gains) as well as Reading.

6.9.4 Evidence from Correlations

Table 6.12 shows how the reading test scores correlate with the two conventional IQ tests and the LETB scores. (For full details of this Matrix see Appendix B, Table B.1.) From Table 6.12 can also be seen that the TRP and TRA have the highest correlations with both READAGEM (r = .66 and .64 respectively; both significant at the 1% level) and READAGED (r = .71 and .65, significant at the 1% level).

Table 6.12 Correlation Matrix of Two IQ Tests, Teacher's Ratings, Reading Tests and Reading Gains.

	READAGEM	READAGED	READGAIN
1) Raven IQ	28	31	11
2) Draw-a-Man	33	48**	32
3) TRA	64**	71**	25
4) TRP	66**	65**	13
5) Piaget	38*	56**	38*
6) SM	53**	60**	24
7) OPAT	32	53**	41*
8) WOPAT	13	37*	42*
9) VSMT	34	56**	43*
10) READAGEM	X	79**	-12
11) READAGED	X	X	52**
12) READGAIN	X	X	X

Decimals Omitted

* Significant at the 5% Level

** Significant at the 1% Level

Out of the five LETB subtests two subtests, Piaget and SM, have substantial correlations with READAGEM. None of the conventional IQ tests have a significant correlation with READAGEM. Of the two IQ tests it is only Draw-a-Man which has a significant relationship with READAGED (r = .48 significant at the 1% level). Barring WOPAT which has a significant correlation at the 5% level,the rest of the LETB subtests have significant correlations with READAGED at the 1% level. Except for the WOPAT which has a marginally lower correlation with READAGED compared to Draw-a-Man, the rest of the subtests have better correlations relative to the two IQ tests.

The picture changes quite dramatically when the correlations amongst all the predictors and reading gain are examined. TRA and TRP do not correlate with READGAIN despite correlating significantly with the initial and final reading scores. Neither do the two conventional IQ tests have a significant correlation with READGAIN: in fact the Raven IQ has the lowest correlation in relation to the rest of the variables. Except for SM, the rest of the LETB subtests have significant correlations with READGAIN - all are significant at the 5% level.

The foregoing correlations show that in so far as READAGEM and READAGED are concerned neither the LET Battery nor the conventional IQ tests have such high correlations with them as TRA and TRP. Does this mean that the other variables compared to TRA and TRP are redundant or less useful either for the purposes of concurrent validity or for predicting these children's reading ability? The answer must lie in the negative. At a superficial level it would appear that TRP and TRA are capable of providing the best estimate of concurrent validity, being the best predictors of reading ability compared to the rest of the variables. However when examined closely it becomes quite plain that the superiority of these two variables in relation to the rest, though impressive, is of least practical use.

It will be recalled from Chapter 4, Section 4.2, that teachers were asked to provide two types of rating of children: one based on children's attainments and the other based on their potential. Since one of the ratings (TRA) was based on the attainments high correlations between TRA, READAGEM and READAGED were to be expected. In a way the correlations between TRA, READAGEM and READAGED are not as high as they should be - the expectation would be near .9 or so and not just .64 or .71 - though both are significant. This shows that teachers' ability to rate children despite their knowledge of their attainments is not really as good as it is popularly believed to be. Further, the high correlation between TRA and TRP (.74 significant at the 1% level; for other details see Appendix B, Table B.1) also explains why TRP too has a higher correlation with READAGEM and READAGED compared to the subtests of the LET Battery and the conventional IQ tests. Thus it would seem the reason that TRA and TRP have better correlations as opposed to the LET Battery or conventional IQ tests with READAGEM and READAGED, is not because the teacher's ability to rate children is better than the LET Battery, but largely due to the artefact of rating.

Even if it was not an artefact and the teachers were in reality better at estimating children's ability to profit from instruction compared to the LET Battery, there are times when teachers cannot provide a measure of a child's ability on the basis of their judgement. An example of such times is when a teacher has not known a child over a long period of time but there is some urgency to ascertain his ability to learn. This could happen in the case of a recently arrived child into this country from abroad - particularly if he/she happens to come from a non-English speaking country. At such times there is a need for an objective test which can yield such information and does not suffer from the constraints of time or language. From the foregoing correlations it would seem that such a need is better met by the LET Battery as against the teacher's rating or the conventional IQ tests.

In conclusion these results quite plainly indicate that the LET Battery is more useful for predicting reading compared to the conventional type of assessment and definitely better for predicting changes in reading levels over a period of time.

6.9.5 Stepwise Multiple Regression Analysis

Stepwise Multiple Regression Analysis was chosen as a second stage of analysis as this seems to be the most appropriate statistical procedure for the proplem at hand. The problem being to determine the relative importance of the two conventional IQ tests, two Teacher's Ratings and the LETB subtests in predicting reading gain. Stepwise Multiple Regression allows the computer to first:-

> "...enter variables in single steps from best to worse provided that they meet the statistical criterion established in the parameters section of the statement. The variable that explains the greatest amount of variance in the dependent variable will enter first; the variable that explains the greatest amount of variance in conjunction with the first will enter second and so on". (Nie <u>et al</u>., 1975, p. 345)

The results of Stepwise Multiple Regression Analysis using READGAIN as the dependent variable are presented in Table 6.13.

Table 6.13 <u>Stepwise Regression Analysis: Using Gain Scores as the Dependent Variable</u>

<u>Dependent Variable: Reading Age Gain in Months</u>

Variable	Multiple R	R Square	RSQ Change	Significance
VSMT	425	180	180	*
WOPAT	496	246	065	ns
DRAW-A-MAN	544	295	049	ns
SM	568	322	026	ns
TRP	580	337	014	ns
OPAT	585	343	006	ns
PIAGET	587	344	001	ns
TRA	587	345	000	ns

* Significant at the 5% Level.

Decimal Points omitted.

From this table it can be seen that the test that explains the maximum amount of variation in the dependent variable, READGAIN, is VSMT, a subtest from the LET Battery. (F ratio of 5.52 with 1,25 df significant at the 5% level.) No further steps are significant; however it may be of interest that a further subtest from the LET Battery, i.e. WOPAT, was selected as it explains the next greatest amount of variance in conjunction with the VSMT: jointly they account for nearly 25% of the variance. None of the conventional measures were selected.

Regression or other analyses using gain scores have been criticised in the literature (Ferguson, 1971; Thorndike & Hagen, 1969; Vernon, 1969; see also Anderson, 1940 and Bloom, 1964 cited in Vernon, 1969). The alternative approach is to seek prediction of the final score, so called covariance analysis. The results of such an analysis are shown in Table 6.14.

Table 6.14 Dependent Variable: Reading Age in Months in December 1980 (Covariance Approach)

Dependent Variable: Reading Age in Months in December 1980

Variable	Multiple R	R Square	RSQ Change	Significance
READAGEM	785	617	617	**
VSMT	844	712	095	**
DRAW-A-MAN	860	740	027	ns
WOPAT	878	771	031	ns
TRA	880	775	003	ns
OPAT	882	778	003	ns
RAVEN IQ	884	782	004	ns
PIAGET	885	784	001	ns
SM	886	785	000	ns

** Significant at the 1% Level

Decimal Points omitted.

As can be seen in Table 6.14 it is again a subtest from the LET Battery, namely, the VSMT that is the main contributor in explaining the variance of the dependent variable, READAGED (F ratio of 7.99 with 1,24 df significant at the 1% level). Again no further steps are significant. The only difference between the two regression analyses is that here the second best predictor in conjunction with the VSMT is DRAW IQ followed by the WOPAT, whereas in predicting the READGAIN, DRAW IQ takes the third place as the best predictor. From these two regression analyses it can be concluded that the LET Battery, or at least one of its subtests, in the main exceeds the conventional measures of assessment in predicting gains in Reading, whether these are measured directly by the variable READGAIN or indirectly through the covariance approach.

To summarise, the subtests of the LETB are homogeneous in content and they have stable and satisfactory alphas and test/re-test reliabilities. In addition the LETB has also acceptable face/content, construct and predictive validities.

CHAPTER 7

RESULTS II: TESTING JENSEN'S HYPOTHESES

7.1 Jensen's Hypotheses of Mental Abilities and Factorial Structure of the LET Battery

One of the aims of the present enquiry, related to Jensen's hypotheses of mental abilities, was to test the factorial structure of the LET Battery to determine whether the various subtests of the LET Battery would form two identifiable orthogonal or near orthogonal clusters as tests of Level I and Level II abilities. The three learning tasks, Objective Picture Association Test, Word Object Picture Association Test and Symbol Manipulation Test, were rooted in the theory of associative learning, and together with the Visual Sequential Short Term Memory Test, should be tests of Level I ability, according to Jensen's hypothesis: the Piaget based test would be a measure of Level II ability (e.g. Jensen, 1969, 1973, 1974 a, 1980 a, 1980 b).

Since a hypothesis of this nature is best handled by factor analysis it was therefore decided to factor analyse the LET Battery together with the Raven's Matrices and Draw-a-Man Tests. The Raven and Draw-a-Man were added on the assumption of Jensen's theory (1969) that they, together with other conventional IQ measures, would also be measures of Level II ability.

On the basis of Jensen's hypothesis the expectation would be that the three learning tests together with the Visual Sequential Short Term Memory (all measures of Level I ability) should form a separate orthogonal cluster from the Raven's Matrices, Draw-a-Man and the Piaget based tests (which are all measures of Level II ability). Since in the development of the various subtests great attention was paid to ensure that these do not favour either English or Asian children (see Principles Section 3.2 Chapter 3, Rationale Behind the Pilot Version of the LETB) it was therefore predicted that there would not be any significant differences in the factorial structure of English and Asian Children.

Again this is in line with Jensen's findings, namely, that although there are differences in means for Level I and Level II between populations, the factorial structure is similar (Jensen, 1973).

7.2 Factorial Structure of the LET Battery and other IQ Measures

As an exploratory step, in order to examine the factorial structure of the LET Battery and the other measures of IQ according to race and age, it was decided to carry out a Varimax factor analysis with a restriction to two factors. (All analyses used the appropriate SPSS procedure, see Nie et al., 1975). Because of the major interest in the factorial structure of the two racial groups and subsidiary interest in age, four separate factor analyses were carried out. Because of the small numbers, Top Infants and First Year Juniors were combined. They were:-

1) Varimax factor analysis for Second Year English children only (n=86).

2) Variamx factor analysis for Second Year Asian children only (n=123).

3) Varimax factor analysis for Top Infants and First Year English children (n=85).

4) Varimax factor analysis for Top Infants and First Year Asian children only (n=90).

As it was simply an exploratory factor analysis, therefore only the correlation coefficients between the variables and plots of rotated factors for each race and age group were examined.

7.2.1 Plots of Rotated Factors

Examination of the plots of rotated factors (see Figures 7.1, 7.2, 7.3 and 7.4) show that there were certain similarities amongst the four groups (Age/Race; Second Year Junior children will be described as older; First Year Junior and Top Infants combined will be referred to

Figure 7.1 Factor Plots of Second Year Junior English Children

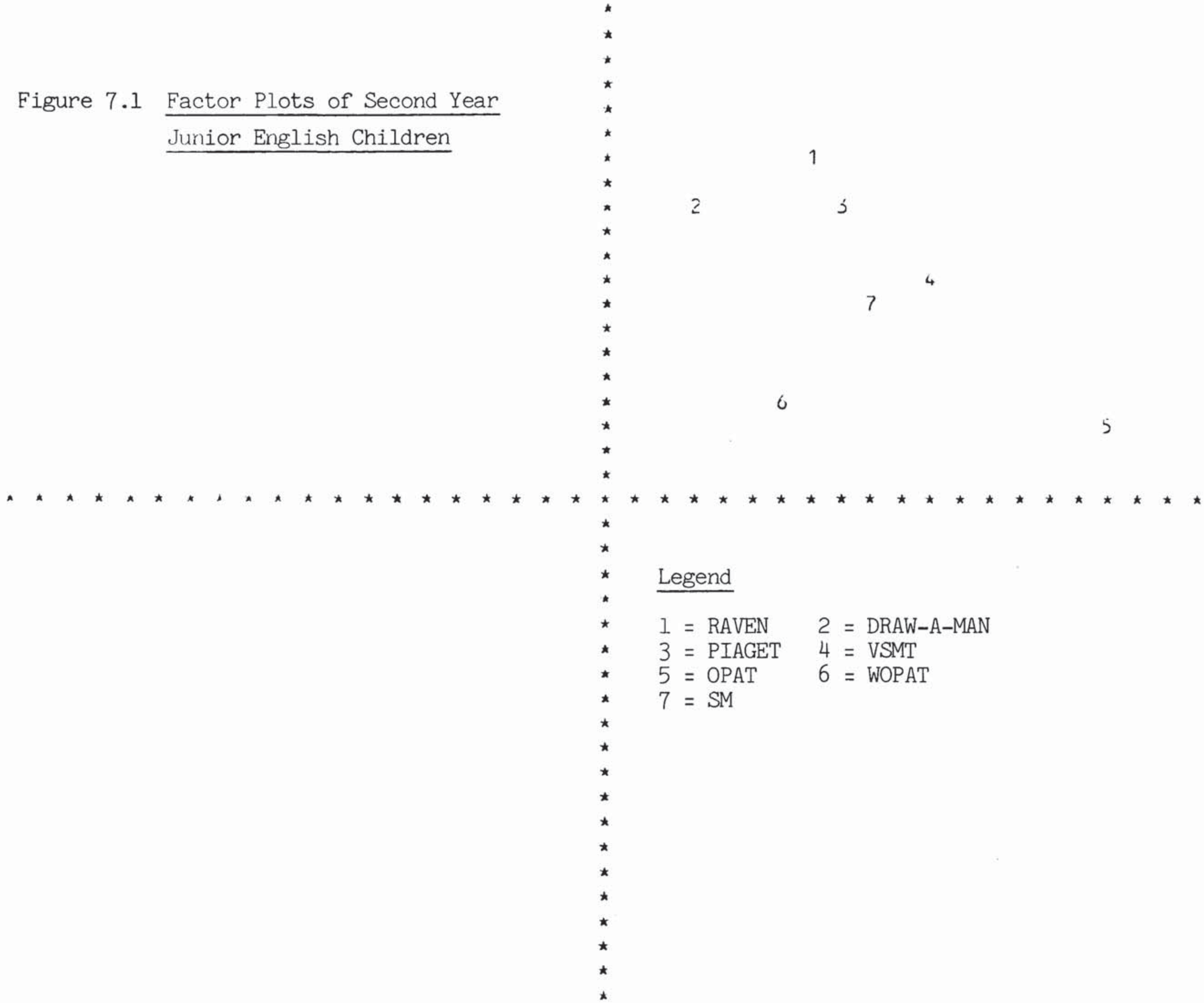


Figure 7.2. Factor Plots of Second Year Junior Asian Children

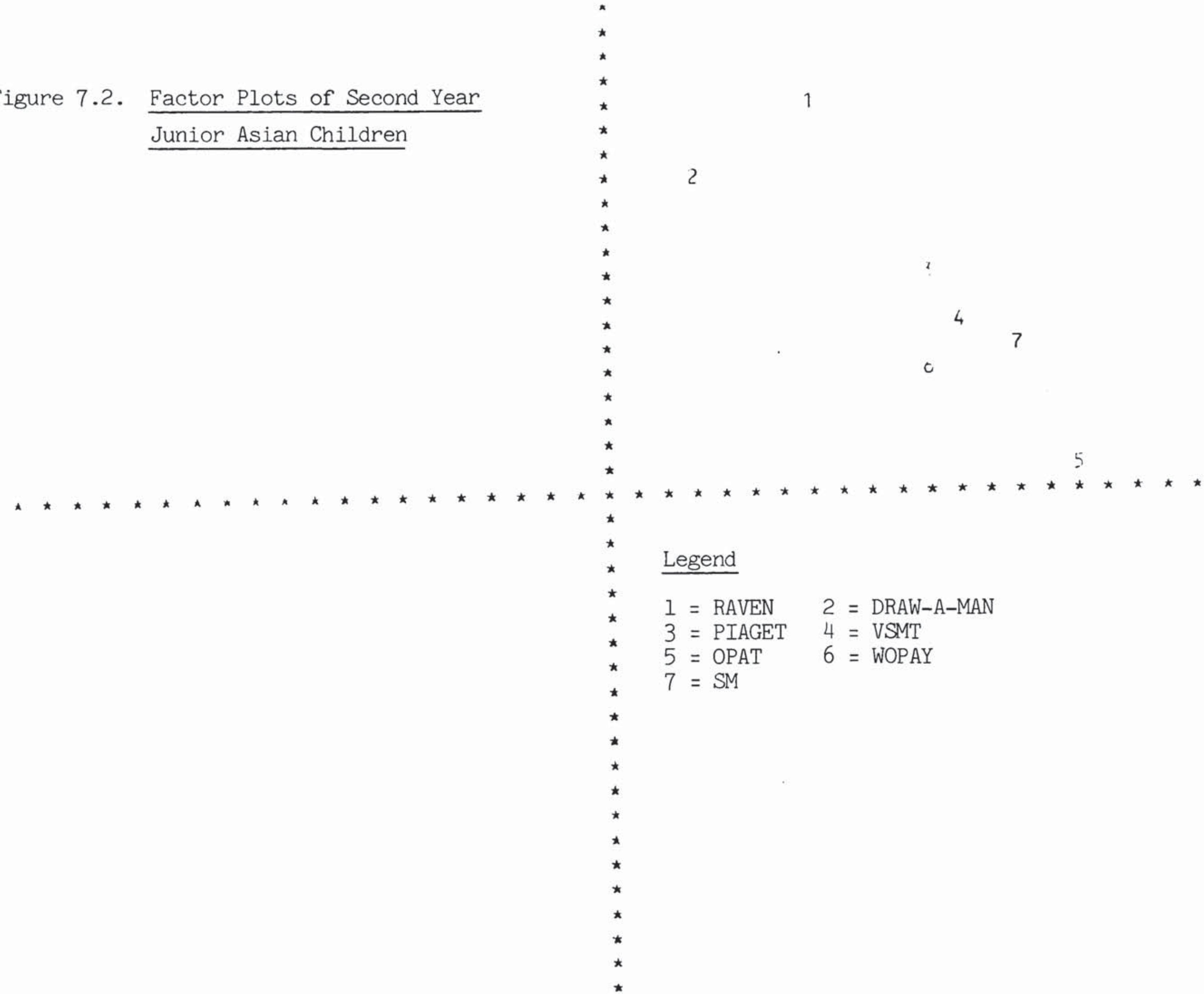


Figure 7.3. Factor Plots of Top Infants And First Year Junior (Combined) English Children

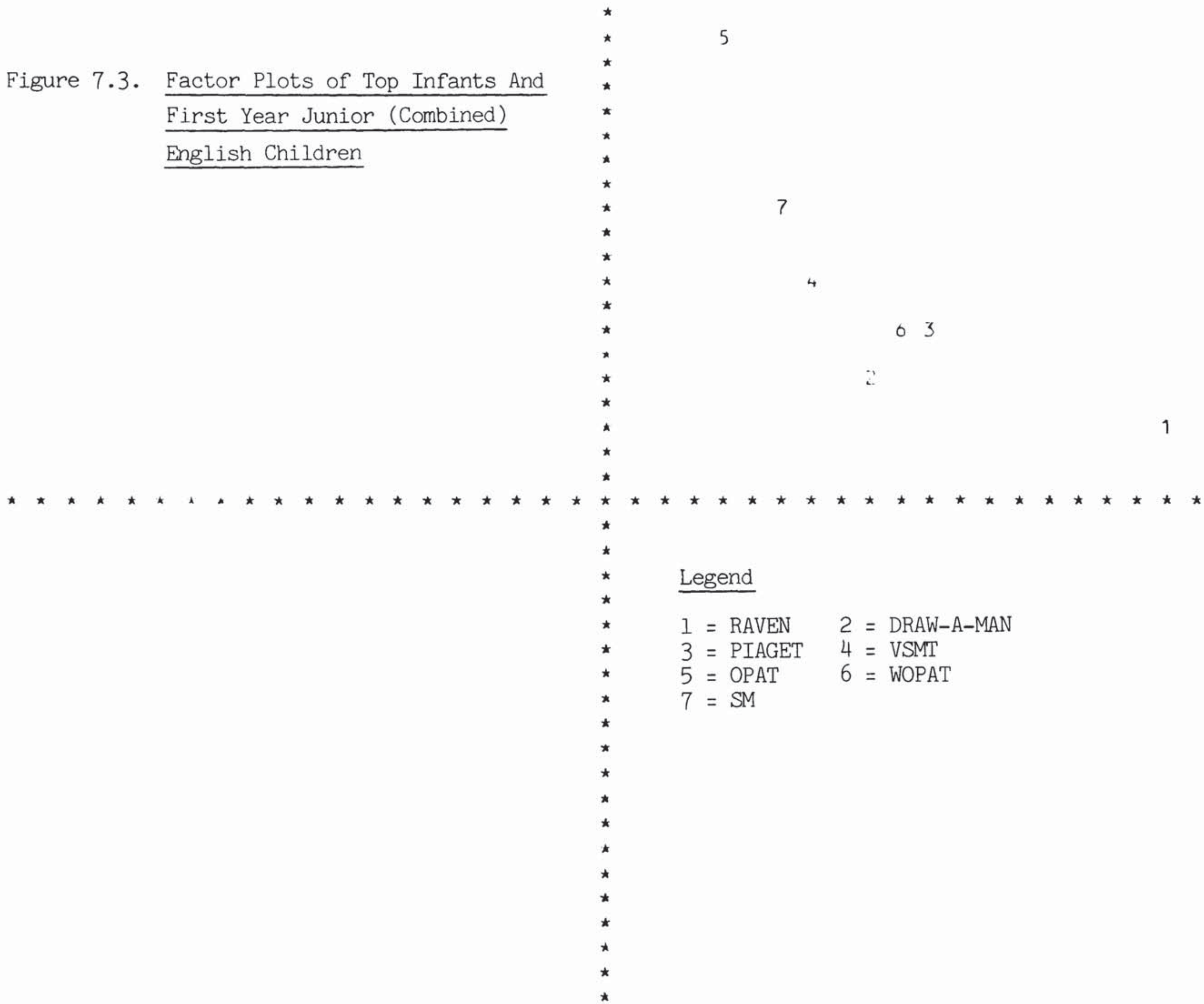
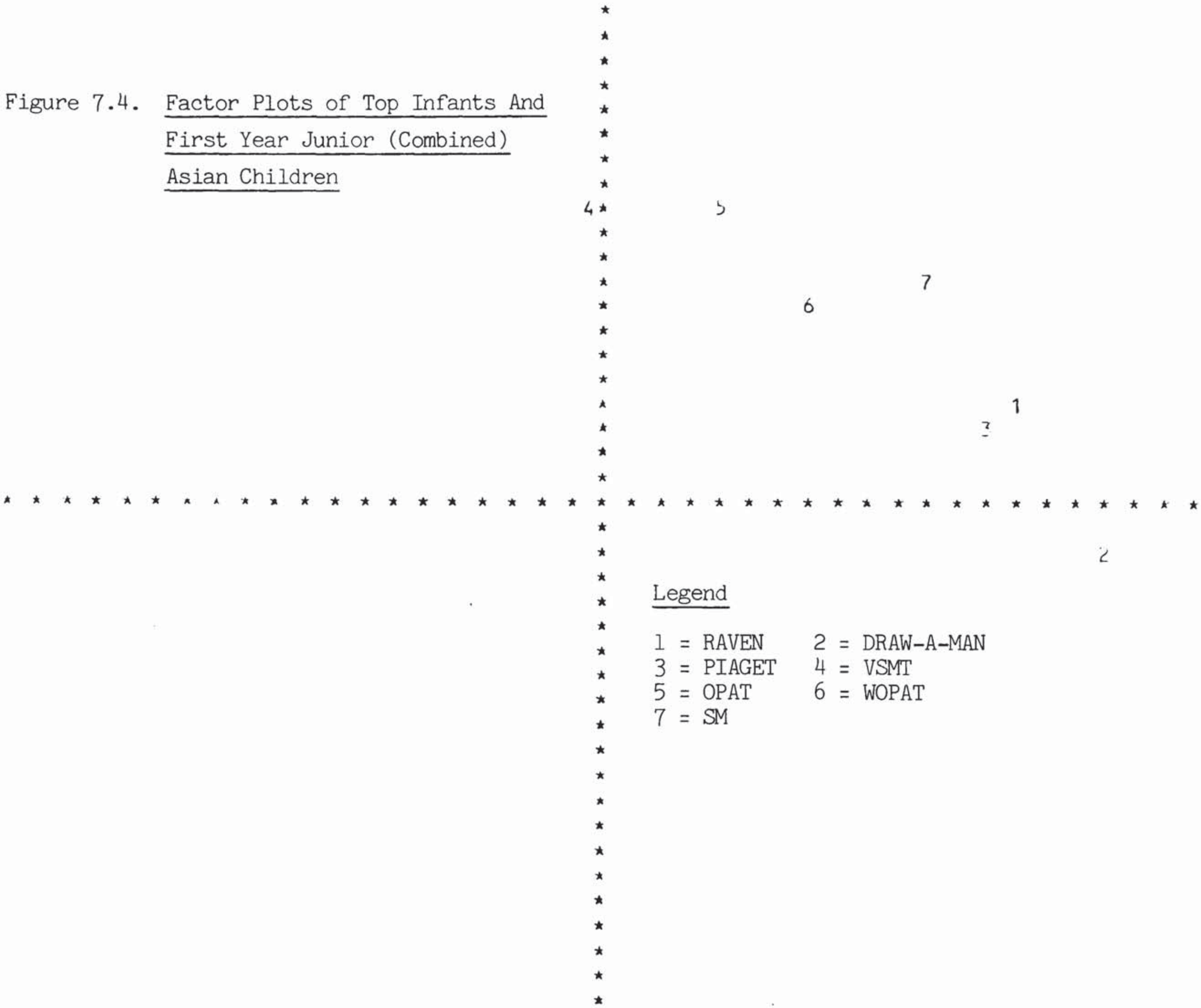


Figure 7.4. Factor Plots of Top Infants And First Year Junior (Combined) Asian Children



as younger group) in their plots of rotated factors. As a matter of fact some similarities between the older English and Asian, and the younger English and Asian were quite marked. First, take for instance, the older group (Figures 7.1 and 7.2): Raven's, Draw-a-Man, Visual Sequential Short Term Memory and the other learning tests would appear to be more or less in the same factor spaces for English as well as Asian children. Now consider the younger group: Visual inspection would suggest that the Raven, Symbol Manipulation and the Object Picture Association Tests for both English and Asian children were, in the main, in the similar factor spaces.

