The University of Hull

# Standardization of the WISC-R for Students Aged 12-15 Years in Iraq

Being a Thesis Submitted in Fulfilment

of the Degree of

Doctor of Philosophy

in the University of Hull

by

ADNAN GHAIB RASHED

M.Ed. Utah State University

1990

## Abstract

The main aim of this research was to provide the educational system in Iraq with a test of intelligence which could have been standardized on a large sample of 12 to 15 year- old students throughout that country. Rather than construct a new intelligence test it was decided to use a well-established test, the Wechsler Intelligence Scale for Children Revised, the WISC-R, which had been previously standardized in both the U.S.A. (where it originated) and Britain and in many other countries throughout the world. It was considered necessary to modify the form of this test before it could be standardized and made available to some of the Arab countries.

The British version of the test (WISC-R, 1974), after initial piloting, was administered to 800 students equally divided among the four age groups - 12 years, 13 years, 14 years and 15 years. The sampling procedure ensured that the selection of students represented the distribution of secondary-school population in the three main regions of Iraq, as well as providing representative samples from both urban and rural communities. For standardization purposes it was decided to have equal numbers of boys and girls in each year- group.

The raw scores on each subtest of the WISC-R within each age group were scaled to give a mean of 10 and Standard Deviation of 3, which was the procedure used in the original standardizations of the WISC-R. In a similar way the overall Verbal, Performance and Full Scale scores, based on the appropriate subscales, were scaled to give a mean of 100 and Standard Deviation of 15. The intercorrelations of all the subscales were compared with those obtained for the British WISC-R. The results in the present study have similar magnitudes. High values of correlation coefficients were obtained for all subscales separately and

with the measures of Verbal, Performance and Full Scale IQs. These again were comparable to the results of previous studies.

Finally, the study reported on the variables 'socioeconomic status', 'education of parents', 'urban-rural status', 'size of family' and 'regional area of Iraq'. Using t- analyses and analyses of variance it was found that many of these relationships for Iraqi students were similar to these found in previous studies on IQ.

## Acknowledgements

First I am exceedingly grateful to my supervisor, Dr. W. J. Wilkinson, who acted as such a generous host during my study years in the Institute of Education, University of Hull. I am grateful not only for his many useful discussions and countless hours supervising my work, but also for his knowledge in the field of assessment, his useful suggestions and helpful comments for the improvement of my approach to the study in hand. I shall not readily forget his kindness, patience and interest in my research.

I wish to express my gratitude to the Iraqi Government, for the opportunity given to me to study for a Ph.D. degree, and for the psychological and financial support which they provided me, especially the Ministry of Education.

Also I feel fortunate in having the assistance of the Educational and Psychological Research Centre, University of Baghdad, the Research Centre, Ministry of Education, for understanding of the nature, scope and problems of my study. My thanks also go to Dr. D. Bartram, Psychology Department, University of Hull, for his helpful comments.

Special thanks goes to my wife, Souha Hamed, for her care, encouragement and sacrifice in making all my academic aims possible for me. My thanks also go to my mother for everything she has done for me.

Mrs. M. Jack also examined the use of the English language in my thesis in detail and I have incorporated many of her useful suggestions with regard to the clear espression of my meaning

## Tables

Tables Page

## Table of Contents

**Chapter One**



**Chapter Two**

**Chapter Three**



**Chapter Four**



**Chapter Five**



**Chapter Six**

## Table of Contents

# CHAPTER ONE

# Intelligence

## CHAPTER ONE

# CHAPTER ONE

## 1.0 Introduction.

Traditionally, individual intelligence testing has played an important role in the intellectual assessment of school age children. Criteria for the designation of handicapping conditions include a specified level of intellectual functioning as measured by individually administered intelligence tests (Thurstone, 1931).

Norm reference intelligence (IQ) tests are useful in the evaluation of intelligence, giving valuable data about an individual's level of functioning when compared to others in the same peer group. These tests provide an efficient method of sampling behaviour in a few hours, and as an integral part of the diagnostic process also enable children to obtain special programmes that can help remedy learning difficulties (Sattler, 1982).

The difficulties presented by the use of standardized tests to assess the abilities and intelligence of children whose language and cultural backgrounds differ from those in the standardization sample have long been recognized. Dickinson (1937) warned that a test score is valid only to the extent that the items of the test are as familiar to each child tested as they were to the children upon whom the norms were based (p. 522).

Investigation of the effectiveness of education, or indeed any research which involves assessing children's cognitive skills, requires reliable measures of achievement or ability. Verbal and conceptual understanding can be assessed and analysed by tests such as the Wechsler Intelligence Scale for Children-Revised (WISC-R), the Stanford-Binet and the British Ability Scales (BRS).