However certain dissimilarities were also discernible in the four groups. In the case of older Asian children, the Piaget based test was closer to the learning tests whilst in the case of English children it was closer to the Raven and Draw-a-Man Tests.

These minor differences in the positioning of some of the subtests in the factor spaces were apparent in the younger group too. In the case of English children (Figure 7.3) the Word Object Picture Association Test and the Symbol Manipulation and the Short Term Memory Tests were closer to the Draw-a-Man Test and the Piaget based Test rather than the Raven's. Whilst this was not the case with respect to Asian children: in their case, the Raven and Piaget would seem to be closely related to each other; the Draw-a-Man considerably away from the learning and Short Memory Tests (see Figure 7.4).

Notwithstanding the similarities and minor dissimilarities of the various tests in the factor spaces certainly there were no two clear cut orthogonal or near orthogonal clusters (in any of the four groups) with learning tests and the Short Term Memory test forming one cluster and conventional IQ tests and Piaget based tests forming a separate cluster. Inspection of the correlation tables (Tables 7.1 & 7.2) quite clearly suggests that in the case of all the groups measures of learning and short term memory were substantially related to the measures of IQ and Piaget based Tests. As a matter of fact some of the correlations between learning tasks and Visual Sequential Short Term Memory (Level I according to Jensen) were greater with the measures of Level II ability (i.e. Raven, Draw-a-Man and Piaget, according to Jensen) than some of the correlations amongst the tests which are supposedly measures of Level II ability. Consider the older Asian children (Tables 7.1) ; the correlation coefficients between the VSMT and the Raven (.43 significant at the 1% level), between the Symbol Manipulation and the Raven (.44 significant at the 1% level) and between the Symbol Manipulation and the Piaget (.45 significant at the 1% level) were in fact substantially greater than the correlation between the Draw-a-Man and the Piaget (.40 significant at the 1% level, see Table 7.2). This was true in the case of English children too: for instance, the correlation between the VSMT and the Raven (.47 significant at the 1% level) was higher than the correlations between the Raven and the Draw-a-Man (.41 significant at the 1% level), and the Draw-a-Man and Piaget (.40 significant at the 1% level) and was almost of the same magnitude as between the Raven and the Piaget (.49 significant at the 1% Level).

It was true in the case of the younger children as well that the correlations of the learning measures and the short term memory with the measures of IQ and the Piaget based tests were fairly substantial. Of the 24 correlation coefficients in Table 7.1 (younger group) there are fourteen which are significant at the 1% level and four at the 5% level; there are only six correlations which are not significant as

Table 7.1 Correlations Between Measures of Learning Tests, Short Term Memory and of Cognitive Ability (Including Piaget Based).

	YOUNGER GROUP						OLDER GROUP					
	English (n=85)			Asian (n=90)			English (n=86)			Asian (n=123)		
	Raven	Draw	Piaget	Raven	Draw	Piaget	Raven	Draw	Piaget	Raven	Draw	Piaget
VSMT	32**	24**	38**	09	-10	03	47**	22**	40**	43**	27*	43**
OPAT	29**	35**	37**	23*	07	12	32**	16	42**	28**	14	42
WOPAT	49**	16	40**	28**	26*	32**	19	13	23*	33**	22**	40**
SM	34**	23*	41**	36**	41**	44**	35**	26**	38**	44**	28**	45**

Note: Decimals omitted
* Significant at the 5% Level
** Significant at the 1% Level

Table 7.2 Correlation Coefficients Between Raven's, Draw-a-Man and Piaget

	YOUNGER GROUP						OLDER GROUP					
	English (n=95)			Asian (n=90)			English (n=86)			Asian (n=123)		
	Raven	Draw	Piaget	Raven	Draw	Piaget	Raven	Draw	Piaget	Raven	Draw	Piaget
Raven	-	49**	51**	-	56**	44**	-	41**	49**	-	55**	47**
Draw	-	-	31**	-	-	45**	-	-	40**	-	-	33**
Piaget	-	-	-	-	-	-	-	-	-	-	-	-

Decimals omitted
** Significant at the 1% Level

would be expected on the basis of Jensen's theory. Further, some of the correlations between the tests of Level I and Level II (as they would be described within Jensen's theoretical framework) were greater or almost as great as among the measures of Level II ability.

The results cannot be regarded as consistent with Jensen's hypothesis with respect to all the four groups, namely that the tests of associative type and short term memory load on a separate and orthogonal factor from the tests of conceptual ability. These results demonstrated that, on the contrary, the learning, short term memory and cognitive tasks were not easily divisible into two non-related identifiable groups. In some other studies too, which have some similarities to the present study, it has been found that learning and memory tests are not orthogonal to the conceptual tasks (Haynes, 1971). Some of the evidence available would also suggest that when only learning tasks are factor analysed they load on more than one factor (Malmi et al., 1979 and a study cited therein by Underwood et al., 1978) - contrary to Jensen's hypothesis.

Cronbach (1970) too, after a survey of a considerable amount of literature concerned with the relationship between intelligence and learning ability, came to the conclusion that there is not just one kind of learning ability and some individuals could excell in one type of learning task but could be poor on others. The inference from Cronbach's conclusion does not suggest that all the learning tasks if factor analysed would load on one factor only. The implication is that they may load on more than one factor (cf. Malmi et al., 1979 and Underwood et al., 1978 cited therein).

Guilford (1967) after surveying the earlier experimental work concerned with the relationship between intelligence and learning ability arrived at the conclusion that there was no general learning ability underlying all learning tasks and:-

> "...both learning ability and intelligence involve many different component abilities and that they share the same components depending upon the nature of the learning task and intelligence test." (p. 20)

The factor analysis results of the present study would appear to be more in line with Rohwer et al. (1971 and the several studies cited therein) that the paired associate learning involves a substantial amount of cognitive processes, perhaps not easily distinguishable from the cognitive processes required in what Jensen describes as conceptual learning tasks. (See also Gagne, 1977 who warns that the simple associative type of learning tests may not after all be that simple.) Cognitive psychologists, too, claim that in the so called simple S-R or simple associative type of learning tasks some of the mental processess that are invoked are: observation of the association, storing of this information in memory and the retrieval of it (Hilgard et al., 1979). It would seem that because some of the mental procesess involved in both Level I and Level II type of learning are more or less the same, therefore these two types of abilities do not form two easily divisible and separate groups when factor analysed. The psychological reality would appear to be that the wide range of measures of mental processes tend to be correlated with each other rather than forming two orthogonal groups (Lunzer & Dolan, 1977, see also Stevensen, 1972 who found correlations ranging from .38 to .56 between measures of Level I and Level II abilities).

Although in the light of the foregoing exploratory factor analysis, Jensen's hypothesis that mental abilities can be categorised into Level I and Level II abilities cannot be accepted, this analysis, however, did not succinctly resolve the problem as to whether the Asian and English have more or less the same type of factorial structure.

Notwithstanding this, however, it needs to be added that there are strong indications that it is unlikely that there will be any major differences in the factorial structure of the two groups although, at this stage, not much reliance can be placed in this assertion. The chief difficulty being that except for Second Year group (n=123) the size of the remaining three groups is in their eighties only. It is well recognised that with small numbers, such as these, the standard errors associated with factor loadings do not permit sufficient localisation of the tests in the factor space to be sure whether the apparent differences, or lack of differences, are statistically significant. Therefore, in order to test these differences between Asian and English children, both in the older group (Second Year Junior) and in the younger group (Top Infants and First Year Junior), some confirmatory rather than exploratory factor analysis was called for.

7.2.2 Confirmatory Factor Analysis

It was decided to use the LISREL Programme (Joreskog, 1971, see also Footnote) as it is possible with this programme to test hypotheses about the equality of factor loadings and/or factor covariances among groups. A series of models were set up and tested for their goodness of fit as shown in Table 7.3.

Grateful thanks to Dr P Coxhead for carrying out job preparation for LISREL analysis and for computing as well which made inordinate demands on his time.

Table 7.3 Summary Results of the LISREL Programme

Model	No of Factors	Equality Constraints	Chi Sqr	df	sig
A	1	None ie separate loadings, factor covariance, unique covariance.	125.12	56	***
B	2	None ie separate loadings, factor covariance, unique covariance.	26.18	32	ns
C	2	Loadings equal across four groups.	77.58	62	ns
D	2	Loadings and factor covariances equal across four groups.	89.11	71	ns
E	2	Loadings and factor covariances and unique variances across four groups.	187.20	92	***

*** = $p < .001$
** = $p < .01$
ns = $p < .05$

First of all there was a need to determine whether a two factor solution was appropriate or whether a different number of factors could be required for all the subtests of the LETB plus the Raven's Matrices and Draw-a-Man Tests. In order to test this hypothesis, firstly, Model A was set up. This model stipulates that there be one factor in each of the four age/race groups with no restriction that the factor loadings, factor covariances or unique covariances should be the same across the groups. Model A clearly did not fit the data as shown by the Chi Square of 125.12 with 56 df ($p < .001$).

Secondly Model B was set up, in which there are two factors in each age/race group, again with no equality contraints across groups. As can be seen from Table 7.3 Model B fits the data extremely well (Chi Square = 26.18 df = 32; $p > .05$). Clearly, then one factor is not adequate but neither are more than two required. Hence the hypothesis of two factors only underlying earlier discussion, is confirmed.

Given that the correct number of factors appear to be two, the next question relates to equalities between the four age/race groups. Models C to E successively equate factor loadings, factor covariances and unique variances across the four groups. Table 7.3 shows that the best of these three models is Model D i.e. loadings and factor covariances equal across groups but separate unique variances for each group. Table 7.4 shows the loadings and factor correlations for Model D after (i) a Varimax rotation and (ii) a Promax rotation.

Table 7.4 Varimax and Promax Factor Loadings of the Four Groups (Age/Race)

	Loadings Varimax Rotation		Loadings Promax Rotation	
	1	2	1	2
Raven	78	25	90	-12
Draw	59	15	70	-14
Piaget	49	39	45	22
WOPAT	29	41	19	35
VSMT	26	52	09	52
SM	22	50	05	51
OPAT	12	76	-21	90

Factor Correlation		
Factors		
1	-	69
2	-	-

Note: Decimals omitted
Salients underlined

The intepretation placed on the loadings would appear to be similar between the Varimax and Oblique solutions. It would also appear that by the Oblique solution the factorial complexity has been somewhat reduced and the obtained solution has become slightly neater, i.e. loadings closer to 1 or 0. For instance, in the case of OPAT, the factor loadings of .12 and .76 in the Varimax solution changes to -.21 and .90 in the Oblique solution. This applies to nearly all the factor loadings, i.e. they have come closer to 1 or 0. Superficially the Promax rotation would suggest that there is substantial support to Jensen's hypothesis that the mental abilities can be viewed as Level I and Level II. All the learning tests and the Short Term Memory Test (measures of Level I according to Jensen) have meaningful loadings (see Child, 1973) on Factor 2, whilst the Raven, Draw-a-Man and Piaget Tests load minimally, the converse is true for Factor I. However, in the light of the correlation between factor I and II ($r = .69$) the inevitable conclusion has to be that there is little evidence to suport Jensen's fundamental thesis that the range of mental abilities fall into a clearcut pattern of Level I and Level II abilities (depending upon the mental manipulation of the sensory input) which are orthogonal to each other. In view of the equalities across the groups, this is true in the case of younger as well as older children, both Asian and English. As it has been stated earlier, the conclusion from these factor analysis results has to be that the factors of mental ability are related to each other across the four groups and they are not independent of each other as Jensen's theory would predict.

A further test of Jensen's theory is possible using LISREL methodology since loadings can be constrained to zero as well as to equality. It was thus decided to test another model, F, in which not only were the factor loadings and factor covariances equal across the four groups but the factor loadings were also constrained to fit the "Jensen pattern". The schematic representation of the "Jensen pattern" would be as presented in Table 7.5.

Table 7.5 Schematic Representation of the "Jensen Pattern"

	Factor I (Level II)	Factor II (Level I)
Raven	+	o
Draw	+	o
Piaget	+	o
VSMT	o	+
OPAT	o	+
WOPAT	o	+
SM	o	+

According to Jensen's theory, the Raven, Draw-a-Man and the Piaget Test would be maximally loaded on Factor I (identified in the above table by '+') and be least loaded (represented by 'o') or orthogonal to Factor II. The obverse would be true of VSMT, OPAT, WOPAT and SM.

The actual factor analysis results obtained are presented in Table 7.6.

Table 7.6 * Factor Analysis of Model F with the Additional Constraint of the "Jensen Pattern"

	Factor Loadings	
	I	II
Raven	75	00
Draw	59	00
Piaget	66	00
VSMT	00	60
OPAT	00	68
WOPAT	00	51
SM	00	56
	Factor Correlations	
Factor I	1.00	70
Factor II	70	1.00

*Decimals omitted.

This model is very similar to the Promax rotation and Model D.

The correlation between Factor I and Factor II ($r = .70$) again clearly suggests that the two factors are highly related to each other and the model does not support Jensen's hypothesis of independence of Level I and Level II abilities. Further Model F does not provide a good fit to the data (Chi Square = 118.21; $df = 76$; $p = <.01$) suggesting that the very clearcut division of tests into Level I or Level II is not correct.

It is therefore concluded that of the six models hypothesised the best fit Model is D which suggests that across the four groups there are no significant differences either in their factor loadings or in their factor covariances.

The fact that the factor loadings and factor covariances were the same across the four groups, helped to meet one of the aims of the study namely, that the LET Battery should be equally useful in assessing English and Asian children - although it needs to be acknowledged that the genesis of its development was to meet the needs of Asian children. As the factorial structure of the LET Battery is not different for either racial group it can be claimed that it is a "culture fair" test in so far as English and Asian children are concerned (See Jensen, 1974b; 1980 a; 1980 b). These findings are a considerable advance compared to the pioneer work of Haynes (1971) and Hegarty and Lucas (1978) which like the present study, were also carried out to meet the assessment needs of children from ethnic minority groups. Neither Haynes nor Hegarty & Lucas provide any information about the comparative factorial structure of the English and ethnic groups they studied. In fact Hegarty & Lucas' (1978) study did not include English children at all.

7.3 Analysis of English and Asian Children's Performance on Measures of Level I and Level II Abilities

The previous section dealt with the factorial structure of the LET Battery. In the present section attention will be focused on the second hypothesis related to Jensen's theory of mental abilities. To recall, this hypothesis was: Asian and English children might differ in means for Level I and Level II abilities.

This hypothesis was developed from Jensen's (1973) following statement concerning his theory:-

> "Population groups that have developed under different selective pressures for different abilities, and through historic, geographic and relative social isolation from one another, might therefore be expected to differ in Level I and Level II abilities..." (1973, p. 264)

Jensen himself does not seem to have studied Asian children (the way they have been defined here and not the way he has defined them, see 1980 a, 1980 b; Jensen & Inouye, 1980) but since Asian children in Britain fulfilled the criterion for growing under different selective pressures and of experiencing differing "historic, geographic and relative social isolation" from English children, it was therefore justifiable to expect, on the basis of his theory, that there might be differences in the performance of English and Asian children in the measures of Level I and Level II abilities.

Perhaps, statistically speaking, in the light of the factor analysis results discussed in the previous section, subtests of the LET Battery and the Raven's Matrices and Draw-a-Man test should no longer be described as measures of Level I and Level II the way they had been classified so far, because the factor analysis has clearly demonstrated that, that hypothesis is not tenable. However, logically, and for the sake of clarity in the exposition, it would seem more appropriate that the learning tests, the Visual Sequential Short Term Memory Test should still be described as measures of Level I, and the rest as measures of Level II until all the hypotheses concerning Jensen's theory have been tested.

As one of the hypotheses was to test differences due to Race on measures of Level I and Level II abilities, the MANOVA Test seemed an appropriate technique to test such differences since this technique is often used to test the "realness" of the differences among the population centroids, or means vectors (Cooley & Lohnes, 1971, p. 224).

Race had two levels: English and Asians; age had two levels: Second Year Junior and Top Infants/First Year Junior (as in the previous analysis these two latter groups were combined in order to have an adequate number of children for the analysis). The thirteen variables studied were: Raven's Coloured Progressive Matrices (Raven), Draw-a-Man Test (Draw), Piaget, Visual Sequential Short Term Memory Test (VSMT), Object Picture Association Test (OPAT), Word Object Picture Association Test (WOPAT), Symbol Manipulation (SM), Graded Arithmetic-Mathematics Test administered at the same time as the LET Battery (MATHAGE), Schonell Word Recognition Test administered at the same time as the LET Battery (READAGE), Graded Arithmetic-Mathematics Test administered the second time after a year of the first testing (MATH 81), Schonell Word Recognition Test administered after a year of the first testing (READ 81), Teachers Rating according to attainments (TRA), Teachers Rating according to potential (TRP). According to Jensen's theory (1969) the IQ tests (Raven and Draw), the Piaget and attainment tests (both reading and maths tests) would be classified as measures of Level II ability, the rest i.e. VSMT, OPAT, WOPAT and SM as measures of Level I ability.

A MANOVA Test showed that the means for the whole set of variables differed among the four groups (Lambda = 0.57; F (39, 1040) = 5.51, significant at the 1% level) so that it was thought legitimate to investigate the differences amongst the means separately.

7.4. Analysis of Variance Results

A series of two way analyses of variance (ANOVA) were undertaken as a follow-up for further interpreting the significant MANOVA (see Bray & Maxwell, 1982 who discuss at considerable length the "proper" method(s) to employ once the MANOVA rejects the null hypothesis of no overall significant differences). The list of variables was the same as for the MANOVA Test and so was the cross classification: Age x Race. Table 7.7 gives a summary of the thirteen 2 x 2 ANOVAS. The other ANOVA details (e.g. Sum of Squares, Degree of Freedom, Mean Square and significance of F) for all the thirteen variables are provided in Appendix C.

It is often recommended in the standard statistical text books (e.g. Lewis, 1968) that in the absence of any significant interactions, the main effects are of chief interest. As none of the Age x Race interactions turned out to be significant, therefore, attention will be focused on the main effects only. In the following sections, first, simply the results of the meaningful main effects will be presented, to be followed by a discussion of those main effects, which significant or not, merit a discussion as they shed some interesting light on some of the hypotheses.

Table 7.7 Summary Table of Thirteen 2 x 2 ANOVAS

Variables	Main Effects F Ratios		Interaction F Ratios
	Race	Age	Race x Age
IQ Tests			
Raven	17.13 **	18.57 **	3.81 ns
Draw-a-Man	.12 ns	8.34 **	1.14 ns
LET Battery			
Piaget	.30 ns	22.10 **	1.03 ns
VSMT	8.70 **	1.66 ns	.414 ns
OPAT	1.50 ns	.12 ns	.95 ns
WOPAT	.04 ns	126.25 **	.25 ns
SM	.07 ns	8.69 **	1.15 ns
Attainment Tests			
MATHAGE	1.92 ns	25.18 **	2.27 ns
READAGE	1.23 ns	13.33 **	1.48 ns
READ 81	.00 ns	14.67 **	1.55 ns
MATH 81	2.25 ns	19.25 **	2.26 ns
Teacher's Ratings			
TRA	1.70 ns	.49 ns	.84 ns
TRP	.54 ns	.00 ns	.87 ns

* Significant at 5% level.
** Significant at 1% level.
ns Not significant.

7.4.1 The Main Effect Race

Of the thirteen there were only two variables, the Ravens and Visual Sequential Short Term Memory, where there were significant differences between English and Asian children at the 1% level. In Table 7.8 are shown means (adjusted for age) on the Raven's and the VSMT for Asian and English children.

Table 7.8 Mean Scores on the Raven and Visual Sequential Short Term Memory as a Function of Race (Adjusted for Age)

	Means	
	English	Asian
Raven	22.80	20.15
VSMT	17.53	19.39

On the Raven's, as expected, the difference of 2.65 points was in favour of English children whilst a slightly smaller difference of 1.86 on the VSMT was in favour of Asian children. The latter difference was neither intended nor expected. On none of the other tests were there any significant differences. In the light of Jensen's theory and the review of literature pertaining to the performance of English and Asian children the main effect, Race, raised a few questions.

1) Why were there significant differences on the Raven's Matrices and not on the Draw-a-Man Test?

2) Why were the differences on the VSMT in favour of Asian children?

3) Why, contrary to a reasonable amount of evidence in the literature, were there non-significant differences on all the attainment tests?

Each question will be taken up in turn following the outlining of the results of the main effect, Age.

7.4.2 The Main Effect Age

In examining the main effect, Age, of interest here are the age differences related to the LET Battery rather than the other variables such as attainment tests, IQ tests or teacher's rating.

Table 7.7 shows that none of the Race x Age interactions are significant. Of the thirteen variables there are nine variables where, as expected, there are significant differences between the younger and older children and these differences favour the later group; in each case they are significant at the 1% level with 1, 384 df. On four variables, namely, TRA, TRP, OPAT and VSMT there are no statistically significant differences. This is expected for TRA and TRP which are presumably norm referenced.

There are two subtests from the LET Battery, VSMT and OPAT, where there are no statistically significant age differences which merit some discussion.

Like the main effect, Race, the main effect, Age, too raised an important issue: why the F ratio did not show any significant differences on the VSMT and OPAT? This question and the other questions related to the main effect, Race, posed earlier, will be considered now. First, questions related to the main effect, Race.

7.4.3 Discussion of the Main Effect Race Results

The statistically significant difference in the scores of English and Asian children on the VSMT would suggest that inadvertently a learning skill has been identified which would appear to be more developed in the Asian culture compared to the indigenous English culture. Although this outcome was not expected, it is consistent with the vast cross-cultural evidence which does suggest that different cultures do favour the fostering of differing cognitive/learning skills and:-

> "There is indeed reason (Goodnow, 1976) to believe that habitual cognitive values and strategies vary much more from one culture or subculture to another than had been supposed..." (McReynolds, 1982, p. 122; see also Anastasi & Foley, 1949; Bruner, 1966; Cole et al., 1971; De Vos & Hippler, 1969; Ghuman, 1975; Vernon, 1969)

Soviet psychologists believe:-

> "...that cultural variability in intellectual activities is the basis for both developmental and cross-cultural differences in memory. This seems to be so because of the important role played by memory strategies, mnemonics or other memory activities (e.g. verbal encoding, perceptual scanning, depth of processing and so on), that may vary across individuals

and contexts for remembering, as well as across societies." (Wagner, 1981, p. 192)

There is further support to this line of thinking if the Indian Sample's mean scores on the VSMT were examined and compared with the mean scores of English and Asian children on the same test. Table 7.9 shows that the mean of Asian children was about the same as the Indian children; whilst the means of Asian and Indian children were greater than the mean of English children by more or less the same magnitude.

Table 7.9 Mean Scores, SDs of English, Asian and Indian Children on VSMT - 8 to 9 Years

	Mean Scores (SDs in Brackets)		
	English (n=87)	Asian (n=101)	Indian (n=30)
VSMT	17.52	19.19	19.50
SDs	(5.85)	(6.13)	(5.43)

The foregoing explanation must apply to a certain extent to the unexpected non-significant differences of Asian and English children on the Draw-a-Man (part of the first question). The mean scores of English and Asian children, when adjusted for age, were 27.28 and 27.55 respectively. One of the plausible explanations of this non-significant difference may well reside in the fact that English and Asian children had a fairly homogenous socio-economic status and probably did not have too dissimilar early experiences. The role of early experiences related to a particular skill are quite critical as this has been shown by numerous researchers working in the field of cross-cultural studies (cf. Anastasi & Foley, 1949; Cronbach, 1970; Ghuman, 1975, 1981 in which he refers to several studies; Vernon, 1969). Further, it also seems reasonable to speculate that the Draw-a-Man is a relatively more culture fair test compared to other IQ tests (e.g. Raven's Matrices. WISC-R, 1974) particularly where the main population is concerned. Haynes (1971) too did not find any significant differences on the Draw-a-Man Test between the performance of Asian and English children whilst she found significant differences on some of the subtests of the Weschler Intelligence Scale for Children (e.g. WISC Picture Completion, WISC Picture Arrangement, WISC Block Design, WISC Object Assembly, WISC Performance Scale IQ, WISC Vocabulary).

The other reason for the non-significant differences on the Draw-a-Man (and this reason would apply to the third question as well: i.e. why, contrary to evidence, there were no significant differences on the attainment tests?) could very well be that. as has been shown, a Western type of education seems to have some influence on certain skills of Asian children and with time they tend to catch up in these areas with their English counterparts (Ghuman, 1975, 1978, see also Bruner, 1966). This line of thinking was confirmed by Monica Taylor who came to a similar conclusion after an exhaustive review of the literature concerned with examining the performance of Asian children on IQ and attainment tests. Taylor, who works for the NFER, has recently carried out a very detailed survey about the performance of Asian children for the Swann Committee who have been asked to look into the performance of Asian children in Great Britain. In a personal communication, she states that in the numerous studies she has reviewed, she too found that, in the main, Asian children who had all, or major part , of their education in Britain scored as well as English children on the Draw-a-Man and attainment tests.

It would, therefore, seem that the Western type education influenced the performance of Asian children on the Draw-a-Man Test (and attainment tests). The mean scores of English, Asian and Indian children on the Draw-a-Man were consistent with this line of thinking (see Table 7.10). Children who had no exposure to the Western type culture (i.e. Indians) had lower mean scores on the Draw-a-Man compared to both Asian (who have had some exposure) and English children.

Table 7.10 Mean Scores and SDs of English, Asian and Indian Children on Draw-a-Man - 8 to 9 Years

	Means (SDs in Brackets)		
	English (n=87)	Asian (n=101)	Indian (n=30)
Mean on Draw-a-Man	28.21	28.88	25.30
SD on Draw-a-Man	(7.56)	(8.41)	(7.81)

Turning to the third question: why did the study not show significant differences on attainment tests? Two reasons have already been offered: namely, the homogenous background and possibly the influence of the Western type of education. Yet another reason could be that Asian parents value education more than the indigenous population and therefore reinforce their children to work harder so that their children's performance in schools reaches a satsifactory standard. Thus even though Asian children start schooling with a certain disadvantage, as time goes on, they tend to catch up by dint of sheer perserverence (Spencer, 1982; see also Allport, 1937; Cattell, 1965; who discuss the role of certain personality traits in achievement).

The issue that yet remains to be discussed is: if Asian children's performance was sensitive to Western type education on attainment and the Draw-a-Man Test why was it not influenced on the Raven's Matrices as well? It is quite possible that Western type education does not affect all the tested abilities equally (cf. Ghuman, 1975). Perhaps certain abilities take considerably more time before they are affected as a consequence of different cultural influences.

Hence, probably the reason that the abilities of Asian children required to perform well on the Draw-a-Man Test and attainment tests were more easily, and relatively within less time, influenced compared to the abilities required to do well on the tasks like the Raven's Matrices or the WISC (cf. Haynes, 1971). How long it would take for the dominant culture to begin to influence Asian children's performance so that they equalled English children's performance, remains yet to be seen. The present research does clearly suggest that the Raven's Matrices discriminated against Asian children. It would therefore seem that despite the popularly held belief about the Raven's Matrices that it is a culture fair test, the present study has shown that in so far as Asian children are concerned, it is not and has also highlighted the complex role of culture on measured abilities (cf. Ghuman, 1980).

7.4.4 Discussion of the Main Effect Age

Of the five subtests from the LETB, for two subtests - VSMT and OPAT - the F ratio showed non-significant differences. Examination of the mean scores on the VSMT and OPAT showed (see Table 7.11) that although the F ratio failed to demonstrate age differences in the mean scores on these two tests, older children did have marginally higher means than the younger children. In the case of OPAT, because of the fairly high mean scores achieved by both the older and younger groups, the "ceiling" effect must have, to some extent, contributed to the non-significant differences between the two groups.

Table 7.11 Mean Scores and SDs of Second Year and First Year Top Infants on VSMT and OPAT

	Mean Scores (SDs in Brackets)	
	First Year & Top Infants n=175	Second Year n=210
VSMT	18.11 (6.27)	18.98 (5.91)
OPAT	31.86 (6.72)	32.17 (7.29)

It is apposite to reiterate here that the "ceiling" effect was inevitable because one of the key aims of the LET Battery was that it should be able to discriminate the bottom 5% of the population, and also because of the artefact of the test administration model which offers a substantial amount of learning experience to the child during the time of testing to determine if he can reach the mastery level. In order to achieve this goal, the discriminating power of these two subtests for these two age groups has been sacrificed. Although discriminating power among the typical Top Infants and First Year Junior and Second Year Junior has not been possible to achieve, the OPAT and VSMT still discriminate between the mainstream sample and the ESN-M children (see Table 7.12) notwithstanding that the mean age of the later sample is somewhat higher than the other groups - as their respective means show.

Table 7.12 Means and SDs of ESN-M, Second Year Junior and First Year and Top Infants Combined on VSMT and OPAT

		Means (SDs in Brackets)	
		VSMT	OPAT
ESN-M n=40	=	3.92 (5.21)	15.60 (12.57)
Second Year n=210	=	18.98 (5.91)	32.17 (7.29)
First Year & Top Infants n=175	=	18.11 (6.27)	31.86 (6.72)

With the ESN-M child's mean scores on these two subtests as a guide, a practicing educational psychologist should still be able to decide whether the child should continue to receive education in the mainstream or be recommended for a special educational provision. Thus, notwithstanding, the inability of OPAT and VSMT to discriminate between the younger and older age groups, these tests still have the potential to be useful in the work of an educational psychologist; but more importantly, they are still able to meet one of the aims of the LET Battery that it should be able to assist in determining the future educational needs of Asian children.

Yet another reason that the F ratio does not reveal significant differences on the OPAT and VSMT could possibly be due to the fact that the learning skills measured by these subtests are such that the performance on them does not improve with age - unlike other subtests of the LET Battery. In view of this these two subtests should not be used on older children. However, the LET Battery still has three subtests (Piaget, WOPAT and SM) which have demonstrated age-related differences which means that these can be usefully employed for the younger as well as older children.

7.5 ANOVA Results and Jensen's Hypothesis of Mental Abilities

In the light of the previous ANOVA results now the following two hypotheses, embedded in Jensen's theory (e.g. 1973) can be considered.

Hypothesis 3: There would be real differences between Asian and English children on the Raven, Draw-a-Man, Piaget, Maths and Reading tests (measures of Level II ability).

Hypothesis 4: There would be no real differences between English and Asian children on the Visual Sequential Short Term Memory Test, Object Picture Association Test, Word Object Picture Association Test and Symbol Manipulation.

Hypothesis 3: The ANOVA results do not provide support to this aspect of Jensen's theory. Except for the Raven's Matrices, on the rest of the measures of Level II ability the analysis of variance results have shown non-significant differences between English and Asian children.

Hypothesis 4: Superficially, it would seem that, on the whole, the analysis of variance results provide support to this hypothesis: after all, out of the four tests of Level I, there is only one subtest (VSMT) where the ANOVA has shown significant differences between the two racial groups. On the other hand, it could be argued that the non-significant differences on the three tests of Level I ability are largely on account of the fact that in selecting these (as a matter of fact this applies to all the subtests of the LET Battery) several factors were taken into consideration to ensure that there should not be significant differences in the performance of English and Asian children on any of the LETB subtests (see principles on pages 56 to 59 and also the Chapter Rationale Behind the Pilot Version of the Learning Efficiency Test Battery). Therefore it cannot unequivocally be said that the ANOVA results support Hypothesis 4.

CHAPTER 8

CONCLUSIONS AND RECOMMENDATIONS

8.1 Conclusions

The present research had two principal aims. Firstly, to develop a test battery for Asian children brought up in England. It was intended that this Battery (Learning Efficiency Test Battery - LETB) would have some special features which none of the available batteries would have, although they might be similar to the LETB in some respects. Although this Battery was developed mainly with Asian children in mind, it was hoped that it would be of equal value for English children as well - in other words, the LETB should be relatively culture fair as far as these two races are concerned.

The second major aim of the study was to test the generalisability of Jensen's theory of mental abilities, Level I and Level II, to populations other than those he had studied. Related to Jensen's theory the main question that was addressed was: if a study similar to that of Jensen's was carried out in Britain involving English and Asian children and assessed them on tasks measuring similar functions (i.e. Level I and Level II) as Jensen has done in his investigations, would there be a similar pattern of results as Jensen had found?

8.2 The Learning Efficiency Test Battery

All the indications are that, as intended, it has been possible to devise a test battery (LETB) which has satisfactory psychometric properties, differs from the conventionally used IQ tests in the procedures of administration, rationale, purpose and content, and does not suffer from some of the criticisms which are often levelled against the conventional IQ tests especially in relation to their use on minority group children (Cole, 1975; Haywood et al., 1975; Gallagher, 1976; Hegarty, 1977; Lambert et al., 1974; McReynolds, 1982; McLoughton & Koh, 1982; see also Sattler, 1974). The LETB also seems free from those weaknesses which even the recently developed learning tests (e.g. Haynes, 1971; Hegarty & Lucas, 1978) suffer from and which have been pointed out in the previous pages. As well as of practical use to psychologists, it would also be possible to link its outcomes with educational objectives. The major features of this Battery are consistent with current thinking about assessment procedures (Hegarty, 1977; Lambert et al., 1974; McReynolds, 1982; McLoughton & Koh, 1982; see also Eaves & McLaughlin, 1977 for the evaluation of the currently popular assessment approaches) and it is likely to be useful for both English and Asian children - although the main inspiration for its development was Asian children living in Britain.

8.2.1 Psychometric Properties

Both varimax and oblique factor rotations confirm that the LETB has five underlying relatively homogeneous factors. The reliability coefficients of the LETB are not only of satisfactory magnitude but they are also somewhat better than Haynes (1971) and the NFER's (Hegarty & Lucas, 1978) later adaptation of her (Haynes) pioneer work in this field in the United Kingdom. The reliability coefficients of the LETB also compare very favourably with the conventional and often used IQ tests such as the WISC-R (1974), Raven's Coloured Matrices (Raven et al., 1978) and the Draw-a-Man Test (Harris, 1963).

Three types of validity were determined: face validity, criterion related or predictive validity and construct validity. In order to determine its face validity several academic and educational psychologists were consulted. They all unanimously agreed that the LETB had an adequate face validity.

Construct validity of the LETB was obtained by the method of using 'contrasted' groups (Anastasi, 1976) - children in the mainstream and children in special schools. (The schools were designated as ESN-M prior to the Education Act 1981 (DES 1981; 1983). From the Analysis of Variance results it can be unequivocally claimed that the LETB has a highly satisfactory construct validity. On all the subtests of the LETB there were significant differences beyond the 1% level in the favour of children in the mainstream. No other tests of its kind provide information about its construct validity based on 'contrasted' groups. Obviously the availability of the information of the construct validity adds to the LETB's usefulness as a decision making instrument, that is, which children should be recommended to stay in the mainstream and which should be recommended for special educational provision.

Predictive validity of the LETB is satisfactory too. It was determined first by involving the main sample and then on a specially selected sample - the details of the latter (Section headed 'Longitudinal Predictive Validity') are provided after the details of the former have been supplied.

Multiple regression analysis, based on the main sample showed that approximately 57% of the variance of Maths is explained by the combination of the two IQ tests and the LETB subtests. There is little difference in the predictive validity of the LETB in relation to the conventional IQ measures for predicting Maths, the latter predictors perhaps being only marginally better. For instance, the two IQ tests explain 45% of the variance in Maths while the LETB accounts for 41%. The two IQ tests and the subtests of the LETB jointly explain 43% of the variance in Reading. Again, as with Maths, with Reading too, there is little difference between the relative predictive abilities of the two conventional IQ tests and the subscales of the LETB. The fact that the LETB is a somewhat better predictor of Maths compared to Reading would appear to be largely due to the relative contributions which Piaget based tests make and which on the whole tend to be better predictors of Maths ability compared to Reading (see Jensen, 1980 b; Lunzer & Dolan, 1977). Again, as with Maths, with Reading too, there is little difference between the relative predictive abilities of the two IQ tests and the LETB sub-tests - the former explains 30% of the variance whilst the latter explains 28%.

8.2.2 Longitudinal Predictive Validity

The major aim of this study was to test empirically, on a longitudinal basis, the validity of the LETB, and to compare it with the conventional assessment procedures based on a specially selected sample. One of the key features of this specially selected sample was that it consisted of only those children who, according to the conventional assessment procedures (i.e. IQ measures and/or teacher's ratings), were in the bottom 5% of the population. However it should be added that the criteria for deciding on the bottom 5% included use of English norms for the Raven's Matrices for Asian children. In view of the preceding discussion, it is likely that the Asian children were actually of rather higher ability but have been under estimated due to the bias in the Raven's Matrices. The results from this study analysed by Stepwise Multiple Regression show that of all the predictors, VSMT (a subtest from the LETB) explains the maximum amount of variation in the dependent variable, reading gain; no further predictors were significant. (However, it may be of interest that a further subtest from the LETB (WOPAT) was selected next as it accounted for the greatest amount of variance (25%) in conjunction with the VSMT.) The rest of the variables including all the conventional IQ measures made little contribution towards explaining the variation in the dependent variable, Reading gain. As an alternative approach prediction of reading was also sought by using the covariance approach since gain scores have been criticised in the literature (Ferguson, 1971; Thorndike & Hagen, 1969; Vernon, 1969). This analysis confirmed the above findings: as with the previous analysis (i.e. using Reading gain as the dependent variable) VSMT still remains the best and only significant predictor. (The only difference being that in this analysis the second best predictor with the VSMT is the DRAW IQ followed by the WOPAT. Thus with the covariance approach the WOPAT takes the third place instead of the second in explaining the variation.) Thus the results of this Longitudinal Study are in favour of the LETB as opposed to the conventional assessment. The overall picture that has emerged is, that for a group such as this, the LETB, in the main, is superior in relation to the conventional measures of assessment in predicting gains in reading whether they are measured directly or indirectly through the covariance approach. What conclusions can be drawn from the two principal multiple regression analyses, one based on a sample which involved children of a wider range of ability (i.e. the main sample) and the other on a very 'restricted' range? Certainly for the low performing group ('restricted' range) the conventional measures should not be relied upon for predicting reading ability or changes in reading ability. Guilford (1967) after an exhaustive review of the earlier literature concerned with the relationship between intelligence and learning ability, came to the conclusion that the commonly held notion of equating intelligence with learning ability was not tenable (see also Feurstein, 1970, 1979; Haywood, et al., 1975; Hegarty & Lucas, 1978; Jensen, 1961, 1963, 1967, 1969). One of the conclusions arrived at by Haynes (1971) is of relevance here: She states:-

> "There was no evidence either from the original correlations or from the factor analysis of any close overlap between intelligence and learning tests, and this was confirmed by the regression analysis" (p.73).

Thus the Longitudinal Study provides support to Guildford's and others' contention that the low IQ should not be viewed as an index of low learning ability, notwithstanding, that there is little difference in the predictive ability of the two IQ measures and the LETB for this group. This is because the IQ tests fail to overcome one of the basic problems that far greater numbers of Asian children are classified as falling in the bottom 5% of the population compared to their English counterparts. For illustration, when English, Asian and Indian children are matched for age (all between the ages of 8 - 9 years) only 11.4 percent of English children fall in the bottom 5% of the population as assessed by the Raven's Matrices and/or Draw-a-Man. Whereas by similar criteria, 26.8 percent of Asian and 27 percent of Indian children fall in the "bottom 5%" of the population. Thus more than twice the number of Asian and Indian children are likely to be classified as being in the bottom 5% of the population or "mentally defective". It is a very serious limitation since results on IQ measures do influence to a considerable extent recommendations as to whether children should receive education in the mainstream or go to special schools (Chazan et al., 1974; Mittler, 1970; Tomlinson, 1981). Therefore there is a danger that reliance on IQ measures for assessing Asian children could lead to their over-representation in special schools - which is socially, politically and ethically indefensible.

The present study has also shown that particularly the Raven's Matrices should not be used on Asian children. It is perhaps worth expanding on the likely problems which will be encountered using the Raven's Matrices on Asian children. English and Asian children have different intercepts when the regression lines are drawn for predicting reading ages from the Raven's Matrices scores. It is well recognised in the literature (e.g. Anastasi, 1976) that children with a higher intercept are discriminated against. Regression analysis results show that the intercept of Asian children is higher than the intercept of English children (see Footnote). Because Asian children have a higher intercept in relation to English children, the former groups' predicted Reading Ages are underestimated by about 3 months. This is shown in Figure 8.1 and in the following example.

Consider a child with a Raven Score of 20. The unstandardised regression coefficients, B, and the respective constants for English and Asian children are shown in the following two equations.

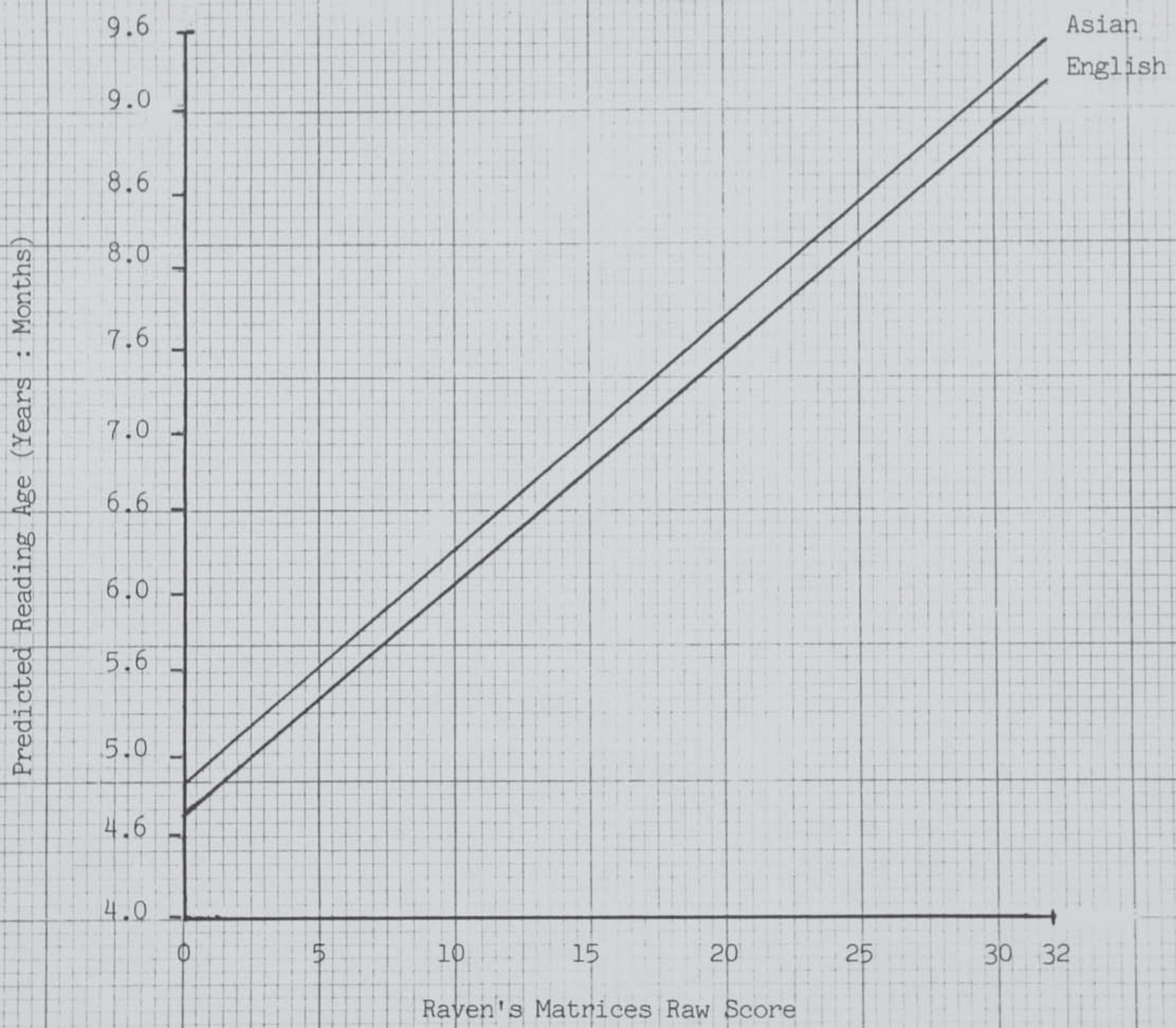
English $\hat{R} = 7.88 + 1.693\ I$

Asian $\hat{R} = 10.23 + 1.723\ I$

$\hat{R}$ stands for the predicted Reading Age and I for IQ - in this case the unstandardised score on the Raven's Matrices. For a child with a Raven score of 20, if English norms are used his predicted Reading Age would be 41.74 months, if the Asian norms are used his predicted Reading Age would be 44.69 months. Thus if an Asian child is judged by the English norms, there is an underestimation of almost three months in the predicted Reading Age.

Two separate regression analyses were carried out for English (n=172 and Asian (n=213) children using the Raven's Matrices as the predictor for the dependent variable Reading Age.

Fig. 8.1 Regression Lines of English and Asian Children to Illustrate Intercept Bias



Unfortunately the limitations of the conventional assessment procedures which the present research has highlighted, are not fully recognised by the vast majority of teachers - and many psychologists. Sometimes they are misconstrued and used to cast aspersions on children of ethnic minority groups. Sometimes such results are interpreted in a way so that they provide pseudo-scientific respectability to many teacher's erroneous thinking and beliefs that children from ethnic minority groups possess low genetic potential (cf. Ghuman, 1980).