The use of intelligence scales or tests in many countries is relatively new. In Iraq there has been a recent trend towards developing intelligence and achievement scales for the purpose of assessing children in educational contexts. As far as the author knows, there is only one instrument, consisting of intelligence scales, that has been translated and standardized for use with Iraqi children. Al-Aubaidy (1987) completed a pre-liminary school standardization of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) (the Performance part only).

The WISC-R was selected as a measure of IQ in this study for several reasons. It is widely used by school psychologists in testing children (Mercer, 1979). Minority children, including some bilingual children, were proportionally represented in the standardization of the WISC-R (Wechsler, 1974). And the internal validity of the WISC-R was demonstrated to be similar for English-speaking and Minority children (Sandoval and Mille, 1980; Gutkin, and Reynolds,1980).

The WISC-R yields IQ scores for Verbal (VIQ) Performance (PIQ) and Full Scale (FSIQ). Several researchers have found that bilingual children are penalized on verbal but not on non-verbal tests (Sattler, 1982; Bernstein, 1961). Moreover, the WISC-R was found to correlate with achievement for learning-disabled students (Anastasi, 1961).

It was realised from the beginning that the more extreme the cultural differences, the more difficult would be the task of identifying common problems such as those concerning sex, occupation and education, particularly if it is recognized that a cognitive problem which may appear on superficial examination to be identical in two cultures may, through differences in social structure, motivation, attitude and the like, actually be quite different when seen from the point of view of the people involved. It is well known that the experience of a child is bound by his culture and that an intelligence test cannot be equally fair to populations with different upbringings.

## 1.1 The Statement of The Problem

It was the purpose of this study to translate the Wechsler Intelligence Scale for Children-Revised (WISC-R), with British amendments into the Arabic language and to adapt it to an Iraqi cultural milieu within the state school population. In order to ensure that the instrument developed for this project and the English WISC-R were comparable, efforts were made to follow the same procedures as those employed by the WISC-R publisher, (the British version) and the Psychological Corporation, in their standardization of the original instrument. Norms for Iraqi children were to be established.

In addition, the research aimed to give evidence on: a) The relationship between intelligence and socio-economic status; b) The relationship between intelligence and educational level of father and mother. c) The relationship between intelligence and family size; and c) the relationship between intelligence and urban--rural location.

## 1.2 Significance of the Study

This study hopes to provide the educational system in Iraq with locally standardized tests of intelligence and aptitude for school children aged 12-15 years. These tests can be used to help psychologists and educators in the identification of children who may later experience difficulties or lack of challenge in different areas of the school curriculum.

Intelligence is held to be a vital factor in describing human thinking. The earliest philosophies made distinct assumptions about the construct 'intelligence', regarding it as either a single or a multiple quality and defining it as either changeable or fixed and immutable (Evans and Waites, 1981). Eysenck (1979) briefly traced the contributions of early philosophies and believed that the psychometric movement had its roots in biology and physiology, and that these same philosophies viewed intelligence as an 'innate, all-round cognitive ability'.

Pintner (1923) concurred with this view in his writings and suggested that the psychometric movement developed as a contribution to the fight for social justice for all classes of individuals, especially the mentally retarded and the feeble-minded. Carrol (1982) on the other hand, feels that the climate of the times has been more influenced by Darwin's formulation of evolution than stirred by an attack of social conscience. These two authors have credited Sir Francis Galton with the introduction of the notions of the experimental method and measurement, in the study of intelligence. In the first sentence of his book Hereditary Genius (1869) Galton stated the premise of his belief concerning mental abilities: "I propose to show in this book that a man's natural abilities are derived by inheritance, under exactly the same limitations as are the form and physical features of the whole organic world". (p.1). Hereditary Genius was the study of a small group of leading families in Great Britain which, according to Galton, produced men of distinction. He firmly believed that, since genius tended to run in families, eminent fathers produced eminent sons,and hence, mental traits must be inherited. It is interesting to note that Galton made no mention of any female descendants in these families. A man of his time, Galton shared the general view that women were inferior in mental capacity and ability.

Although Galton (1869) never formally defined intelligence, he acknowledged its value and necessity. He stated that:

> Civilization is the necessary fruit of high intelligence when found in social animal and there is no plainer lesson to be read off the face of nature than that the result of the operation of her laws is to evoke intelligence in connection with sociability. Intelligence is as much an advantage to an animal as physical strength or any other natural gift, and therefore, out of two varieties of any race of animal who are equally endowed in other respects, the most intelligent variety is sure to prevail in the battle of life. (p. 336)

Butcher (1968) and DuBois (1970) credited Galton with devising many tests of simple sensory and motor functions. In 1882 Galton established an anthropometric laboratory at the South Kensington Museum, London, where people could learn of their capabilities through certain tests for a price of three pence. Among the techniques Galton employed for measurement were "several purely physical measures, keenness of eyesight and hearing, colour-sense and highest audible note, dynamometer pressure, reaction-time and errors in dividing a line and angles". (Galton, 1883). Gould (1981) stated that Galton employed anthropometrics measuring people's skills at his laboratory in London, and that he seriously believed that the size of skull circumference was an indicator of intelligence.