8.2.3 LETB as a "Culture Fair" Test

All the evidence seems to suggest that, as intended, the LETB is fair to both English and Asian children - although it needs to be added that the construction of it began with the assessment needs of the latter group. In attempting to develop this Battery two areas have been identified, learning by association (Gagne, 1970; Hilgard & Atkinson, 1967; Woodworth & Schlosberg, 1954) and those mental processes as measured by the Piaget based tests (Piaget, 1952; Inhelder & Piaget, 1964). It would seem that the apparent cultural variance between English and Asian children makes little difference in their performances on the measures of these functions. This is borne out as a result of the non-significant mean differences on 4 out of 5 LETB subtests, the similar type of factorial structure for both the race and age groups, and similar type of regression weights across the four groups for predicting Maths and Reading. On the basis of these analyses a modest claim can be made that for Asian and English children the LETB is a culture fair test and is not prejudicial to children of either of these two races.

In due course, there is likely to be further evidence of the LETB'S culture fairness as it is being tried in America by Dr Gredler

(Professor of Psychology, University of South Carolina, South Carolina) and Dr Sethi (Professor of Child Development and Education, California State College, Bakersfield, California) on some of their ethnic groups. When the details of their findings become available it would then be possible to claim - or otherwise - whether the areas of learning by association and some of the Piaget based tests are culture fair for other ethnic minority groups as well. In the meantime, it appears a reasonable inference from the findings of the present investigation, that theoretically it seems feasible to devise a culture fair test for children with different cultural backgrounds so long as it is possible to identify in their cultures some common areas of mental functioning which receive equal emphasis.

As well as providing some support to the fact that there are perhaps some mental skills which can be identified across cultures which have been more or less equally developed, the present study has also unwittingly identified an area, short-term memory, which appears better developed both in Asian children and children living in India compared to English children. Besides it would also appear that short-term memory, unlike some of the other domains of mental skills (e.g. Maths, English, "mental maturity" as measured by the Draw-a-Man Test (Harris, 1963) which are sensitive to English education (TES, 1982 and several studies reported in the Chapter: 'Review of the Literature') is least influenced either by English schooling or Western culture. This seems an area where the influence of the home on Asian children is stronger than the outside influences. More importantly, this finding discredits the commonly held notion of regarding children from different ethnic and cultural backgrounds who are born and brought up in Britain, as likely to have developed the similar type of cognitive processes as the children of the dominant culture: the rationale behind this type of thinking being that all these children have had the equal opportunities to learn and been exposed to the dominant culture. Untfortunatley, this type of erroneous rationale grossly underestimates the powerful influence of home which the present study has attempted to highlight here (cf. Ghuman, 1975). However, it remains to be seen whether, with the passage of time and with the further experience of living in Britain, Asian children's performance on tests like the Raven's Matrices will improve and their performance on short term memory would regress so that eventually their performance in these two areas become at par with the indigenous children.

These findings, like the findings of many other cross cultural studies, some of which have been discussed in the earlier pages, shed a modest light on the complex role of culture and its interaction with other cultures and the impact they have on the measured abilities.

8.3 Hypotheses Relating to Jensen's Theory of Mental Ability

The present investigation tested four hypotheses which are embedded in Jensen's theory of mental abilities which has been stated in several of his studies (e.g. Jensen, 1969; 1973; Jensen & Inouye 1980). The first hypothesis was concerned with the independence (ie orthogonality) of tests of Level I ability and tests of Level II ability. In order to test this hypothesis exploratory as well as confirmatory factor analyses were carried out. At the exploratory stage an examination of the plots of the rotated factors showed that there were quite marked similarities between the second year English and Asian, and between the first year and top infants (combined) English and Asian children in the positioning of their various tests in the factor spaces - though there were a few minor dissimilarities as well. What was not at all evident from the positioning of the various tests in the factor spaces was that in any of the four Age/Race groups, tests of associative learning and memory (Level I) were forming a separate and identifiable cluster from the tests of conceptual learning tests (Level II). An inspection of the correlations amongst all these tests also quite clearly suggested that tests of Level I ability and tests of Level II ability had substantial correlations with each other and this was true in the case of all the four Age/Race groups. These preliminary findings were fully supported by the final confirmatory analysis. This analysis showed that although the measures of Level I and the measures of Level II ability loaded on two separate factors, the two factors were highly correlated.

On the basis of these analyses it can be firmly stated that contrary to Jensen's hypothesis, there was absolutely no evidence of tests of Level I and Level II ability forming two separate identifiable clusters: instead, they are all substantially correlated with each other. This is true in the case of all the four Age/Race groups. In some other studies too(e.g.Haynes, 1971) similar types of results have been obtained.

The reason that measures of Level I and Level II ability correlate with each other could very well be due to the fact that, contrary to Jensen's thinking, it is plausible that the mental processes required by the two types of measures are not too dissimilar (cf. see the cognitive psychologist's view point as outlined in Hilgard et al., 1979). Rohwer et al. (1971, and several studies reported therein) contend that associative types of learning tasks involve a substantial amount of "conceptual activity . Gagné (1977) too asserts that simple associative type of tasks may not after all be that simple. The psychological reality, as pointed out by Lunzer & Dolan (1977), would appear to be that a wide range of measures of mental processes tend to be correlated with each other rather than forming two independent clusters. Some of the other evidence available would also suggest that when only learning tasks are factor analysed they load on more than one factor (Malmi et al., 1979 and a study cited therein by Underwood et al., 1978; see also Guilford, 1967 for an earlier review of the literature pertaining to the relationship between learning and intelligence). Thus neither from the present study nor from the other studies can it be concluded that the mental abilities are divisible into two neat independent factors.

The second hypothesis was concerned with the factorial structure of the four Age/Race groups. In line with Jensen's theory the expectation was that there would be no real differences in the factorial structure of English and Asian children on the subtests of the LETB and the two conventional IQ tests. Like the previous hypothesis this hypothesis too was tested at two levels: exploratory and confirmatory. With the later technique it was possible to test hypotheses about the equality of factor loadings and/or factor covariances among groups. At both these levels of analyses it was confirmed that English and Asian children did not differ in their factorial structure on the subtests of the LETB and the two IQ tests. More precisely, after having set up several Models the factor analysis results confirmed that only two factors underlie the LETB and the other two IQ tests. Across the four Age/Race groups there were no significant differences in their factor loadings or in their factor covariances. Because of the equality of factor loadings across the four Age/Race groups (Jensen, 1974 b; 1980 a; 1980 b) a modest claim may be made that the LETB for English and Asian children is "culture fair". These findings deserve to be viewed as a somewhat considerable advance compared to the other similar studies (Haynes, 1971; Hegarty & Lucas, 1978) which provide little information about the comparative factorial structure of English and the ethnic groups they studied. In fact Hegarty & Lucas (1978) did not include English children at all in their study.

At their face value these results suggest that they are consistent with Jensen's hypothesis; later on (while discussing the last hypothesis) it will be argued that, strictly speaking, the underlying reason for the non significant differences in the factorial structure of Asian and English children is not because Jensen's theory is correct but mainly because of other reasons which really have nothing to do with his theory.

The last two hypotheses, that there might be real differences between English and Asian children on the tests of Level II ability (Hypothesis 3) and might not be statistically significant differences on the tests of Level I ability (Hypothesis 4), were tested first by the MANOVA test and then by a series of two way analysis of variance. For both these analyses there were thirteen variables: five LETB subtests, two IQ tests, two teachers ratings, two reading test scores and two maths test scores. The MANOVA results showed that the means for the whole set of variables differed significantly among the four groups: thus providing justification for investigating the differences amonst the means separately (Bray & Maxwell, 1982). This was followed by a number of two way analyses of variance (ANOVA) for further interpreting the significant MANOVA. The ANOVA results showed that none of the Race x Age interactions were significant; of the thirteen variables there were only two, the Raven's Matrices and the VSMT, where there were Race differences. On the Raven's Matrices

the differences were statistically significant (at the 1% level) in favour of English children; on the VSMT they were in favour of Asian children (significant at the 1% level).
Thus out of the seven measures which Jensen would regard as measures of Level II ability (i.e. two IQ tests, Piaget based test, and the four attainment tests) it is only on one test, the Raven's Matrices, where there are any significant differences between English and Asian children. Surely, these results cannot be regarded as consistent with Jensen's theory. Thus the hypothesis that there might be real differences in the conceptual ability of English and Asian children as measured by the tests of Level II ability is rejected.

Other workers too (Green & Rohwer, 1971; Rohwer et al., 1971) who have tried to test Jensen's theory of mental abilities have found that it does not extend even to those populations which Jensen claims he has studied and on whom his theory applies. Green & Rohwer (1971) came to the conclusion that Jensen's theory may hold for SES differences in the white population but not in the black population; or it is also possible that Jensen's model is not satisfactory. Even his own researches (Jensen, 1974 a) do not fully support all his hypotheses pertaining to his Levels theory; nor does it unambiguously extend when Jensen has included other ethnic groups (e.g. Mexicans, Chinese and Japanese) in his researches (Jensen, 1973; Jensen & Inouye, 1980).

The last hypothesis was concerned with the expectation that there would be non-significant differences on tests of Level I ability between English and Asian children. Of the four tests of Level I ability the ANOVA results showed that on the three (OPAT, WOPAT and SM) the differences between English and Asian children were not statically significant. Whilst on the fourth test, VSMT the difference between the two groups were not only significant but were in favour of Asian children. In a way, this would suggest that these results support Jensen's hypothesis. On the other hand, these results deserve to be examined in the light of several sections in Chapter 3 (e.g. 3.2, 3.3, 3.4, 3.5, 3.6 and 3.7). It was emphasized there that in the development of the LETB every effort was made to ensure that the test items should be such that the differential experiences of these two racial groups should make little difference in their performance on them. It is contended here that it was this consideration that is the underlying explanation of non significant [illegible] differences between English and Asian children and not because this aspect of Jensen's theory is generalisable to Asian and English children as well. It is again for this reason that there were non significant differences in the factorial structure of English and Asian children. It is therefore concluded that this research at best provides only nominal support to Jensen's theory of mental abilities; and the author is inclined to agree with Green & Rohwer (1971) that possibly Jensen's model is far from satisfactory.

Because Jensen's theory cannot unreservedly be generalised to English and Asian children, it is therefore recommended that it should not be used to explain significant differences on IQ and attainment tests. Uncritical acceptance of Jensen's theory, as it has been pointed out by Ghuman (1980), has the potential danger of exonerating teachers and others concerned with children from taking any "positive and constructive" steps to reduce the gap between the performance of indigenous children from ethnic minority groups.

8.4 Special Features and Potential Applications of the LETB

In the development of the LETB in addition to making some modest contribution towards the theoretical issues of cross-cultural testing and test development, the Battery, as hoped, has also some special features and potential applications which may be summarised as follows.

8.4.1 Use of the LETB for Determining Future Educational Needs

There is a both theoretical and empirical justification for using for example the Piaget based test of the LETB for identifying children who may have special educational needs. This seems possible to achieve if the child's performance can be interpreted in the light of Piaget and his associates' work. Piaget and his co-workers (Inhelder & Piaget, 1964; Inhelder et al., 1974; Piaget, undated, cited in Gruber & Vonèche, 1977) claim that there are three stages in the acquisition of seriation - and this applies to serial and approximately to ordinal correspondence as well. The three stages are:

Stage 1a. During this stage the child shows complete lack of understanding of seriation. The child's attempts largely consist of arranging a "few sticks more or less parallel to each other, horizontally or vertically, in no particular order". (Inhelder et al., 1974, p. 295.) This stage is observed when the child is around three to four years of age.

Stage 1b. However, during the second subphase of this stage the child attempts to arrange sticks in sub-series of 2, 3 and 4 but he is unable to put them together. Inhelder et al. (1974) describe it as one of the more advanced responses of this phase of the development.

Stage 2. This occurs when the child is around six years of age and the child achieves success in organising the ten elements but only by 'groping'. Although the child is able to make up a seriation, he does not show any understanding of a system of relations which is a marked characteristic of the next stage.

Stage 3. At this stage, which starts at 7 - 8 years, the child uses a systematic method first by looking for the smallest (or largest) element, then looking for the next one among those remaining and so on. This method of organising the elements is described as "properly operational" because this method implies an awareness that any given element is both larger than the preceding and smaller than those that succeed it (e.g. E$>$D C etc. and E$<$F, G etc.)" (Gruber & Vonèche, 1977, p. 385).

Gruber & Venèche (1977) report an experiment conducted by Piaget & Szeminska (undated) which gives the various stages of the concept of seriation. This experiment lends support to what has been described above about the various stages in the understanding of seriation. The results are summarized in the following table.

Table 8.1 Development of Seriation (in percentages).

Age No. of subjects		4 (15)	5 (34)	6 (32)	7 (32)	8 (21)
Stage 1.A.	No attempt at seriation.	53	18	7	0	0
Stage 1.B.	Small uncoordinated series.	47	61	34	22	0
Stage II.	Success by trial and error.	0	12	25	15	5
Stage III.	Success with operational method.	0	9	34	63	95

After Piaget & Szeminska cited in Gruber & Vonèche, 1977, p. 385.

These results demonstrate, as Piaget and his co-workers claim, that a four year old child does not begin to organise the ten sticks either by trial and error (Stage II) or by the "operational method" (Stage III).

The majority of the children (53%) are at substage la or substage lb (47%). Whilst when the child attains the age of 8, a high percentage of them (95%) can seriate "operationally . In the light of the foregoing discussion it seems possible to interpret a child's performance on the Piaget based subtest of the LETB in order to obtain an approximate index of his current level of cognitive functioning and then employing this information as a guideline for identifying his educational needs. One possible way of interpreting the child's score on the Piaget based test is to consider a score between 9 - 12 as being at Stage 3, a score between 6 - 8 as being at Stage 2, a score between 3 - 5 as being at Stage lb and a score between 0 - 2 as being at Stage la.

For illustration, take an eight year old child. Suppose his score is either between 0 - 2 (Stage la) or between 3 - 5 (Stage lb).

From this child's performance on Seriation and other subtests of the Piaget based test of the LETB, it can be inferred that his level of mental processes is roughly comparable to a typical four year old child. Thus at a chronological age of eight, this hypothetical child's cognitive level is retarded by about four years.

Clearly this is a useful piece of information for determining such a child's future educational needs. It is important to emphasize here - in case it is misinterpreted - the fact that because the Piaget based test is capable of providing information about the child's level of cognitive functioning, therefore only this information should be used for determining a child's future educational needs. Clearly for a comprehensive assessment several other factors would need to be taken into consideration as well (Chazen et al., 1974; DES, 1981 and 1983; Home Office, 1978; Mittler, 1970: discuss the other variables which should be taken into account when recommending for special educational provision).

The availability of the information about the performance of the contrasted group (i.e. ESN-M children) is also useful in comparing with the performance on the LETB of the referred child and then deciding whether such a child should be recommended to stay in the mainstream of education or should be recommended for a transfer to a special school.

8.4.2 Designing of the Curriculum and the LETB

In line with current thinking (see Ainscow & Tweddle, 1979; Alper et al., 1974; Becker & Engelmann, 1976; Eaves & McLaughlin, 1977; Newland, 1973; Toepher, 1981; Woodward, 1970; Yesseldyke & Salvia, 1974, who evaluate some of the current ideas related to assessment) it was also intended that the LETB should be able to assist in designing curriculum appropriate to the child's level of thinking. This intended characteristic of the LETB is best served by the Piaget based test. It has been shown above that with the aid of Piaget based test it is possible to estimate a child's level of cognitive processes (approximately between the ages of 3+ and 8+).

Once some estimate of the child's current level of mental processes is available, then this information can be usefully employed in designing a curriculum which matches his developmental level. For instance, if on testing one finds that the child is unable to seriate, Piaget's theory would suggest that the child at this stage is pre-numerical or does not yet have stable number concepts. Such a child therefore is not likely to understand certain Mathematics concepts (for details see Wadsworth, 1978). Matching of the curriculum with the child's level of mental processes is absolutely vital to general cognitive development as providing just rich or varied experiences can at best result in rote learning only.

More recently Toepher (1981) has advanced the view that learning, particularly high level learning, should be matched with the individual's cognitive skills. High IQ and 'satisfactory record' alone are not sufficient to cope with high level learning. Toepher came to this conclusion on the basis of his enquiry based on 1700, 12 to 14 year old children with IQ's above 120.

Thus the outcome from the Piaget based test can be employed for designing curriculum for children between the ages of 3+ and 8+ years.

8.4.3 LETB as a Test for Assessing Learning Efficiency

Results obtained from the child's performance on the LETB, besides being useful in determining the child's future educational needs, can also be used for inferring about the teaching effort he is likely to require in order to master or grasp new concepts or skills. Efficient performance on the various subtests of the LETB is inversely related to the teaching effort; in other words, high scorers on the LETB require less teaching input to reach the mastery level as opposed to the low scorers. In the testing situation, were the child to require an inordinate amount of teaching effort to reach criterion, this would suggest that in the classroom situation he would be likely to need a substantial amount of teaching help in order to grasp new concepts. For any decision making then the central question would be: can the teacher concerned provide the teaching input the child is likely to require or the psychologist needs to explore some kind of special educational provision where the teaching could be carried out at the child's pace of learning? This interpretation from the LETB is justifiable as its validity coefficient is quite satisfactory. (See espectally Section 6.9, Longitudinal Predictive Validity in Chapter 6.)

8.4.4 LETB and Recently Arrived Children from India

The availability of information about a small Indian sample not only adds a cross cultural dimension to the Battery, but may also legitimize the LETB's use on recently arrived children from India who have had no exposure to the Western culture at all.

8.5 Other Features of the LETB

The foregoing are the special features of the LETB, and as far as the author's knowledge of similar tests goes none of them have these special features built into them. However, the LETB has two features too which cannot be claimed as its novel features.

It was intended that the LETB should be only minimally dependent upon language. After the experience of testing nearly 500 children it can be safely claimed that the LETB, in its administration, requires little use of the language either on the part of the examiner or the child. Confirmation to this assertion has also been provided by some of the colleagues in the profession who have tried the Battery while still in its developmental stage, as well as by a few teachers who also tried it on small groups of children. Personal experience of administering the Battery has shown that although little language is used in the actual administration of the Battery children still enjoy taking the test. The fact that the LETB can be administered with the minimal use of language cannot be considered as its novel feature. There are other tests (e.g. Learning Ability Test produced by the NFER (Hegarty & Lucas, 1978) which also do not depend upon verbal instructions in their administration. Neither can special claims be made about the model

(Demonstration, Demonstration and Practice, Testing) on which the adminstration and some of its features are embedded. There are some similarities between the LETB's model and the model on which the Learning Ability Test (Hegarty & Lucas, 1978) is based. As a matter of fact, the LETB model owes a good deal to the model used for the administration of the Learning Ability Test .

8.6 Recommendations

8.6.1 The Need for Further Research Arising Out of the Present Study

The development and the standardisation of the experimental test materials so that the norms can be established is a major priority. The availability of norms would enable test users to compare the referred child's performance with the normative sample. Additionally, this would also help meet the needs of several professionals, academics and students both in Britain and America who have shown keen interest in the Learning Efficiency Test Battery during its developmental stages and have shown considerable dissatisfaction with the currently used standardised tests on children from ethnic minority groups.

Although the norms for the Learning Efficiency Test Battery are not yet available, the results so far would warrant the development of teaching materials linked with the assessment which take into account the child's level of mental processes and rate of learning.

Particularly the learning and short term memory subtests from the LETB should be tried on children from other ethnic minority groups as well. Not only would this determine the culture fairness of these subtests, but more importantly, it should be possible to test the hypothesis whether the process of learning by association receives equal emphasis across the cultures investigated. Should the null hypothesis be confirmed, this means an area of mental processes has been identified which can form the basis for developing culture-fair tests.

Since the Learning Efficiency Test Battery can be administered with minimal use of language therefore it has potential relevance for such diagnostic categories as hearing and language impaired chidlren. The responses of these children might shed some light on whether the mental processes of these children differ from 'normal' children.

Throughout this research it has been highlighted that the current assessment techniques of testing children from ethnic minority groups are far from satisfactory. Notwithstanding this it would seem that little systematic research has been carried out with regards to the various assessment techniques that are employed for testing children from ethnic minority groups accross the country. There is a tendency to talk about it in terms of "my experience" and not on the basis of well documented evidence. With the availability of such information, it is then possible to examine it in the light of the current thinking concerning the testing of children from ethnic minority groups. The evaluation of current practices may lead to formulating and empirically testing approp-riate and non-discriminatory assessment techniques. The end product of such research could be the production of a manual with the necessary guidelines for evaluating children from ethnic minority groups and can be made available to all the professionals who are usually involved in making placement decisions for special educational provision.

8.7 Implications for Policy

From the present research the following implications have arisen; most of them are relevant for both Asian and English children.

The assessment of learning potential by the standardised tests of learning ability should play a greater part in decision making as to which children should be recommended for special educational provision. This is particularly important for children who have been rated as being in the bottom 5% of the population by the conventional assessment procedures (e.g. IQ tests, teachers ratings). The use of conventional assessment procedures for such children either for purposes of prediction or for inferring their learning potential is indefensible.

Piaget based tests should be used more widely for determining the level of cognitive processes of Asian children in preference to conventional methods of assessment since they are far less likely to misclassify disproportionate number of them being in the bottom 5% of the population. Furthermore, the information obtained about the child's level of cognitive functioning can also be of tremendous use in designing a curriculum which matches that level thereby allowing teaching to be optimally effective.

Efforts need to be directed towards helping those teachers revise their opinions who, as a result of the researches of Jensen, have come to believe that Asian children are genetically inferior in relation to their English couterparts. Additionally, more recognition needs to be given to the influences of the home environment in shaping our mental abilities. This should enable them to counterbalance the commonly held notion that children from different cultural backgrounds if they have lived in Britain long enough, their pattern of cognitive abilities is likely to be more or less the same as their English counterparts. Such an awareness should help those involved with the education of children to understand as to why many children from ethnic minority groups continue to perform poorly compared to their English counterparts notwithstanding that they were born and brought up in the UK.

APPENDICES

APPENDIX A

MANUAL FOR THE LEARNING EFFICIENCY TEST

CONTENTS

Testing Considerations

Outlined below are some obvious but nevertheless very important testing considerations which should be borne in mind while testing children. No attempt is made here for these considerations to be comprehensive as the topic is quite adequately covered in most books on psychological testing.

1 A quiet room with as few distractions as possible should be used.

2 The examiner or the child should not be unduly disturbed.

3 Make absolutely sure that the child appears to be relaxed and comfortably seated.

4 Do not commence testing unless rapport with the child is fully established.

5 Make sure that the child is reasonably motivated and apying full attention to the task.

6 Stop testing if the child begins to betray any signs of boredom or tiredness.

7 Try to make the experience as enjoyable as possible.

8 For the sake of testing, the child should not be made to miss his favourite lesson(s).

9 Be as encouraging as possible.

10 As much as possible, try to adhere to the testing instructions.

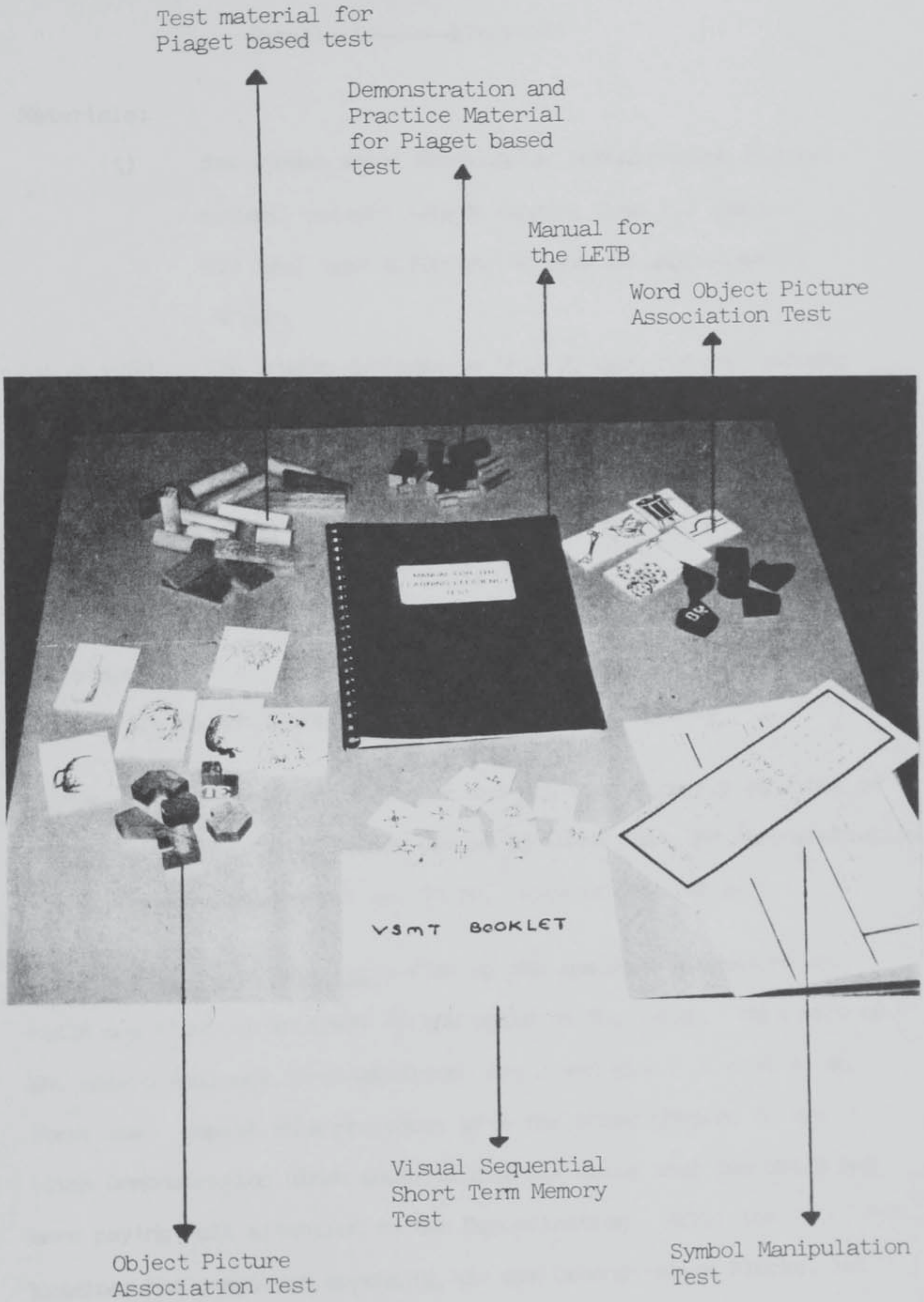
11 The testing should not be carried out by a tester:-

   (i) Whom the child does not like;

   (ii) Who, in the past, had to administer considerable punishment on several occasions and has had little opportunity of rewarding the child for his desirable behaviour or good work.

Note: The examiner would require a reasonable amount of surface area (at least 120cms x 120cms) so that the testing materials can be laid out in front of the child - and the child can work with the test items too - without getting it too cluttered.

LEARNING EFFICIENCY TEST BATTERY



SERIATION A

(Smallest ⟶ Largest)

Materials:

(i) Six wooden small Rectangular Demonstration Blocks; natural colour; length ranging from 1.3 cms to 4.3 cms, each differing in size by approximately .6 cms.

(ii) Ten wooden Rectangular Test Blocks; natural colour; length varying from 2.9 cms to 7.4 cms, each differing in size by approximately .5 cms.

STAGE 1

Demonstration

Materials -

Six Wooden Demonstration Blocks (same as above)

Randomise the Demonstration Blocks and put them directly in front of the child, on the table. Care should be taken that the Demononstration Blocks are well spaced out and in full view of the child.

Procedure for Demonstration: Pick up the smallest Demonstration Block and place it in front of the child on the table. Then pick up the second smallest Demonstration Block and place it next to the first one. Repeat this procedure with the third through to the sixth Demonstration Block ascertaining all along that the child has been paying full attention to the Demonstration. After the Examiner has completed arranging the six Demonstration Blocks, let the child observe the arrangement for a few seconds (approximately

3-5 seconds). Then pick up all the Demonstration Blocks and place them in front of the child in a random order.

STAGE 2

Demonstration and Practice 1 : Tapping once or twice near the Demonstration Blocks and then on the table where you want the child to arrange the Blocks is a sufficient non-verbal cue for some children to embark upon the task. However, some children may need more help than just tapping, in which case the Examiner places the first two smallest Demonstration Blocks in front of the child on the table. Then, indicate to the child to place the remaining four Demonstration Blocks. If the child is still unsure, place the third and fourth Demonstration Blocks for him and encourage him to do the rest.

If necessary, continue demonstrating in a similar manner until all the Demonstration Blocks have been arranged in front of the child. After the child has observed for some time, pick up all the Demonstration Blocks and place them again in front of the child in a mixed up order for a second attempt.

In the case of a child who does not seriate correctly, point to his mistake and place the Demonstration Block(s) in the correct position(s). Let the child observe for some time and then mix them up and place them in front of the child for a second attempt.

Demonstration and Practice 2-3 : Up to a maximum of three more Demonstration and Practice Trials are permissible, if felt necessary, in the same manner as described in Demonstration and Practice Trial 1.

Criterion for Discontinuing after the last Demonstration and Practice Trials

If, by the end of the last Demonstration and Practice Trial, the child is still unable to seriate, even once, then further testing of this concept need not be pursued.

STAGE 3

Testing

Materials

Ten Rectangular Blocks (same as materials (ii)

Presentation of the materials both in the case of children who are not reluctant to start and are reluctant is exactly the same. In fact presentation of the materials for the purposes of testing is almost similar to the procedure described in Demonstration and Practice 1(p223-224). The only difference being that the child is now presented with ten Test Blocks instead of six Demonstration Blocks. However, the procedure for the presentation of the materials is described here again.

Children who are not reluctant to start (spontaneous)

Presentation of Materials: Place the ten Rectangular Test Blocks in a random order in front of the child. It is important to ensure that each Test Block is spaced out so that all of them are in full view of the child. Tap once or twice near the Test Blocks and then on the table where you want the child to arrange the Blocks and try to convey through gestures and expressions that he should start to seriate.

After this, allow the child a few seconds to respond.

Obviously, the child who starts spontaneously and whose response is without error does not need any correction. In this case, testing on Seriation A need not continue any further.

<u>1st Cue</u>: If, on the other hand, the child, out of the ten Test Blocks, does not place one or more in their correct position, the Examiner should shake his head and say "No" to him. He should then correct the child's mistake(s) and draw his attention to the corrected arrangement. Allow the child a few seconds to observe the corrected position of the ten Test Blocks. Remove all the Test Blocks and place them in front of the child. Place the two smallest Test Blocks in front of the child on the table. Using the method of non-verbal cues, try to convey to the child that he has to complete the rest.

<u>2nd Cue</u>: If the child makes a mistake again, shake your head and say "No". Point and correct his mistake(s), then, as above, remove all the Test Blocks and place them in front of him. This time place the first four Test Blocks (in their order of magnitude) in front of the child and inidcate that he should place the next six. No further help should be given after this. Testing of Seriation A is discontinued.

<u>Child who is reluctant to start</u>

<u>1st Cue</u>: If the child does not start spontaneously, the Examiner should place the two smallest rectangular Test Blocks in front of him. Allow the child to observe for a few seconds.

<u>2nd Cue</u> If the child still makes no response, the Examiner should place two more Rectangular Test Blocks, (i.e. a total of four Test Blocks) of the next size, on the table and should wait for about 10 seconds. No further help is given after this.

Criterion for Discontinuing Testing: This is applicable to children who are reluctant to start and those who are not. Further testing of Seriation A should stop:

(i) After spontaneous correct response

(ii) After the first cue, if the child's response is correct

(iii) After the second cue, irrespective of correct or incorrect response.

Scoring: Scoring system applies equally to both categories of children, i.e. those who are reluctant to start and those who are not.

Total Score	No. of Cues	Explanation
3	Spontaneous	Refers to those children who did not need any help <u>at all</u>, who, as soon as the rectangular Test Blocks were placed in front of them, started seriating straight away. In order to be in this category they should have received little help apart from Demonstration(s) with the Demonstration Blocks. These children receive one point for their spontaneous correct response; two points for the trials which they did not need to undertake as their performance is an indication of their full understanding of the concept.
2	1	This category refers to children who were unable to complete the task spontaneously, but needed one cue. They score one point for this attempt and earn a credit of one point for the next trial which they did not need to undertake as their performance demonstrates that they have a reasonable grasp of the concept.
1	2	This category refers to those children who in order to perform the task required two cues. They earn one point, if after the two cues, they can seriate the six Rectangular Test Blocks correctly without any mistake in their arrangement.

SERIATION B

(Largest ⟶ Smallest)

Materials

(i) Six wooden small Rectangular Demonstration Blocks (same as in Seriation A)

(ii) Ten wooden Rectangular Test Blocks (same as in Seriation A)

STAGE 1

Demonstration

Materials

Six wooden Demonstration Blocks

Randomise the Demonstration Blocks and put them directly in front of the child, on the table. Care should be taken that the Demonstration Blocks are well spaced out and in full view of the child.

Procedure for Demonstration: Pick up the largest Demonstration Block and place it in front of the child on the table. Then pick up the second largest Demonstration Block and place it next to the first one. Repeat the procedure with the third through to the sixth Demonstration Block ascertaining all along that the child has been paying full attention to the Demonstration. After the Examiner has completed arranging the six Demonstration Blocks let the child observe the arrangement for a few seconds (approximately 3-5 seconds). Then pick up all the Demonstration Blocks from the table and place them in front of the child in a random order.

STAGE 2

Demonstration and Practice 1: Tapping once or twice near the Demonstration Blocks and then on the table where you want the child to arrange the blocks is a sufficient non-verbal cue for the child to embark upon the task. However, some children may need more help than just tapping in which case the Examiner places the first two largest Demonstration Blocks in front of the child. Then, indicate to the child to place the remaining four Demonstration Blocks. If the child is still unsure, place the third and fourth Demonstration Blocks for him and encourage him to do the rest. If necessary, continue demonstrating in a similar manner until all the Demonstration Blocks have been arranged in front of the child. After the child has observed for some time, pick up all the Demonstration Blocks and place them again in front of the child in a mixed up order for a second attempt.

In the case of a child who does not seriate correctly, point to his mistake and place the Demonstration Blocks in the correct position(s). Let the child observe for some time and then mix them up and place them in front of the child for a second attempt.

Demonstration and Practice 2-4: Up to a maximum of three more Demonstration and Practice Trials are permissible, if felt necessary, in the same manner as described in Demonstration and Practice Trial 1.

Criterion for Discontinuing after the last Demonstration and Practice Trials

If by the end of the last Demonstration and Practice Trial the child is still unable to seriate even once, then further testing of this concept need not be pursued.

STAGE 3

Testing

Materials

Ten Rectangular Test Blocks (same as in Seriation A)

Children who are not reluctant to start (spontaneous)

Presentation of Materials: Place the Ten Rectangular Test Blocks in a random order in front of the child. It is important to ensure that each Test Block is spaced out so that all of them are in full view of the child. Tap once or twice near the Test Blocks and then on the table where you want the child to arrange the Blocks and try to convey through gestures and expressions that he should start to seriate.

After this, allow the child a few seconds to respond.

Obviously, the child who starts spontaneously and whose response is without error does not need any correction. In this case, testing on Seriation B need not continue any further.

1st Cue: If, on the other hand, the child, out of the ten Test Blocks, does not place one or more in their correct position, the Examiner should shake his head and say "No" to him. He should then correct the child's mistake(s) and draw his attention to the corrected arrangement. Allow the child a few seconds to observe the corrected position of the ten Test Blocks. Remove all the Test Blocks and place them in front of the child again. This time place the two largest Test Blocks in front of the child. Using the method of non-verbal cues, try to convey to the child that he has to complete the rest.

2nd Cue: If the child makes a mistake again, shake your head and say "No". Point and correct his mistake(s). Then, as above, remove all the Test Blocks and place them in front of him. This time place the first four Test Blocks (in their order of magnitude) and indicate to the child that he should place the next six. No further help should be given after this. Testing of Seriation B is discontinued.

Child who is reluctant to start

1st Cue: If the child does not start spontaneously, the Examiner should place the two largest Rectangular Test Blocks in front of him. Allow the child to observe for a few seconds.

2nd Cue: If the child still makes no response, the Examiner should place two more rectangular Test Blocks (i.e. a total of four Test Blocks) of the next size in front of the child and should wait for about 10 seconds. No further help is given after this.

Criterion for Discontinuing Testing: This is applicable to children who are reluctant to start and those who are not. Further testing of Seriation B should stop:

(i) After spontaneous correct response

(ii) After the first cue, if the child's response is correct.

(iii) After the second cue, irrespective of correct or incorrect response.

Scoring: Scoring system applies equally to both categories of children, i.e. those who are reluctant to start and those who are not.

Total Score	No. of Cues	Explanation
3	Spontaneous	Refers to those children who did not need any help at all, who, as soon as the Rectangular Test Blocks were placed in front of them started seriating straight away. In order to be in this category they should have received little help apart from demonstration(s) with the Demonstration Blocks. These children receive one point for their spontaneous correct response; two points for the trials which they did not need to undertake as their performance is an indication of their full understanding of the concept.
2	1	This category refers to children who were unable to complete the task spontaneously, but needed one cue. They score one point for this attempt and earn a credit of one point for the next trial which they did not need to undertake as their performance demonstrates that they have a reasonable grasp of the concept.
1	2	This category refers to those children who in order to perform the task required two cues. They earn one point if, after the two cues, they can seriate the six Rectangular Test Blocks correctly without any mistake in their arrangement.

SERIAL CORRESPONDENCE

(Serial Arrangement)

Materials:

Six Demonstration Blocks (same as in Seriation A)

Six Wooden Demonstration Rods of a natural colour with a flattened edge; each 9 mm in diameter; length ranging from 1.3 cms to 4.3 cms; each differing in size by approximately 0.6 cms

Ten Rectangular Test Blocks (same as Seriation A)

Ten Wooden natural coloured Rods, round and with flattened edge; 2 cms in diameter; length varying from 2.7 cms to 7.2 cms; each differing in size by approximately .5 cms

STAGE 1

Demonstration

Materials

(i) Six Demonstration Blocks

(ii) Six Demonstration Rods (smaller in length and diameter)

Place the six Demonstration Blocks directly in front of the child on the table. The Demonstration Blocks should be well spaced so that all of them are in full view of the child. Likewise, some distance away, but alongside the six Demonstration Blocks, place the six Rods. Care should be taken about the random arrangement of both the Demonstration Blocks and the six Rods.

Presentation of the materials is alike both at the time of Demonstration and Demonstration and Practice Trials.

Procedure for Demonstration: Beginning with the smallest, arrange the six Demonstration Blocks in their ascending order of magnitude. Next, pick up the smallest Demonstration Rod and place it directly below the smallest Demonstration Block. Then pick up the next smallest Demonstration Rod and place it directly below the next smallest Demonstration Block. This procedure is followed until all the six Demonstration Rods have been placed directly below the six Demonstration Blocks. The ordinal position of each Rod should correspond with the ordinal position of each Demonstration Block.

When this arrangement is complete allow the child to observe it for a few seconds.

Remove the Demonstration Rods and place them in front of the child in the random order, leaving the six Demonstration Blocks in their position.

STAGE 2

Demonstration and Practice Trial 1: Using the method of gestures and tapping (already familiar both to the child and to the Examiner) encourage the child if he, without any help, can arrange the six Demonstration Rods in the manner demonstrated above. If the child cannot, the Examiner should pick up the two smallest Demonstration Rods and place them directly below the two ordinally corresponding Demonstration Blocks. Now use tapping and non-verbal cues again to convey to the child that he should try to match the remaining four Demonstration Rods with their corresponding four Demonstration Blocks.

If the child still does not make a response or makes a mistake, place two more Demonstration Rods of the next size directly below the next two Demonstration Blocks. Thus the ordinal position of the four Demonstration Rods corresponds with the ordinal position of the four Demonstration Blocks.

The same procedure is continued, if necessary, until the six Demonstration Rods ordinal position matches with the six Demonstration Blocks.

Demonstration and Practice 2-4: Up to a maximum of three more Demonstration and Practice Trials are permissible if felt necessary, in the same manner as described in Demonstration and Practice Trial 1.

Criterion for Discontinuing after Demonstration and Practice Trials: Prior to proceeding assessment with the testing materials, the child should be able to show that he, by the end of the last Demonstration and Practice Trial, can do ordinal correspondence with the demonstration materials with minimum or little help (up to one incorrect matching is permissible).

Criterion for proceeding to Testing: If during the trial of Demonstration and Practice sessions the child can perform the task independently (up to one error is permissible) discontinue further help and proceed with the testing.

STAGE 3

Testing

Materials:

(i) Ten Rectangular Test Blocks (same as in Seriation A)

(ii) Ten Rods

Procedure

Presentation of the Materials: The Examiner should have the above materials placed in front of him. The procedure for presenting the materials is essentially the same as during Stage 1 and Stage 2 of Serial Correspondence. The only difference is that now the Examiner and the child work with the Testing Materials instead of Demonstration Materials . Thus, before the child starts to respond he will have in front of him ten Rectangular Test Blocks arranged in their ascending order and ten Rods in a mixed up order. The child has to match the ten Rods according to their ordinal position with the ten Rectangular Test Blocks.

Children who are not reluctant to start (spontaneous): Using the method of gestures and tapping, well familiar to the child, encourage the child if he can perform the task without any assistance. If his response is correct, there is no need to test him any further on this test.

1st Cue: If his response is not correct the Examiner corrects it and lets the child observe this for a few seconds. Then remove the ten Rods (leaving the ten Rectangular Test Blocks in their position) and place them in front of the child in the manner they were originally placed when he embarked upon the task.

Place the two smallest Rods directly below the two smallest Test Blocks. The ordinal position of the Rods should correspond with the ordinal position of the Test Blocks. Now invite the child to try again.

2nd Cue: Should the child's response be incorrect again, the Examiner should correct it. Allow the child to observe the correct arrangement for a few seconds again.

Remove the Rods and place them in front of the child in the manner that they were when the Testing began. This time place the four smallest Rods below the four ordinally corresponding Test Blocks. Encourage the child to try again.

After the second cue, irrespective of whether the child's response is correct, no further help is given. Further testing on Serial Correspondence is discontinued.

Children who are reluctant to start: The child has in front of him all the materials that he needs. Tap once or twice near the Rectangular Test Blocks (already arranged in the ascending order), and the Rods and try to convey through gestures and expressions that the child should begin the task. Allow the child a few seconds before the Examiner should give any cues.

1st Cue: If the child fails to respond, place the smallest two Rods directly below the two Rectangular Test Blocks of the similar ordinal position. Again allow the child to observe for a few seconds.

2nd Cue: If the child still does not start, place two more Rods of the next size alongside to the first two. Thus the child should have placed in-front of him four Test Blocks and directly below them four Rods. The ordinal position of each Test Block should match the ordinal position of each Rod.

If the child still does not respond, or his response is correct or incorrect, no further help is given. Further Testing on Serial Correspondence is discontinued.

Criterion for Discontinuing: This is applicable to children who are reluctant to start and those who are not:

(i) Spontaneous correct response

(ii) After the first cue, if the child's response is correct

(iii) After the second cue, irrespective of correct or incorrect response.

Scoring: Scoring system applied equally to both categories of children, i.e. those who were reluctant to start and those who were not.

Total Score	No. of Cues	Explanation
3	Spontaneous	Refers to those children who did not need any help at all. As soon as the testing material was placed in front of them they started to respond. In order to be in this category, they should have received little help apart from demonstration(s) with the Demonstration Material. These children receive one point for their correct response and two points for the trials which they did not need to undertake as their performance is an indication of their full understanding of the concept.
2	1	This category refers to children who were unable to complete the task spontaneously, but needed one cue. They score one point for this attempt and earn a credit of one point for the next trial which they did not need to undertake as their performance demonstrates that they have a reasonable grasp of the concept.
1	1	This category refers to those children who in order to perform the task required two cues. They earn one point if, after the two cues they can do the rest of the task.

ORDINAL CORRESPONDENCE

(Random Arrangement)

Materials:

Six Demonstration Blocks (same as in Seriation A)

Six Demonstration Rods (same as in Serial Correspondence)

Ten Rectangular Test Blocks (same as in Seriation A)

Ten Rods (Same as in Serial Correspondence)

STAGE 1

Demonstration

Materials:

Six Demonstration Blocks

Six Demonstration Rods

Presentation of the Materials: The Examiner presents the materials in exactly the same manner as he did during Stage 1 of the Serial Correspondence.

Procedure for Demonstration: Arrange slowly one at a time, the six Demonstration Blocks in a random order. Pick up the smallest Rod and place it directly below the Demonstration Block so that the ordinal position corresponds with it. Then pick up the second smallest Rod and place it directly below the corresponding Block so that the ordinal position matches. Repeat the procedure with the remaining four Demonstration Rods and the Demonstration Blocks.

Irrespective of their random arrangment, the ordinal position of each Demonstration Rod should correspond with the ordinal position of each Demonstration Block.

When this arrangement is complete, allow the child to observe it for a few seconds.

Leave the six Demonstration Blocks in thier random positions. Pick up all the Demonstration Rods and place them nearer to the child.

STAGE 2

Demonstration and Practice Trial 1: Using the method of gestures and non-verbal cues (already well familiar both to the child and the Examiner) encourage the child if he can, without any help, arrange the six Demonstration Rods to correspond with the six Demonstration Blocks. If the child cannot, the Examiner should pick up the two smallest Rods and place them directly below their appropriate two Demonstration Blocks.

Now use the procedure of non-verbal cues again and try to convey to the child that he should try to match the remaining four Demonstration Rods with the four Demonstration Blocks.

If the child still does not respond (or makes a mistake), the Examiner should place two more Rods of the next size, directly below the two appropriate Demonstration Blocks.

Now the child has in front of him four Demonstration Blocks and four Demonstration Rods so that the ordinal positions match each other. The same procedure should continue, if necessary, until the six Demonstration Rods ordinal positions correspond with the six Demonstration Blocks.

Demonstration and Practice 2-4: Up to a maximum of three more Demonstration and Practice Trials are permissible, if felt necessary, in the same manner as described in Demonstration and Practice Trial 1.

Criterion for Discontinuing after Demonstration and Practice Trials: Prior to proceeding assessment with the testing materials, the child should be able to show that he, by the end of the last Demonstration and Practice Trial, can do ordinal correspondence (random arrangement) with the demonstration materials with minimum or little help (up to one incorrect matching is permissible).

Criterion for Proceeding to Testing: If during any of the Demonstration and Practice sessions the child can perform the task independently (up to one error is permissible), discontinue further help and proceed with the testing.

STAGE 3

Testing

Materials

(i) Ten Rectangular Test Blocks (same as in Seriation A)

(ii) Ten Rods (same as in Serial Correspondence)

Procedure

Presentation of the Materials: As in Serial Correspondence the Examiner has all the materials in front of him. Arrange slowly and one at a time, the Rectangular Test Blocks in a random order. Randomise the Rods and place them in front of the child.

By use of non-verbal cues and tapping, try to convey to the child that now his task is to match Rods with the Rectangular Test Blocks according to their ordinal position, although their arrangement is in a random order. If the child can correspond the Rods with the Rectangular Test Blocks correctly, this is described as spontaneous correspondence, i.e.the child did not need any cues. No further testing is carried out.

1st Cue: In case the child makes an error(s) in his correspondence, shake your head and say "No". The Examiner corrects the position of the Rods whose ordinal position does not match. Let the child observe the corrected arrangement. Mix up the Rods and place them in front of the child again.

Place the smallest Rod under the smallest Rectangular Test Block and the next smallest Rod under the next smallest Rectangular Test Block Having given the first cue, encourage the child to try again.

2nd Cue: Should the child again make a mistake(s) in his correspondence with the remaining eight Rectangular Test Blocks and eight Rods, shake your head and say "No". Like before correct the mistakes that the child has made in his arrangement. Let him observe for some time.

Remove the Rods and place them in a mixed up order again in front of the child. The Rectangular Test Blocks remain in their position. Put the smallest Rod below the smallest Rectangular Test Block, the next size Rod below the next smallest Rectangular Test Block and so on until the four smallest Rods have been placed directly below the four Rectangular Test Blocks of the corresponding size. Let the child complete the rest. After the second cue, whether the child's response is correct or incorrect, no further help should be given.

Criterion for Discontinuing:

(i) Spontaneous correct response

(ii) After the first cue, if the child's response is correct

(iii) After the second cue, irrespective of correct or incorrect response.

Scoring

Total Score	No. of Cues	Explanation
3	Spontaneous	Refers to those children who did not need any help at all. As soon as the testing material was placed in front of them they started to respond. In order to be in this category, they should have received little help. These children receive one point for their correct response and two points for the trials which they did not need to undertake as their performance is an indication of their full understanding of the concept.
2	1	This category refers to children who were unable to complete the task spontaneously, but needed one cue. They score one point for this attempt and earn a credit of one point for the next trial which they did not need to undertake as their performance demonstrates that they have a reasonable grasp of the concept.
1	1	This category refers to those children who in order to perform the task required two cues. They earn one point if, after the two cues they can do the rest of the task.

Visual Sequential Short-term Memory Test

Materials:

12 Small Cardboard Squares (small squares) 3.5 cm x 3.5 cm

each with a different symbol

Symbol Sequencing Test Booklet

Stopwatch

STAGE 1

Demonstration

Materials:

Twelve Small Squares

Symbol Sequencing Test Booklet

Stopwatch

Procedure

Presentation of Materials: All the small Squares should be randomly placed in front of the Examiner on the table. They should be reasonably spaced so that all of the small Squares are in full view of the Examiner.

Procedure for Demonstration: The Examiner should work from the child's left to right. Pick up the two appropriate small Squares and place them in front of the child. The order of presentation of the small Squares is shown in the Sequencing Test Booklet (which should not be shown to the child) as well as reproduced on the Record Form. Allow the child a few seconds to observe. Mix the small Squares up (only those which were used at the time of the Demonstration), and move them slightly nearer to the child.

STAGE 2

Demonstration and Practice Trial 1: By using non-verbal cues, indicate to the child that he should try now. If the child's response is correct, proceed to Demonstration and Practice 2. If the sequencing of the small Squares is not as it was shown to him, shake your head and say "No" and the Examiner should do it again for the child. Again allow the child to observe for a few seconds. Mix them up (only those which were used at the time of Demonstration) and move them nearer to the child. Then encourage the child to try. Repeat this procedure up to a maximum of three times, if necessary. Proceed to the Demonstration and Practice Trial 2, even if the child fails on Trial 1.

Demonstration and Practice Trial 2: Follow the same procedure as described in the Demonstration and Practice Trial 1. Now place three small Squares in the sequence as shown in the Symbol Sequencing Test Booklet. During this Trial the child is corrected up to a maximum of three times. If the child can reproduce the correct sequence after the first demonstration, proceed to Test Item 1. Proceed to Test Item 1 even if the child fails to reproduce the correct sequence.

STAGE 3

Testing

Materials:

Twelve Small Squares

Symbol Sequencing Test Booklet

Stopwatch

Procedure

Presentation of the Material: It is essentially the same as during Stage 1.

<u>ITEM 1-4</u>: Place the three appropriate small Squares in front of the child in the sequence shown in the <u>Booklet</u>. Allow the child to observe for seven seconds. Then mix them up (only those which were used at the time of the Demonstration), and slide them nearer to the child. Using the method of non-verbal cues, indicate to the child that he should try now. If his response is correct proceed to the next item. If the response is not correct, the Examiner should shake his head and say "No". Then correct it and allow the child to observe for seven seconds again. Mix them up again and place them in front of the child for him to have a second trial. If he can make a correct response, he should not be given the remaining trials of the same item. In that case, proceed to the next item. If the child's response is incorrect during the second attempt as well, the procedure of correction and allowing the child to observe is followed. A maximum of three trials is permitted. <u>ITEM 1-2</u>: For each item three small Squares are used. <u>ITEM 3-4</u>: For each item four small Squares are used. <u>ITEM 5-8</u>: The procedure for administration and correction is essentially the same as during the first four items. The only difference is that now the child is given a maximum of four trials; time remains the same,i.e.seven seconds. <u>ITEM 5-6</u>: For each item five small squares are used. <u>ITEM 7-8</u>: For each item six small Squares are used.

<u>Criterion for Discontinuing</u>:

(i) Further testing and teaching should stop if by the end of the first Test Item the child is still unable to make an error-free response.

(ii) Anytime during the testing stage after two consecutive failures. Example: A child fails to make an error-free response during the Demonstration and Practice Trials.

He, however, makes a correct response during the second trial of item 1, but he fails to make a correct response on any trial of Item 2 or 3. In this case, after the child has been given the last trial of Item 3, no further testing and teaching should be carried out.

Scoring: Demonstration and Practice Trials are not scored but the child's responses are recorded on the form for qualitative interpretation. One point for each correct response and one point for each trial(s) which the child did not take.

Example. A child makes a correct response on Item 1 during the first trial. On Item 2 his second trial is correct. On Item 3, he passes on the third trial. He fails on all the trials of Items 4 and 5. Because of his two consecutive failures, no further testing and teaching is carried out. His total score would be as follows.

	Item 1	
Tr 1	1	0
Tr 2	✓1	0
Tr 3	✓1	0
	Total = 3	

	Item 2	
Tr 1	1	0
Tr 2	1	0
Tr 3	✓1	0
	Total = 2	

	Item 3	
Tr 1	1	0
Tr 2	1	0
Tr 3	1	0
	Total = 1	

In addition to one point on trial 1, he also earns a credit of two points for trials 2 and 3 of Item 1. There was no point in assessing him on trials 2 and 3 as he could make a correct response during trial 1. For Item 2, he gets two points; one for the correct response and a credit of one point for trial 3. For Item 3 he gets only one point for his correct response during the last trial. Thus his total score is six.

WORD OBJECT - PICTURE ASSOCIATION TEST

Materials:

Six (5 cm x 5 cm) Picture Cards with illustrations of a Femur, a Murex, a Gimlet, a Libra, a Brazier, a Cacti

Six Wooden Blocks

STAGE 1

Demonstration

Materials

Six (5 cm x 5 cm) Picture Cards

Six Wooden Blocks

Presentation of the Material: Lay the six Picture Cards on the table in front of the child in the following order:

Picture Cards

Femur	Murex	Libra	Cacti	Brazier	Gimlet

Wooden Blocks

FR	MX	LA	CI	BR	GT

In between the Picture Cards and the child place the six wooden blocks on the table. These should be in a mixed up order and in full view of the child. The sides with the letters on should never face upwards.

NOTE: To help the Examiner to remember which wooden block and picture have to be associated, each wooden block has the first and the last letters of the illustration to which it has to be paired. For instance the wooden block which resembles an incomplete round shape has the letters FR. This means that this block has to be paired with the Picture Card with the illustration of Femur. Care should be taken that the child does not use these cues while making a response. To avoid circumlocution in instructions, each Block would be described by the letters that it has underneath; each Picture Card as Picture Card Femur rather than the Picture Card with the illustration of Femur on it. Example: Put the Block FR directly below the Picture Card Femur.

Procedure for Demonstration: This should be carried out in a random fashion and not in any set order.

The Examiner should say one of the six words(e.g.Brazier) and almost simultaneously pick up the Wooden Block (marked BR on one side), that has to be associated with this word, and place it under the Picture Card Brazier.

Likewise, again at random, say one of the remaining five words (e.g. Murex) and select from the remaining five Wooden Blocks and place it below the Picture Card Murex.

Repeat the procedure with the remaining four Blocks and the Picture Cards.

After each Wooden Block has been placed under the appropriate card, let the child observe the arrangement for a few seconds (approximately five seconds).

Then, remove all the Wooden Blocks and place them in a mixed up order in front of the child. The arrangement of the Picture Cards should be left in the same order.

During each Stage of this administration of the Test, the association of the Wooden Blocks with their respective Picture Cards should remain the same,i.e.the Block FR should always be paired with Femur, Block MX with Murex, Block LA with Libra and so on.

STAGE 2

Demonstration and Practice 1: This should be carried out in a random fashion and not in any set order.

The Examiner should say one of the six words(e.g.Cacti) and by non-verbal cues try to convey to the child that he should pick up the appropriate Wooden Block (in this case CI) and place it directly below the appropriate Picture Card (Cacti). The Examiner should help the child to place it under the appropriate Picture Card, if necessary.

Again, at random, say one of the remaining five words(e.g.Libra) and, by non-verbal gestures, encourage the child that he should select the right Wooden Block (LA) from the remaining five Blocks. If the child places it under the incorrect Picture Card, the Examiner should move it from there (while the child is watching), and place it directly below the appropriate Picture Card (Libra).

Repeat the above procedure (of associating the word with its corresponding Block and the two with the Picture Card) with the remaining items.

After each Block has been placed directly below its respective Picture Card, allow the child a few seconds (approximately 5) to observe the arrangement. Remove all the Blocks and place them in front of the child. Leave the Picture Cards in their original position.

Demonstration and Practice Trial 2: Follow the same procedure as outlined above.

STAGE 3

Testing

Materials:

Same as during the previous two stages

Procedure

Presentation of the Materials: The presentation of the materials is exactly the same as during the Stage 1. From the previous Trial the Picture Cards are already arranged and placed in front of the child in the following order:

Femur	Murex	Libra	Cacti	Brazier	Gimlet
FR	MX	LA	CI	BR	GT

In between the Picture Cards and the child are the six Blocks on the table. These should be in a mixed up order in full view of the child. As before, work from the child's left to right. Testing should be carried out in a random fashion and not in any set order.

Trial 1 Say the word "Femur" and indicate to the child that he should pick up the appropriate Block from the six lying in front of him in a random order, which during the previous trials has been associated with Femur. Care should be taken that in so doing the Examiner does not inadvertently point to the right Picture or Block. If the child picks up the Block LA (which has not been associated with Femur) let him do so and also let him place it under whichever Picture Card he thinks it should go.

Repeat the procedure with the remaining five Blocks (in a random order).

When the child has completed placing all the Blocks under the Picture Cards, the Examiner should interchange all the Blocks which have been incorrectly placed and put them at their appropriate places. The Examiner should then, in turn, lift each Block, one at a time, and say the name of the Picture Card to which it has been paired. For instance, lift the Block FR and say "Femur" and put it at the appropriate place. Repeat the procedure with the Blocks MX, LA, CI, BR and GT.

However, in a situation when the Examiner says the word "Femur" but the child does not pick up any Block at all, the Examiner should help the child to select the right Block and place it below the appropriate card. Should the child continue to display reluctance (or lack of confidence) in picking up the remaining Blocks during this trial and placing them under the correct or incorrect Picture Cards, the Examiner should continue to help the child in the manner described above.

No further help should be given during this trial.

Leave the Blocks under their respective Picture Cards for some time (approximately five seconds) for the child to observe. As during the previous trials, leave the Picture Cards in their original position. Remove all the Blocks and place them in front of the child in a mixed up order.

Trial 2 The same procedure should be adopted during this Trial as in Trial 1.

Trial 3 Change the order of the Picture Cards as follows:

Gimlet	Libra	Cacti	Femur	Murex	Brazier

Apart from this positional change of the Picture Cards, the procedure for this Trial remains the same as in Trials 1 and 2.

Trial 4 Leave the Picture Cards as in Trial 3. The rest of the procedure for administration remains the same as in the earlier trials.

Trial 5 Prior to starting Trial 5, the order of the Picture Cards should be changed again, as follows:

Libra	Cacti	Murex	Gimlet	Femur	Brazier

After the Picture Cards have been changed as above, the rest of the administration procedure remains the same as in the previous Trials.

Trial 6 Leave the cards as in Trial 5. The rest of the procedure is the same as during the earlier Trials. This is the last Trial and, irrespective of whether the child has learnt to associate all the words with their respective Blocks and Cards correctly, the testing of this subtest is discontinued.

Criterion for Discontinuing

After the sixth trial.

Scoring: The Demonstration and Demonstration and Practice are not scored. On Trials 1 through to 6 for each correct response, without any help, the child earns one point.

NOTE: The situation may arise during any trial(s) that the child let the Examiner carry on showing him the association of the first five Blocks with their appropriate Picture Cards. However, when there is only one Block left to be paired with the last Picture Card, he may then decide to respond himself. In this case, the child is permitted to respond but does not earn a score for this item.

OBJECT-PICTURE ASSOCIATION TEST

Materials

Six (5 cm x 5 cm) Picture Cards with the illustrations of a Tree, a Bottle, a Face, a Dog, an Apple and a Cup

Six Wooden Blocks

STAGE 1

Demonstration

Materials

Six (5 cm x 5 cm) Picture Cards with the illustrations of a Tree, a Bottle, a Face, a Dog, an Apple and a Cup

Six Wooden Blocks

Presentation of the Materials: Lay the Six Picture Cards on the table in front of the child in the following order:

Picture Cards

TREE	BOTTLE	FACE	DOG	APPLE	CUP

Wooden Blocks

TE	BE	FE	DG	AE	CP

In between the Picture Cards and the child place the six Wooden Blocks on the table.

These should be in a mixed up order and in full view of the child.

The sides with the letters on should never face upwards.

NOTE : To help the Examiner to remember which Wooden Block and Picture have to be associated, each Wooden Block has the first and the last letters of the illustration to which it has to be paired. For instance, the Wooden Block which is almost like a boot has the letter TE. This means that this Block has to be paired with the Picture Card with the illustration of Tree.

Care should be taken that the child does not use these cues while making a response. To avoid circumlocution in instructions, each Block would be described by the letters that it has underneath; each Picture Card as Picture Card Tree rather than the Picture Card with the illustration of Tree on it. Example : Put the Block TE directly below the Picture Card Tree.

Procedure for Demonstration: Working from the child's left to right pick up the Block TE and place it directly below the Picture Card Tree. The sides with the letters on should never face upwards. The Examiner should ensure that the child is paying full attention to the Demonstration procedure. Next pick up the Block BE and place it under the Picture Card Bottle. The orientation of the Blocks is not critical. Continue with the same procedure with the remaining four Wooden Blocks. The Block TE has to be placed under the Picture Card TE; the Block FE under the Picture Card Face; the Block AE under the Picture Card Apple, and the Block CP under the Picture Card Cup and so on.

After each Wooden Block has been placed directly below the appropriate Picture Card, let the child observe the object-picture associations for a few seconds (approximately five seconds). Then remove all the Wooden Blocks and place them in a mixed up order in front of the child. The arrangement of the Picture Cards should be left in the same order.

During each Stage of this test administration the association of the Wooden Blocks with their respective Picture Cards should remain the same, i.e. Wooden Block TE should always be paired with the Picture Card Tree; the Block BE with the Picture Card Bottle; the Block FE with the Picture Card Face and so on.

STAGE 2

Demonstration and Practice Trial 1: Pick up the Block TE and give it to the child. By using non-verbal cues try to convey to the child that he has to place it directly below the Picture Card Tree. The Examiner should help the child to place it directly below the appropriate Picture Card, if necessary. Then pick up the Block BE and assist the child to place it directly below the Picture Card Bottle. Repeat the procedure with the remaining four Wooden Blocks and the four Picture Cards.

After each Block has been placed directly below its respective Picture Card, allow the child a few seconds to observe the arrangement.

Remove all the Wooden Blocks and place them in front of the child. Leave the Picture Cards in their original position.

Demonstration and Practice Trial 2: Pick up the Block TE and give it to the child - make sure that he does not see the letters TE at the bottom of the Block. Try to indicate to the child that he should have a go at placing it directly below the Picture Card to which it has been associated during the Demonstration and Practice Trial 1. If the child's response is incorrect, the Examiner should correct it.

Then pick up the Block BE and repeat the procedure. If the child's response is incorrect, the Examiner should correct it. The above procedure should be repeated with the Blocks FE, DG, AE and CP.

After all the Blocks have been placed directly below their respective Picture Cards, allow the child a few seconds to observe.

Remove all the Blocks and place them in front of the child in a mixed up order as before. The Picture Cards should be left in their original position.

STAGE 3

Testing

Materials

Same as during the previous two trials.

Procedure

Presentation of the Materials: The presentation of the materials is exactly the same as during the Stage 1. From the previous trial the Picture Cards are already arranged and placed in front of the child in the following order:

TREE	BOTTLE	FACE	DOG	APPLE	CUP

In between the Picture Cards and the child are the six Wooden Blocks on the table. These should be in a mixed up order and in full view of the child. As before, work from the child's left to right.

Trial 1: Pick up one of the six Wooden Blocks and give it to the child. Tap near the Picture Cards and try to convey non-verbally that the child has to pair it(i.e.place it directly below) with the appropriate Picture Card. Encourage the child to attempt and let him place the first Block wherever he thinks it should go.

Repeat the procedure with the remaining five Blocks. After the child has placed all the Blocks below their respective pictures, the Examiner should correct all the incorrect responses. For instance, if the child has placed the Block BE below the Picture Card Tree and vice versa, but the rest of his responses are correct, the Examiner should interchange these two Blocks (BE and TE) and allow the child to observe the corrected arrangement.

After that, all the Blocks are removed and placed in front of the child - again in a random order. Prior to the second Trial the Picture Cards are rearranged in the following order:

FACE	DOG	CUP	APPLE	BOTTLE	TREE

Trials 2-4: The method of administration is exactly the same as described during the first trial. The only difference is that the Picture Cards have been arranged differently.

Trails 5-6: Follow the same procedure for administration as outlined above. The arrangement of the Picture Cards for the next two Trials should be as follows.

CUP	TREE	FACE	BOTTLE	APPLE	DOG

Sixth Trial is the last trial. Irrespective of whether the child has learnt to associate all the Blocks with their respective Picture Cards correctly, the testing of this test is discontinued.

<u>Criterion for Discontinuing</u>

(i) After the sixth Trial.

(ii) After three fully correct, without any help, consecutive matchings of all the Blocks with their respective Picture Cards during any trial.

Example: If a child, during the second trial, places all the Blocks below the correct Picture Cards and during the third and fourth trials again makes no mistake, he should not be tested on the remaining two trials. He should, however, be credited for those trials according to the procedure explained in the Scoring Section.

<u>Scoring</u>: The Demonstration and the Demonstration Practice are not scored.

On Trials 1 through to 6 for each correct response, without any help, the child earns one point. The child is also credited six points for each trial which he did not have to take because he reached the specified criterion prior to the last trial. Example: If the test was discontinued at the fourth trial, the child should be credited for the fifth and sixth trials(i.e.12 points).

NOTE: The situation may arise during any trial(s) that the child may let the Examiner carry on showing him the association of first five Blocks and their five respective Picture Cards. However, when there is only one Block left to be paired with the last Picture Card, he may then decide to respond himself. In this case the child is permitted to respond but does not earn a score for this item.

Maximum Score = 36

SYMBOL MANIPULATION

Materials:

1 Work Sheet

1 Pencil, 1 Pen, a Rubber

1 Stimulus Card

STAGE 1

(Items 1 - 10)

Demonstration

Materials

1 Work Sheet

1 Pencil, 1 Pen, a Rubber

1 Stimulus Card

Procedure

Presentation of the Materials: Place the work sheet in front of the child. Make sure that the surface of the table is reasonably smooth. Also place the pencil and the rubber next to the work sheet. When the Demonstration and Practice Trial 1 begins, the Examiner should pick up the pencil and offer it to the child.

Demonstration: Point to the first Demonstration Symbol (Ʒ) on the work sheet and then to the identical symbol on the Stimulus Card. Point to the first Demonstration Symbol again. When the child's attention is drawn the second time to the identical symbol on the stimulus Card, the Examiner also points to the symbol that it represents (Ч). The Examiner should draw this symbol (Ч) against the first Demonstration symbol (Ʒ). Repeat this procedure through all the Demonstration 1 items.

STAGE 2

Demonstration and Practice: Essentially the procedure here is the same as during Stage 1; the only difference is that now the child is encouraged - with the help of the Examiner - to see if he can work out the answers himself. If the child is still uncertain or makes an incorrect response, the Examiner should help him.

STAGE 3

Testing

Materials

1 Work Sheet

1 Pencil, 1 Pen, a Rubber

1 Stimulus Card

Procedure

ITEMS 1-10: First tap once or twice against the blank space of the Test Item No 1 (ξ); then tap on the Stimulus Card. Allow a few seconds for the child to work out the answer. If after a few seconds the child appears to be unsure or reluctant to write the answer, the Examiner should supply the answer.

On the other hand, if the child responds, but his response is incorrect, the Examiner should shake his head and say "no". Supply the correct answer either next to the child's response or directly below it. At the same time, also draw the child's attention to the identical symbol (ξ) on the Stimulus Card and to the Symbol that it represents. Throughout this section(i.e.Test Items 1 to 10) the Examiner either supplies the right answer, if the child is hesitant to respond, or corrects his answer if it is not right.

Follow the same procedure for Items 1 - 10.

STAGE 1

(Items 11 - 20)

Demonstration: The second Demonstration follows after the first 10 Test Items. The procedure for demonstrating this section is exactly the same as during the above demonstration. Prior to proceeding to this section of the test, please consult the Criterion for Discontinuing.

STAGE 2

Demonstration and Practice: This section is provided at the end of the first ten Test Items and three Demonstration Items. The procedure for administering these Items is identical to the procedure used above during Demosntration and Practice. However, prior to proceeding with this section, please consult the Criterion for Discontinuing.

STAGE 3

Testing

Materials

1 Work Sheet

1 Pencil, 1 Pen, a Rubber

1 Stimulus Card

Procedure

ITEMS 11-20: Prior to proceeding with this section, please consult the Criterion for Discontinuing. The procedure for administering this is identical to the procedure used during Items 1-10.

STAGE 1

(Items 21 - 28)

Demonstration: The third Demonstration is provided at the end of the first 20 Test Items. Follow the same procedure for this Demonstration as during the first Demonstration. However, prior to proceeding to this section please consult the Criterion for Discontinuing.

NOTE: If during this stage the child volunteers to do the items rather than letting the Examiner demonstrate, he should not be discouraged from attempting. However, if he makes an error(s) he should be corrected.

STAGE 2

Demonstration and Practice: This section is provided just before proceeding to the Test Items beginning with three symbols. As during the earlier two phases of Demonstration and Practice 1 and 2, the procedure for administration remains the same. However, prior to proceeding with this section, please consult the Criterion for Discontinuing.

STAGE 3

Testing

ITEMS 21-28: Prior to proceeding with this section please consult the Criterion for Discontinuing. The procedure for administering these items is almost identical to the procedure used for Items 1-20. The only difference is that after the first three Items (i.e.Items 21, 22, 23) the Examiner neither supplies the answer if the child does not respond, nor does he correct, if the child's response is incorrect.

<u>ITEMS 29-35</u>: Prior to proceeding with this section, please consult the Criterion for Discontinuing.

The procedure for administering these Items(i.e.29-35) is identical to the procedure used for previous items. The only difference is that after the first two items(i.e.29 and 30), the Examiner neither supplies the answer if the child does not respond, nor does he correct it if the child's response is incorrect.

Criterion for Discontinuing

(i) Do not continue if the child fails to make even one correct response during the first ten Test Items.

(ii) Irrespective of any number of correct responses during the first ten items, if the child fails to produce even one correct response between the Test Items 11-20, further assessment should stop.

(iii) Irrespective of any number of correct responses during the first 20 items, if the child fails to score at least two points betweeen Test Items 21 and 28, do not proceed any further.

SCORING One point for each correct response without any assistance from the Examiner. Maximum score possible is 35.

NOTE: In this test we are not concerned with the child's drawing ability. A recognisable approximation to the actual symbol should be acceptable. However, if for any reason the child finds it difficult to draw any symbol he should be permitted to point to the correct response on the Stimulus Card and this should be considered as a correct response and the child should receive credit for it.

LEARNING EFFICIENCY TEST BATTERY

RECORD FORM

Name:

Sex:

Race:

Date of Birth:

C.A.

Date of Test:

Examiner

		Cues	Raw Score	Std. Sc.
1.	Seriation A	--------	---------	---------
2.	Seriation B	--------	---------	---------
3.	Serial Correspondence	--------	---------	---------
4.	Ordinal Correspondence	--------	---------	---------
		Corrections	Raw Score	Std. Sc.
5.	Visual Sequential Short Term Memory Test	--------	---------	---------
		Trials	Raw Score	Std. Sc.
6.	Word Object Picture Association Test	--------	---------	---------
7.	Object-Picture Association	--------	---------	---------
8.	Symbol Manipulation Test	--------	---------	---------

Psychological and Child Guidance Service,
Ward End Centre,
201 Sladefield Road,
Ward End,
BIRMINGHAM B8 2SY

Photo Reduced

SERIATION

Seriation (A) Smallest ——> Largest

Demonstration and Practice*
(circle appropriately)

Trials	Right	Wrong
1	✓	✓
2	✓	✓
3	✓	✓
4	✓	✓

Test

Attempts	Type of Response	Score	
1st	Spontaneous	0	1
2nd	1 Cue	0	1
3rd	2 Cues	0	1
	Total Raw Score		

* For Qualitative Interpretation only

Maximum 3

Seriation (B) Largest ——> Smallest

Demonstration and Practice*
(circle appropriately)

Trials	Right	Wrong
1	✓	✓
2	✓	✓
3	✓	✓
4	✓	✓

Test

Attempts	Type of Response	Score	
1st	Spontaneous	0	1
2nd	1 Cue	0	1
3rd	2 Cues	0	1
	Total Raw Score		

* For Qualitative Interpretation only

Maximum 3

Photo Reduced

SERIAL CORRESPONDENCE

Demonstration and Practice*
(circle appropriately)

	Right	Wrong
Trial 1	✓	✓
Trial 2	✓	✓
Trial 3	✓	✓
Trial 4	✓	✓

Attempts	Type of Response	Score (circle appropriately)	
1st	Spontaneous	0	1
2nd	1 Cue	0	1
3rd	2 Cues	0	1
	Total Raw Score		

* For Qualitative Interpretation only

Maximum Score 3

ORDINAL CORRESPONDENCE

(Random Arrangement)

Ordinal Correspondence

Demonstration and Practice only*
(circle appropriately)

	Right	Wrong
Trial 1	✓	✓
Trial 2	✓	✓
Trial 3	✓	✓
Trial 4	✓	✓

Attempts	Type of Response	Score (circle appropriately)	
1st	Spontaneous	0	1
2nd	1 Cue	0	1
3rd	2 Cues	0	1
	Total Raw Score		

* For Qualitative Interpretation only

Maximum Score 3

Photo Reduced

VISUAL SEQUENTIAL SHORT TERM MEMORY TEST

Demonstration and Practice* 1
(circle appropriately)

	Right	Wrong
Tr_1	✓	✓
Tr_2	✓	✓
Tr_3	✓	✓

Demonstration and Practice* 2

	Right	Wrong
Tr_1	✓	✓
Tr_2	✓	✓
Tr_3	✓	✓

* For Qualitative Interpretation Only

TEST

Maximum = 28

Total Raw Score =

Total Trials =

Items 1-4 Exposure 7 seconds

TRIALS	SCORE	
1. Tr_1	0	1
Tr_2	0	1
Tr_3	0	1

Total =

2. Tr_1	0	1
Tr_2	0	1
Tr_3	0	1

Total =

3. Tr_1	0	1
Tr_2	0	1
Tr_3	0	1

Total =

4. Tr_1	0	1
Tr_2	0	1
Tr_3	0	1

Total =

Items 5-8 Exposure 7 seconds

TRIALS	SCORE	
5. Tr_1	0	1
Tr_2	0	1
Tr_3	0	1
Tr_4	0	1

Total =

6. Tr_1	0	1
Tr_2	0	1
Tr_3	0	1
Tr_4	0	1

Total =

7. Tr_1	0	1
Tr_2	0	1
Tr_3	0	1
Tr_4	0	1

Total =

8. Tr_1	0	1
Tr_2	0	1
Tr_3	0	1
Tr_4	0	1

Total =

Photo Reduced

WORD OBJECT-PICTURE ASSOCIATION TEST

	FEMUR		MUREX		LIBRA		CACTI		BRAZIER		GIMLET	
Demonstration and Practice 1 and 2					NOT SCORED							
Trial 1	0	1	0	1	0	1	0	1	0	1	0	1
Trial 2	0	1	0	1	0	1	0	1	0	1	0	1
Change the order of the Picture Cards	GIMLET		LIBRA		CACTI		FEMUR		MUREX		BRAZIER	
Trial 3	0	1	0	1	0	1	0	1	0	1	0	1
Trial 4	0	1	0	1	0	1	0	1	0	1	0	1
	LIBRA		CACTI		MUREX		GIMLET		FEMUR		BRAZIER	
Trial 5	0	1	0	1	0	1	0	1	0	1	0	1
Trial 6	0	1	0	1	0	1	0	1	0	1	0	1

Total Score = ______

Number of Trials = ______

Maximum = 36

Photo Reduced

OBJECT-PICTURE ASSOCIATION TEST

	TREE		BOTTLE		FACE		DOG		APPLE		CUP	
Demonstration & Practice 1 & 2				NOT SCORED								
Trial 1	0	1	0	1	0	1	0	1	0	1	0	1
Change the order of the Picture Cards	FACE		DOG		CUP		APPLE		BOTTLE		TREE	
Trial 2	0	1	0	1	0	1	0	1	0	1	0	1
Trial 3	0	1	0	1	0	1	0	1	0	1	0	1
Trial 4	0	1	0	1	0	1	0	1	0	1	0	1
Change the order of the Picture Cards again	CUP		TREE		FACE		BOTTLE		APPLE		DOG	
Trial 5	0	1	0	1	0	1	0	1	0	1	0	1
Trial 6	0	1	0	1	0	1	0	1	0	1	0	1

Total Score = ______

Number of Trials = ______

Maximum = 36

Photo Reduced

Symbol Manipulation

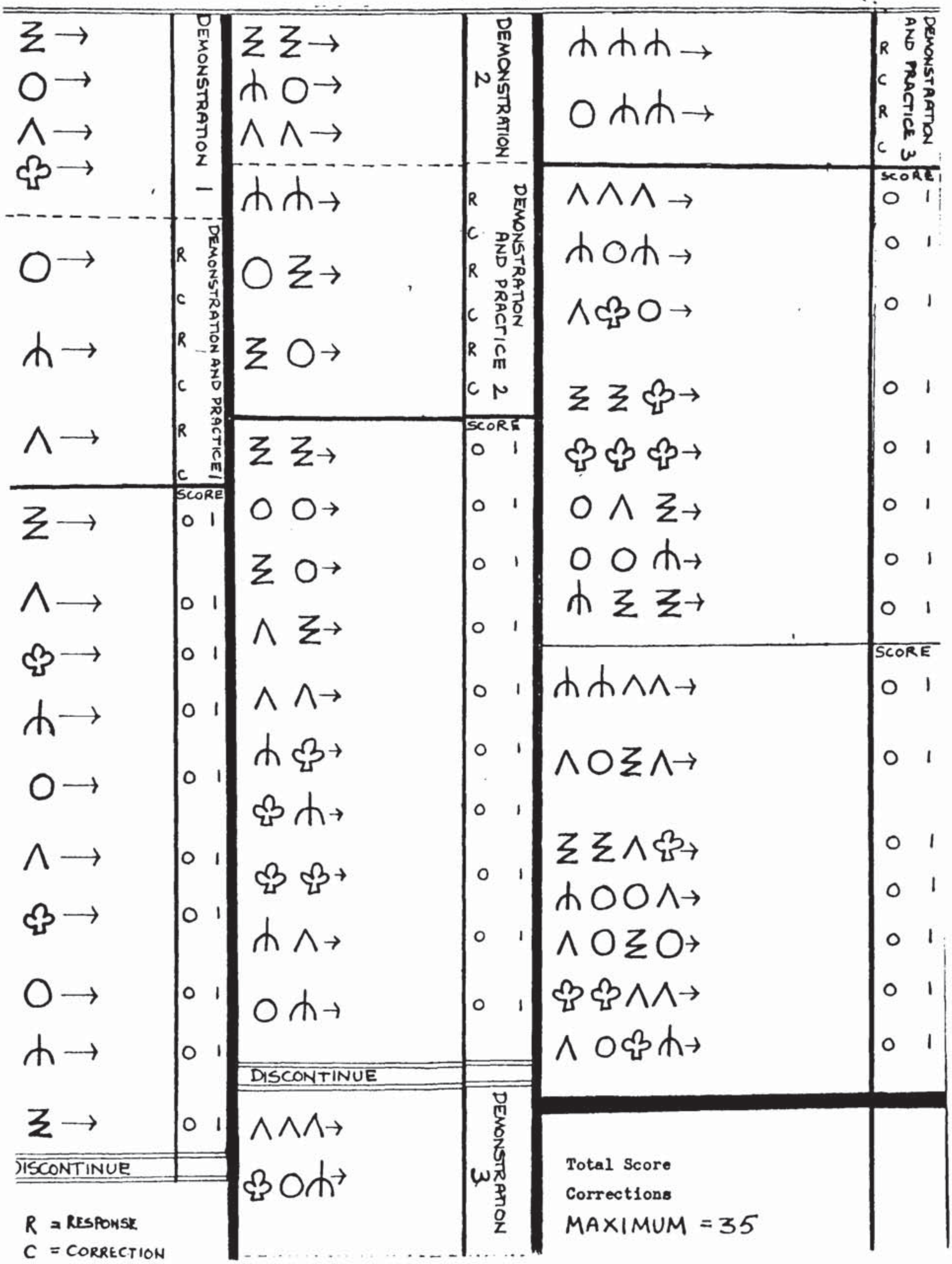


Photo Reduced

APPENDIX B

STATISTICAL TABLES

Table B.1 Corelation Matrix of Two IQ tests, Teacher's Ratings LET Battery and Reading Test Scores and the Reading Gain Scores (n = 27)

	1	2	3	4	5	6	7	8	9	10	11	12
1 RAVEN IQ	X	07	06	02	30	59**	32	19	35	28	31	11
2 DRAW IQ		X	60**	17	20	20	13	04	33	33	48**	32
3 TRA			X	74**	35	35	32	23	52**	64**	71**	25
4 TRP				X	55**	50**	49**	42*	45*	66*	65**	13
5 Piaget					X	78**	82**	73**	61**	38*	56**	38*
6 SM						X	79**	57**	66**	52**	64**	31
7 OPAT							X	76**	67**	32	53**	41*
8 WOPAT								X	44*	13	37*	42*
9 VSMT									X	34	56**	43*
10 READAGEM										X	79**	-12
11 READAGED											X	52**
12 READGAIN												X

Decimals omitted

* significant at the 5% level

** significant at the 1% level

Table B.2 Details of 2 x 2 Analysis of Variance:

Raw Score on Ravens Matrices

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig. of F
Race	664.11	1	664.11	17.13	.001 **
Age	719.74	1	719.74	18.57	.001 **
Race x Age	147.70	1	147.70	3.81	.05
Residual	14767.87	381	38.76		
Total	16207.11	384	42.21		

** Significant at 1% level

Table B.3 Details of 2 x 2 Analysis of Variance:

Raw Score on Draw a Man Test

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig. of F
Race	7.40	1	7.40	.12	.73
Age	522.98	1	522.98	8.34	.004 **
Race x Age	71.16	1	71.16	1.36	.29
Residual	23897.34	381	62.72		
Total	24510.56	384	63.83		

** Significant at 1% level

Table B.4 Details of 2 x 2 Analysis of Variance:

PIAGET

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig. of F
Race	2.15	1	2.15	.30	.59
Age	159.86	1	159.86	22.09	.001 **
Race x Age	7.46	1	7.46	1.03	.31
Residual	2757.18	381	7.24		
Total	2924.82	384	7.62		

** Significant at 1% level

Table B.5 Details of 2 x 2 Analysis of Variance :

Visual Sequential Short Term Memory Test (VSMT)

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig. of F
Race	313.01	1	313.01	8.70	.003 **
Age	59.70	1	59.70	1.66	.20
Race x Age	14.91	1	14.91	.41	.52
Residual	13063.79	363	35.99		
Total	13472.82	366	36.81		

** Significant at 1% level

Table B.6 Details of 2 x 2 Analysis of Variance:

Object Picture Association Test (OPAT)

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig. of F
Race	74.13	1	74.13	1.50	.22
Age	5.99	1	5.99	.12	.73
Race x Age	46.78	1	46.78	.95	.33
Residual	18840.35	381	49.45		
Total	18970.69	384	49.40		

Table B.7 Details of 2 x 2 Analysis of Variance:

Word Object Picture Association Test (WOPAT)

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig. of F
Race	3.05	1	3.05	.04	.85
Age	10705.95	1	10705.95	126.25	.001 **
Race x Age	21.31	1	21.31	.25	.62
Residual	32308.11	381	84.80		
Total	43119.51	384	112.29		

** Significant at 1% level

Table B.8 Details of 2 x 2 Analysis of Variance:

Symbol Manipulation Test (SM)

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig. of F
Race	2.23	1	2.23	.07	.79
Age	273.06	1	273.06	8.69	.003 **
Race x Age	36.16	1	36.16	1.15	.28
Residual	11977.56	381	31.44		
Total	12293.99	384	32.02		

** Significant at 1% level

Table B.9 Details of 2 x 2 Analysis of Variance:

Maths Age in months over 4 years (Math Age)

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig. of F
Race	655.22	1	655.22	1.92	.17
Age	8605.51	1	8605.51	25.18	.001**
Race x Age	773.90	1	773.90	2.27	.13
Residual	124055.28	363	341.75		
Total	133797.74	366	365.57		

** Significant at 1% level

Table B.10 Details of 2 x 2 Analysis of Variance:

Reading Age in months over 4 years (READAGE)

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig. of F
Race	334.43	1	334.43	1.23	.27
Age	3632.20	1	3632.20	13.33	.001 **
Race x Age	401.83	1	401.83	1.48	.23
Residual	98915.59	363	272.50		
Total	103146.76	366	281.82		

** Significant at 1% level

Table B.11 Details of 2 x 2 Analysis of Variance:

Reading Age in months over 4 years 1981 (READ 81)

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig. of F
Race	1.34	1	1.34	.00	.95
Age	4548.55	1	4548.55	14.67	.001 **
Race x Age	479.32	1	479.32	1.55	.22
Residual	112551.69	363	310.06		
Total	117615.12	366	321.35		

** Signficant at 1% level

Table B.12 Details of 2 x 2 Analysis of Variance:

Maths Age in months over 4 years 1981 (MATH 81)

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig. of F
Race	811.80	1	811.80	2.25	.14
Age	6955.47	1	6955.47	19.25	.001 **
Race x Age	817.33	1	817.33	2.26	.13
Residual	131191.28	363	361.41		
Total	139475.90	366	381.08		

** Significant at 1% level

Table B.13 Details of 2 x 2 Analysis of Variance:

Teachers Rating on Attainment (TRA)

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig. of F
Race	1.51	1	1.51	1.70	.19
Age	.44	1	.44	.49	.48
Race x Age	.78	1	.78	.87	.35
Residual	340.36	381	.89		
Total	342.98	384	.89		

Table B.14 Details of 2 x 2 Analysis of Variance:

Teachers Rating on Potential (TRP)

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig. of F
Race	.00	1	.00	.00	.96
Age	1.03	1	1.03	1.06	.30
Race x Age	.82	1	.82	.84	.36
Residual	372.30	381	.98		
Total	374.17	384	.97		

REFERENCES

Ainscow, M. & Tweddle, D.A. (1979). Preventing Classroom Failure: An Objectives Approach. Chichester: John Wiley & Sons.

Allport, G.W. (1937). Personality: A Psychological Interpretation. London: Constable & Company Ltd.

Alper, T., Nowlin, L., Lemoine, K., Perine, M. & Bettencourt, B. (1974). The rated assessment of academic skills. Academic Therapy, ix, 3, 151-164.

Anastasi, A. (1961). Psychological tests: uses and abuses. Teachers College Record, 62, 389-393.

Anastasi, A. (1976). Psychological Testing. 4th ed. New York: Macmillan Publishing Company Inc.

Anastasi, A. & Foley, J.P. (1949). Differential Psychology. New York: Macmillan.

Becker, W.C. & Engelmann, S. (1976). Evaluation of Instruction: Teaching 3. Chicago: Science Research Associates Inc.

Berry, J.W. (1966). Temne and Eskimo perceptual skills. International Journal of Psychology, 1, 207-229.

Berry, J.W. (1981). Cultural systems and cognitive styles. In M.P. Friedman, J.P. Das, N. O'Connor (eds). Intelligence and Learning. New York: Plenum Press.

Berry, J.W. & Dason, P.R. (1974). Culture and Cognition: Readings in Cross-Cultural Psychology. London: Methuen.

Bersoff, D.N. (1981). Testing and the Law. American Psychologist, 36, 10, 1047-1056.

Bower, T.G.R. (1982). Development in Infancy. 2nd ed. San Francisco: Freeman & Company.

Braine, M.S. (1959). The ontogeny of certain logical operations. Piaget's formulation examined by non-verbal methods. Psychological Monographs, 73, 5, whole No.

Bray, J.H. & Maxwell, S.E. (1982). Analyzing and interpreting significant MANOVAs. Review of Educational Research, 52, 3, 340-367.

Brimer, M.A. & Dunn, L.M. (1963). The English Picture Vocabulary Test. Slough: NFER.

Bruner, J.S. (1966). On cognitive growth 1 & 11. In J.S. Bruner, R. Oliver, P.M. Greenfield. Studies in Cognitive Growth. London: John Wiley & Sons.

Bruner, J. (1975). The ontogenesis of speech acts. Journal of Child Language, 2, 1-19.

Bryant, P. (1974). Perception and Understanding in Young Children: An Experimental Approach. London: Methuen.

Bryant, P. (1982). Editorial: Piaget's questions. British Journal of Psychology, 73, 2, 157-161.

Bryant, P. & Trabasso, T. (1971). Transitive inferences and memory in young children. Nature, 232, 456-458.

Budoff, M. (1967). Learning potential among institutionalised young adult retardates. American Journal of Mental Deficiency, 72, 404-411.

Buddoff, M. (1969). Learning potential: a supplementary procedure for assessing the ability to reason. Seminars in Psychiatry, 1, 278-290.

Budoff, M. (1973). Learning Potential and Educability among the Educable Mentally Retarded. (Progress report, Grant No. OEG-0-8-080506-45977 from National Institute of Education, HEW). Cambridge, Mass: Research Institute for Educational Problems.

Butcher, H.J. (1968). Human Intelligence: Its Nature and Assessment. London: Methuen & Company Ltd.

Call, R. (1973). Verbal abstracting performance of low-SES children: an exploration of Jensen's theory of mental retardation. Unpublished doctoral dissertation, George Peabody College.

Carroll, J.B. & Horn, J.L. (1981). On the scientic basis of ability testing. American Psychologist, 36, 10, 1012-1020.

Cattell, R.B. (1965). The Scientific Analysis of Personality. Penguin Books.

Cattell, R.B. (1966). Handbook of Multivariate Experimental Psychology. Chicago: Rand McNally.

Cattell, R.B. (1971). Abilities: Their Structure, Growth and Action. Boston Houghton-Mifflin.

Centre for Educational Research and Innovation (1977). Piagetian Inventories: The Experiment of Jean Piaget. Paris: OECD.

Chazan, M., Moore, T., Williams, P. and Wright, J. (1974). The Practice of Educational Psychology. London: Longman Group Ltd.

Child, D. (1970). The Essentials of Factor Analysis. London: Holt, Rinehart & Winston.

Clark, E.V. (1973a). What's in a word? On the child's first acquisition of semantics in his first language. In T.E. Moore (ed.), Cognitive Development and the Acquisition of Language. New York: Academic Press.

Clark, E.V. (1973b). Non-linguistic strategies and the acquisition of word meanings. Cognitive Psychology, 2, 161-182.

Clarke, A.M. & Clarke, A.D.B. (eds) (1974). The Mental Deficiency: The Changing Outlook. 3rd ed. London: Methuen & Company Ltd.

Clarke, A.M., Cooper, G.M. & Clarke, A.D.B. (1967). Task complexity and transfer in the development of cognitive structures. Journal of Experimental Child Psychology, 5, 562-576.

Clarke, A.M. Cooper, G.M. and Loudon, E.H. (1969). A set to establish equivalence relations. Journal of Experimental Child Psychology, 8, 180-189.

Cole, M., Gay, J., Glick, J.A., Sharp, D.W. (1971). The Cultural Content of Learning and Thinking: An Exploration Experimental Anthropology. New York: Basic Books.

Cole, M. (1975). Culture, cognition and IQ testing. The National Elementary Principal, 54, 4, 49-52.

Cooley, W.W. & Lohnes, P.R. (1971). Multivariate Data Analysis. New York: John Wiley & Sons Inc.

Copeland, R.W. (1979). How Children Learn Mathematics. New York: Macmillan Publishing Company Inc.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 29, 159-163.

Cronbach, L.J. (1970). Essentials of Psychological Testing. 3rd ed. New York: Harper International Edition.

Cronbach, L. (1977). Edcational Psychology. 3rd ed. New York: Harcourt Brace Jovanovich.

Dasen, P.R. & Heron, A. (1981). Cross-cultural tests of Piaget's theory. In H.C. Triandis & A. Heron (eds), Handbook of Cross-Cultural Psychology: Developmental Psychology. Vol. 4. Boston: Allyn & Bacon, Inc.

Davidson, G.R. & Klich, L.Z. (1980). Cultural factors in the development of temporal and spatial ordering. Child Development, 51, 569-371.

Davis, W.M. Jr. (1971). Are there solutions to the problems of testing black Americans? In M.M. Meir, some answers to ethnic concerns about psychological testing in the schools. Symposium presented at the American Psychological Association, Washington, D.C.

Department of Education & Science (1981). Education Act 1981. Circular No. 8/81. London: DES.

Department of Education & Science (1983). Assessments and Statements of Special Educational Needs. Circular 1/83. London: DES.

De Vos, G.A. & Hippler, A. (1969). Cultural psychology: comparative studies of human behaviour. In G. Lindzey & E. Aronson (eds), Handbook of Social Psychology, Vol.VI. Cambridge, Mass: Addison-Wesley.

Donaldson, M. (1978). Children's Minds. London: Fontana.

Donaldson, M. & Balfour, G. (1968). Less is more. A study of language comprehension in children. British Journal of Psychology, 59, 461-471.

Donaldson, M. & Wales, R.J. (1970). On the acquisition of some relational terms. In J.R. Hayes (ed.), Cognition and the Development of Language. New York: Wiley.

Duffy, J.B., Salvia, J., Tucker, J. & Yesseldyke, J. (1981). Non-biased assessment: A need for operationalism. Exceptional Children, 47, 6, 427-434.

Duncanson, J.P. (1964). Intelligence and the Ability to Learn. Princeton: Educational Testing Service.

Dunn, J.A. (1967). Inter- and intra-rater stability of the new Harris-Goodenough Draw-a-Man Test. Perceptual & Motor Skills, 24, 269-270.

Eaves, R.C. & McLaughlin, P. (1977). A systems approach for the assessment of the child and his environment: getting back to basics. Journal of Special Education, 11, 1, 99-113.

Edwards, R.P.A. & Gibbon, V. (1973). Words Your Children Use. London: Burke Books.

Elliot, A. & Donaldson, M. (1981). Piaget on language. In S. Modgil & C. Modgil (eds) Jean Piaget: Consensus and Controversy. London: Holt, Rinehart & Winston.

Elliot, C. (1981). The measurement characteristics of developmental tests. In S. Modgil & C. Modgil (eds) Jean Piaget: Consensus and Controversy. London: Holt, Rinehart and Winston.

Elliot, C.D., Murray, D., Pearson, L. (1977). The British Ability Scales. Slough: NFER.

Essen, J. & Ghodsian, M. (1979). The children of immigrants: school performance. New Community, 8, 3, 422-429.

Ferguson, G.A. (1954). Learning and human ability. In S. Wiseman, (ed.), Intelligence and Ability. Penguin Books.

Ferguson, G.A. (1971). Statistical Analysis in Psychology and Education. New York: McGraw-Hill Book Company.

Feuerstein, R. (1968). Learning Potential Assessment Device. In B.W. Richards (ed.), Proceedings of the first congress of the International Association for the Scientific Study of Mental Deficiency. Reigate, Surrey: Michael Jackson.

Feuerstein, R. (1970). A dynamic approach to the causation, prevention and alleviation of retarded performance. In H.C. Haywood (ed.), Social-cultural Aspects of Mental Retardation. New York: Appleton-Century-Crofts.

Feuerstein, R., Hoffman, M., Shalom, H., Kiram, L., Narrol, H., Schachter, E., Katz, D., & Rand, Y. (1972). The dynamic assessment of retarded performers: the Learning Potential Assessment Device, theory, instruments, and techniques. Studies in Cognitive Modifiability. (Report No. 1, Vol. I.). Jerusalem: Hadassah-Wizo- Canada Research Institute.

Feuerstein, R. (1979). The Dynamic Assessment of Retarded Performers: The Learning Potential Assessment Device, Theory, Instruments and Techniques. Baltimore: University Park Press.

Feuerstein, R. (1980). Instrumental Enrichment: An Intervention Program For Cognitive Modifiability. Baltimore: University Park Press.

Flavell, J.H. (1963). The Development of Psychology of Jean Piaget. New York: Van Nostrand Reinhold.

Foster, M. (1970). The effects of different levels of enriched stimulus input on the abstracting ability of slow learning children. Unpublished Master's Thesis, George Peabody College.

Gagné, R.M. (1970). The Conditions of Learning. London: Holt, Rinehart & Winston.

Gagné, R.M. (1977). The Conditions of Learning. 3rd ed. New York: Holt, Rinehart & Winston.

Gallagher, G.A. (ed.) (1976). National Association for the Advancement of Coloured People Report on Minority Testing. New York: NAACP Special Contribution Fund.

Garrett, H.E. (1961). Statistics in Psychology and Education. Bombay: Vakils, Feffer & Simons Private Ltd.

Ghuman, P.A.S. (1975). The Cultural Context of Thinking. Windsor: NFER Publishing Company.

Ghuman, P.A.S. (1978). Nature of intellectual development of Punjabi children. International Journal of Psychology, 13-14, 281-294.

Ghuman, P.A.S. (1980). A comparative study of cognitive styles in three ethnic groups. International Review of Applied Psychology, 29, 75-86.

Ghuman, P.A.S. (1981). An evaluation of Piaget's theory from a cross-cultural perspective. In S. Modgil & C. Modgil (eds.), Jean Piaget: Consensus and Controversy. London: Holt, Rinehart & Winston.

Gibson, E. (1969). Principles of Perceptual Learning and Development. New York: Appleton-Century-Crofts.

Glaser, R. (1981). The future of testing: a research agenda for cognitive psychology and psychometrics. American Psychologist, 36, 9, 923-936.

Goodnow, J.J. (1976). The nature of intelligence: questions raised by cross-cultural studies. Cited in P. McReynolds, The future of psychological assessment, International Review of Applied Psychology, 31, 117-139.

Gordon, J.E. & Haywood, H.C. (1969). Input deficit in cultural-familial retardation: effect of stimulus enrichment, American Journal of Mental Deficiency, 73, 604-610.

Green, R. & Rohwer, W.D. Jr. (1971). SES differences on learning and ability tests in black children. American Educational Research Journal, 8, 601-609.

Gruber, M.E. & Vonèche, J. (eds) (1977). The Essential Piaget. New York: Basic Books.

Guilford, J.P. (1967). The Nature of Human Intelligence. New York: McGraw-Hill Book Company.

Harris, D.B. (1963). Goodenough-Harris Drawing Test Manual. New York: Harcourt, Brace & World, Inc.

Hathaway, W.E. & Hathaway-Theunissen, A. (1974). The unique contributions of Piagetian measurement to diagnosis, prognosis, and research of children's mental development. Paper presented at the fourth annual conference on Piaget and the Helping Professions. Los Angeles, California.

Haynes, J.M. (1971). Educational Assessment of Immigrant Pupils. Slough: NFER.

Haywood, H.C., Filler, J.W., Shifman, M.A., & Chateldnat, G. (1975). Behavioural assessment in mental retardation. In P. McReynolds (ed.), Advances in Psychological Assessment. Vol. 3., San Francisco: Jossey-Bass Publishers.

Hebb, D.O. (1949). The Organisation of Behaviour. New York: John Wiley.

Hegarty, S. (1977). Test of Children's Learning Ability: Individual Version. Slough: NFER.

Hegarty, S. (1978). Manual for the NFER Test of Learning Ability (individual form). Slough: NFER.

Hegarty, S. & Lucas, D. (1978). Able to learn? The Pursuit of Culture-Fair Assessment. Slough: NFER.

Hilgard, E., & Atkinson, B. (1967). Introduction to Psychology. 4th ed. New York: Harcourt, Brace and World, Inc.

Hilgard, E.R., Atkinson, R.L. & Atkinson, R.C. (1979). Introduction to Psychology. 7th ed. New York: Harcourt Brace Jovanovich, Inc.

Hilgard, E.R. & Bower, G.H. (1975). Theories of Learning. 4th ed. Englewood Cliffs, New Jersey: Prentice-Hall.

Hill, D. (1976). Teaching in Multicultural Schools: A Guidebook. London: Methuen.

Hill, W.F. (1980). Learning: A Survey of Psychological Interpretations. 3rd ed. London: Methuen.

Home Office (1978). Special Educational Needs. Report of the Committee of Enquiry into the Education of Handicapped Children and Young People. London: HMSO.

Hudson, W. (1967). The study of the problem of pictorial perception among unacculturated groups. International Journal of Psychology, 2, 89-107.

Inhelder, B. & Piaget, J. (1964). The Early Growth of Logic in the Child. London: Routledge & Kegan Paul.

Inhelder, B., Sinclair, H. & Bovet, M. (1974). Learning and the Development of Cognition. Cambridge, Mass.: Harvard University Press.

Irvine, S.H. (1966). Towards a rationale for testing attainments and abilities in Africa. British Journal of Educational Psychology, 36, 24-32.

Jahoda, G. (1956). Assessment of abstract behaviour in a non-Western culture, Journal of Abnormal Social Psychology, 53, 237-243.

Jahoda, G. (1980). Sex and ethnic differences on a spatial-perceptual task: some hypotheses tested. British Journal of Psychology, 71, 425-431.

Jensen, A.R. (1961). Learning abilities in Mexican-American and Anglo-American children. California Journal of Educational Psychology, XII, 4, 147-159.

Jensen, A.R. (1963). Learning ability in retarded, average and gifted children. Merrill-Palmer Quarterly of Behaviour and Development, 9, 2, 123-140.

Jensen, A.R. (1967). The culturally disadvantaged: psychological and educational aspects. Educational Research, 10, 1, 4-20.

Jensen, A.R. (1969). Intelligence, learning ability and socioeconomic status. In S. Wiseman (ed.), Intelligence and Ability. 2nd ed. Harmondsworth: Penguin Books.

Jensen, A.R. (1970). Hierarchical theories of mental ability. In B. Dockrell (ed.), On Intelligence. Toronto, Ontario: Institute for Studies in Education.

Jensen, A.R. (1973). Level I and level II abilities in three ethnic groups. American Educational Research Journal, 10, 4, 263-276.

Jensen, A.R. (1974a). Interaction of level I and level II abilities with race and socioeconomic status. Journal of Educational Psychology, 66, 1, 99-111.

Jensen, A.R. (1974b). How biased are culture loaded tests? Genetic Psychology Monographs, 90, 185-244.

Jensen, A.R. (1975). Forward and backward digit span interaction with race and IQ: predictions from Jensen's theory. Journal of Educational Psychology, 67, 6, 882-893.

Jensen, A.R. (1979). The nature of intelligence and its relation to learning. Journal of Research and Development in Education, 12, 2, 79-95.

Jensen, A.R. (1980a). Précis of bias in mental testing. Behavioural and Brain Sciences, 3, 325-371.

Jensen, A.R. (1980b). Bias in Mental Testing. London: Methuen.

Jensen, A.R. & Inouye, A.R. (1980). Level I and level II intelligence. Intelligence, 4, 1, 41-49.

Johnson, O.G. (1976). Tests and Measurements in Child Development: Handbook II. Vol. 2. San Francisco: Jossey-Ban Publishers.

Jöreskog, K.G. (1971). Simultaneous factor analysis in several populations. Psychometrika, 36, 409-426.

Jöreskog, K.G. (1977). Structural equation models in the Social Sciences: specification, estimation and testing. In P.R. Krishnaiah (ed.), Application of Statistics. Amsterdam: North-Holland Publishing Company.

Jöreskog, K.G. & Sörbom, D. (1978). LISREL IV User's Guide. Chicago, Illinois: International Educational Services.

Kamii, C. (1974). Pedagogical principles derived from Piaget's theory: relevance for educational practice. In M. Schwebel & J. Ralph (eds), Piaget in the Classroom. London: Routledge & Kegan Paul.

Kee, D.W. & Rohwer, W.D. Jr. (1973). Noun-pair learning in four ethnic groups: conditions of presentation & response. Journal of Educational Psychology 65, 2, 226-232.

Kerlinger, F.N. (1973). Foundations of Behavioural Research. 2nd ed. London: Rinehart & Winston.

Kerlinger, F.N. (1979). Behavioural Research: A Conceptual Approach. Holt, Rinehart & Winston.

Kirk, S.A., McCarthy, J.J. & Kirk W.D. (1976). Examiner's Manual: Illinois Test of Psycholinguistic Abilities. Rev. ed. Urbana, Illinois: University of Illinois Press.

Kroeger, E. (1980). Cognitive development in the acculturation of migrant children: the role of training in the assessment of learning ability. International Review of Applied Psychology, 29, 1/2, 103-118.

Lambert, N.M., Wilcox, M.R. & Gleason, W.P. (1974). The Educationally Retarded Child. New York: Grune & Stratton.

Lawson, G., Baron, J. & Siegel, L.S. (1974). The role of number and length cues in children's quantitative judgements. Child Development, 45, 731-736.

Lerner, J.W. (1976). Children with Learning Disabilities: Theories, Diagnosis, Teaching Strategies. 2nd ed. Boston: Houghton Mifflin.

Lesser, G.S., Feifer, G. & Clark, D.H. (1965). Mental abilities of children from different social-class and cultural groups. In S. Wiseman (ed.) (1973), Intelligence and Ability. 2nd ed. Penguin Books.

Lewis, D.G. (1967). Statistical Methods in Education. London: University of London Press Ltd.

Lewis, D.G. (1968). Experimental Design in Education. London: University of London Press Ltd.

Lewis, D.G. (1974). Assessment in Education. London: University of London Press Ltd.

Little, A. (1975). Performance of children from ethnic minority backgrounds in primary schools. Oxford Review of Education, 1, 2, 117-135.

Little A. (1982). Swann: search for clues. TES, 23 July.

LLoyd, F. & Pidgeon, D.A. (1961). An investigation into the effects of coaching on non-verbal test material with European, Indian and African children. British Journal of Educational Psychology, 31, 145-151.

Longstreth, L.E. (1978). Level I - level II abilities as they affect performance of three races in the college classroom. Journal of Educational Psychology, 70, 3, 289-297.

Lunzer, E.A. (1973). Recent Studies in Britain Based on the Work of Jean Piaget. New ed. Slough: NFER.

Lunzer, E.A. (1981). Review symposium: Piaget re-examined. The Durham & Newcastle Research Review, IX, 46, 238-240.

Lunzer,E. & Dolan, T. (1977). Making Sense. Vol. 1 & 2. Report to the Social Science Research Council. Nottingham University, School of Education.

Mabey, C. (1981). Black British literacy: a study of reading attainment of London black children from 8 to 15 years. R.S. Report, Inner London Education Authority, 1-30.

Mackay, G.W.S. & Vernon, P.E. (1963). The measurement of learning ability. British Journal of Educational Psychology, 33, 177-186.

Mackenzie, A. (1980). Are ability tests up to standard? Australian Psychologist, 15, 3, 335-348.

MacKenzie, B. (1980). Fallacious use of regression effects in the IQ controversy. Australian Psychologist, 15, 3, 369-384.

Malmi, R.A., Underwood, B.J. & Carroll, J.B. (1979). The inter-relationships among some associative learning tasks. Bulletin of the Psychonomic Society, 13, 3, 121-123.

McLoughton, C.S. & Koh, T. (1982). Testing intelligence: a decision suitable for the psychologist? Bulletin of the British Psychological Society, 35, 308-311.

McReynolds, P. (1982). The future of psychological assessment. International Review of Applied Psychology, 31, 117-139.

McShane, D.A. & Plas, J.M. (1982). Wechsler Scale performance patterns of American Indian children. Psychology in Schools, 19, 1, 8-17.

Miller, S.A. (1976). Non verbal assessment of Piagetian concepts. Psychological Bulletin, 83, 405-430.

Mittler, P.J. (ed.) (1970). The Psychological Assessment of Mental and Physical Handicap. London: Methuen.

Modgil, S. (1974). Piagetian Research: A Handbook of Recent Studies. Windsor: NFER Publishing Company Ltd.

Modgil, S. & Modgil, C. (1976). Piagetian Research: Compilation and Commentary. Vol. I-VIII. Windsor: NFER Publishing Company Ltd.

Modgil, S. & Modgil, C. (1976). Piagetian Research: Compilation and Commentary. Vol IV. Windsor: NFER Publishing Company Ltd.

Modgil, S. & Modgil, C. (1976). Piagetian Research: Compilation and Commentary. Vol. VIII. Windsor: NFER Publishing Company Ltd.

Modgil, S. & Modgil, C. (1980). Toward a Theory of Psychological Development. Windsor: NFER Publishing Company Ltd.

Newland, T.E. (1973). Assumptions underlying psychological testing. Journal of School Psychology, 11, 4, 316-322.

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K. & Bent, D.H. (1975). Statistical Package for the Social Sciences. 2nd ed. New York: McGraw Hill Book Company.

Nunnally, J.C. (1978). Psychometric Theory. 2nd ed. New York: McGraw.

Oleron, P. (1977). Language and Mental Development. New Jersey: Lawrence Erlbaum Associate, Publishers.

Ortar, G.R. (1960). Improving test validity by coaching. Educational Research, 2, 3, 137-142.

Ortar, G.R. (1963). Is a verbal test cross-cultural? Scripta Hierosoly-mitana, 13, 219-235.

Pallack, M. (ed.) (1981). American Psychologist, 36, 10. Washington: APA.

Paraskevopoulos, J.N. & Kirk, S.A. (1969). The Development and Psychometric Characteristics of the Revised Illinois Test of Psycho-linguistic Abilities. Urbana: University of Illinois Press.

Phillips, C.J. (1979). Educational underachievement in different ethnic groups. Educational Research, 21, 116-130.

Piaget, J. (1952). The Child's Conception of Number. London: Routledge & Kegan Paul Ltd. Translated by Gattegno, C.E. & Hodgson, F.M. from the first French edition published in 1941.

Piaget, J. (1971). The theory of stages in cognitive development. In D.R. Green, M.P. Ford & G.B. Flames (eds), Measurement and Piaget. New York McGraw-Hill Book Company.

Piaget, J., Inhelder, B. & Szeminska, A. (1960). The Child's Conception of Geometry. London: Routledge & Kegan Paul.

Piaget, J. & Inhelder, B. (1969). The Psychology of the Child. New York: Basic Books.

Price-Williams, D.R. (ed.) (1969). Cross-Cultural Studies. Penguin Books.

Raven, J.C. (1962). Guide to Using The Coloured Progressive Matrices. London: H.K. Lewis & Company Ltd.

Raven, J.C., Court, J.H. & Raven, J. (1978). Manual for Raven's Progressive Matrices and Vocabulary Scales. London: H.K. Lewis & Company Ltd.

Resnick, L.B. & Resnick, D.P. (1982). Testing in America: the current challenge. International Review of Applied Psychology, 31, 75-90.

Rohwer, W.D. Jr., Ammon, M.S., Suzuki, N. & Levin, J.R. (1971). Population differences and learning proficiency. Journal of Educational Psychology, 62, 1, 1-14.

Salvia, J. & Yesseldyke, J.E. (1978). Assessment in Special and Remedial Education. Boston: Houghton-Mifflin.

Sattler, J.M. (1974). Assessment of Children's Intelligence. Philadelphia: W.B. Saunders and Company.

Scar, S., Caparulo, B.K., Ferdman, B.M., Tower, R.B. & Caplan, J. (1983). Developmental status and school achievements of minority and non-minority children from birth to 18 years in a British Midlands town. British Journal of Developmental Psychology, 1, 31-48.

Schaffer, H.R. (1977). Studies in Mother-infant Interaction. London: Academic Press.

Schonell, F.J. (1951). Diagnostic and Attainment Testing. Edinburgh: Oliver & Boyd.

Schools Council (1970). Immigrant Children in Infant Schools. (Working Paper 31.) London: Evans/Methuen.

Schwarz, P.A. (1961). Aptitude Tests for Use in Developing Nations. Pittsburgh: American Institute for Research.

Schwebel, M. & Raph, J. (1974). Piaget in the Classroom. London: Routledge & Kegan Paul.

Scott, G.C. (1950). Measuring Sudanese intelligence. British Journal of Educational Psychology, 20, 1, 43-54.

Siegel, L.S. (1977). The cognitive basis of the comprehension and production of relational terminology. Journal of Experimental Child Psychology, 24, 40-52.

Siegel, L.S. (1978). The relationship of language and thought in the preoperational child: a reconsideration of non-verbal alternatives to Piagetian tasks. In L.S. Siegel & C.J. Brainerd, Alternatives to Piaget: Critical Essays on the Theory. New York: Academic Press.

Siegel, L.S. & Brainerd, C.J. (1978). Alternatives to Piaget: Critical Essays on the Theory. New York: Academic Press.

Silvey, J. (1963). Aptitude testing and educational selection in Africa. Rhodes Livingstone Journal, 34, 9-22.

Sinclair, H. (1981). Piaget on language: a perspective. In S. Modgil & C. Modgil (eds) Jean Piaget: Consensus and Controversy. London: Holt, Rinehart & Winston.

Southgate, V. (1959). Southgate Group Reading Tests, Manual of Instructions: Tests 1 (Forms A, B and C) - Word Selection Test. London: University of London Press.

Spencer, D. (1982). Staying on helps blacks to exam success. TES, 8 October.

Stake, R.E. (1961). Learning parameters, aptitudes and achievements. Psychometric Mon. No. 9.

Stevenson, H.W. (1972). The taxonomy of tasks. In F. J. Monks, W.H. Hartup & J. De Wit, (eds), Determinants of Behavioural Development. New York: Academic Press.

Sullivan Associates (1969). Programmed Reading. St. Louis: Webster-McGraw-Hill.

Taylor, J.H. (1973). Newcastle-upon-Tyne: Asian pupils do better than whites. British Journal of Sociology, 24, 4, 431-447.

Telford, C.W. & Sawrey, J.M. (1977). The Exceptional Individual. 3rd ed. New Jersey: Prentice-Hall Inc.

Terman, L.M. & Merrill, M.A. (1960). Standford-Binet Intelligence Scale: Manual for the Third Revision Form L-M. London: Harrap.

Thorndike, R.L. & Hagen, E.P. (1969). Measurement and Evaluation in Psychology and Education. New York: John Wiley & Sons Inc.

Thorndike, E.L. & Lorge, I. (1944). The Teacher's Word Book of 30,000 Words. New York: Teacher's College, Columbia University.

Thurstone, L.L. (1947). Multiple Factor Analysis. Chicago: University of Chicago Press.

Toepher, C. (1981). No details about the title etc. In C. Cookson, Learning is linked to spurts in development of brain. TES, 15 May.

Tomlinson, S. (1981). Special Education. London: Harper & Row.

Townsend, H.E.R. (1971). Immigrant Pupils in England. Slough: NFER.

Townsend, D.J. (1974). Children's comprehension of comparative forms. Journal of Experimental Child Psychology, 18, 293-303.

Townsend, H.E.R. & Brittan, E.M. (1972). Organisation in Multiracial Schools. Slough: NFER.

Tucker, J.A. (1980). Ethnic proportions in classes for the learning disabled issues in non-biased assessment. Journal of Special Education, 14, 1, 93-105.

Tyrnchuk, A.J. (1973). Effects of concept familiarization vs. stimulus enhancement on verbal abstracting in institutionalized retarded delinquent boys. American Journal of Mental Deficiency, 77, 551-555.

Vernon, P.E. (1960). Intelligence and Attainment Tests. London: University of London Press.

Vernon, P.E. (1969). Intelligence and Cultural Environment. London: Methuen.

Vernon, P.E. (1972). The distinctiveness of field independence. Journal of Personality, 40, 366-391.

Vernon, P.E. (1979). Intelligence: Heredity and Environment. San Francisco: W.H. Freeman & Company.

Vernon, P.E. & Miller, K.M. (1976). Graded Arithmetic-Mathematics Test. Metric ed. Seven Oaks: Hodder and Stoughton Educational.

Wadsworth, B.J. (1978). Piaget for the Classroom Teacher. New York: Longman.

Wagner, D.A. (1981). Culture and memory development. In H.C. Triandis & A. Heron (eds), Handbook of Cross-Cultural Psychology: Developmental Psychology. Vol. 4. Boston: Allyn and Bacon Inc.

Wechsler, D. (1949). The Wechsler Intelligence Scale for Children. New York: The Psychological Corporation.

Wechsler, D. (1974). Wechsler Intelligence Scale for Children. Rev. NFER-Nelson Publishing Company Ltd.

Winer, G.A. (1980). Class inclusion reasoning in children. A review of the empirical literature. Child Development, 51, 309-328.

Winter, W. (1963). The perception of safety posters by Bantu industrial workers. Psychology Africana, 10, 127-135.

Wober, M. (1967). Adapting Witkin's field independence theory to accommodate new information from Africa. British Journal of Psychology, 58, 29-38.

Woodrow, H. (1938a). The effect of practice on test intercorrelations. Journal of Educational Psychology, 29, 268-278.

Woodrow, H. (1938b). The relation between abilities and improvement with practice. Journal of Educational Psychology, 29, 215-230.

Woodrow, H. (1939a). Factors in improvement with practice. Journal of Educational Psychology, I, 55-70.

Woodrow, H. (1939b). The application of factor analysis to problems of practice. Journal of General Psychology, 21, 457-460.

Woodrow, H. (1939c). The relation of verbal ability to improvement with practice in verbal tests. Journal of Educational Psychology, 30, 179-186.

Woodward, W.M. (1970). The assessment of cognitive processes: Piaget's approach. In P. Mittler (ed.), The Psychological Assessment of Mental and Physical Handiaps. London: Tavistock Publications.

Woodworth, R.S. & Scholosberg, H. (1954). Experimental Psychology. Rev. ed. Calcutta: Oxford & IBH Publishing Company.

Youniss, J. & Furth, H.G. (1973). Reasoning & Piaget. Nature, 244, 314-315.

Ysseldyke, J.E. & Salvia, J. (1974). Diagnostic prescriptive teaching: two models. Exceptional Children, 41, 181-185.