Cattell was responsible for introducing the Galton tradition of testing into the United States. He had been an assistant in Galton's London laboratory and expressed great interest in individual differences. Cattell (1890) advocated that tests be given in schools and initiated "the collection of physical and mental measurement of students in Cambridge University, the University of Pennsylvania, and Bryn Mawr College in 1887-1888" (Cattell and Farrand, 1896, p.20). Perhaps the most famous of Cattell's experiments was conducted by Cattell and Farrand (1896) with 100 students from the University of Columbia. This elaborate programme produced much data regarding student ability, and directed research in two areas:the interrelation of measured traits, and the interdependence of the senses and the mind. Cattell shared Galton's view of intelligence as a general capacity of cognition.Underlying Cattell's psycho-physical approach to the measurement of intelligence was the assumption that physical tests measure mental ability. (Stenberg and Powell, 1983). An alternative approach for defining the construct of intelligence was developed by Binet.

Alfred Binet is best known for his construction of the early forms of the Binet-Simon Scale and for its contribution to the development of psychometrics. Perhaps Binet's greatest contribution was his discovery that tests were needed to measure the more complex tasks and mental abilities used in everyday life. (Anastasi, 1976; Burt, 1955).

In 1892 Alfred Binet began to work in the laboratory of physiological psychology at the Sorbonne. As both a worker and--later--director of the faculty, Binet experimented with young children and searched for relationships between individual differences, age and grade--placement. He quickly lost his early fascination with cranial measurement for developing a definition of intelligence level, realizing that no major development was possible through this avenue of study, and turned his attention to the construction of other assessment techniques.

Binet developed a description of intelligence and concentrated upon a psychological method of measurement "because it aims to measure the state of intelligence as it is at the present moment" (Binet and Simon 1916/1973, p. 40). The Minister of Public Instruction in Paris appointed Binet to a Commission whose duty was to identify and provide instructional material for children who did not profit from regular classroom instruction. In writing about the Commission, Binet and Simon (1916/1973) stated that "our purpose is to be able to measure the intellectual capacity of a child who is brought to us in order to know whether he is normal or retarded." Binet made a clear distinction between intelligence and the effects of schooling. The tests constructed by him attempted to limit the degree of instruction to which the subject had been exposed; the student was given nothing to read or to write in these tests in order to control for effects of learning.

In 1905 Binet and his colleague Theodore Simon constructed their first formal scale for appraising intelligence abilities. According to Binet (1916/1973) the tests started "from the lowest intellectual level that can be observed" (p.40)and ended with "an average, normal intelligence", and explained a "hierarchy among divers intelligence." (p.41). For Binet and Simon this classification was equivalent to measured intelligence.

Binet believed that his scale was an instrument which would both identify intelligence and distinguish between the two kinds of intelligence "maturity" and "correctness." He and Simon (1916/1973) defined intelligence as the ability" to judge well, to comprehend well, to reason well" (p.43). An important but relatively ignored outcome of this intelligence testing which arose from their reviewing of the results, was their identification of sections of the test which needed revision and clarification.

In 1911 Binet and Simon revised their tests first revision. Anastasi (1976) listed these modifications as consisting of: five tests being provided for each level from three to ten years of age; new tests added for ages 12-15 years and adult levels; and a minor adjustment in scoring. The procedures established by Binet and his staff for individual testing took place alongside assessment instruments for his continuing experiments on intelligence.

Charles Spearman, a contemporary of Alfred Binet, formulated a theory of general intelligence that attempted to go beyond mere description. (Pintner, 1923). While in the army in England, Spearman collected data on Hampshire village school--children which served as a basis for two articles published in 1904. The first paper, "The Proof and Association Between Two Things", (in the, American Journal of Psychology) introduced a new concept--'correlation' between test score and academic standing. Spearman developed a method of ranking, and listed the following as its greatest advantage: "From the practical point of view, it is so urgently desirable to obtain the smallest probable error with a given number of subjects that the method of rank must often have preference when we are dealing with two series of measurements properly comparable with one another." (Burt, 1955).

The second article, "General Intelligence Objectively Determined and Measured, "(1904) presented Spearman's two-factor theory of intelligence (Eysenck and Kamin, 1981; Burt 1955). He originally classified four different kinds of intelligence: present efficiency, native capacity, general impression and common sense (pp. 250--251),and investigated them empirically. He termed the two classes of elements common to the four categories "general intelligence" and "general discrimination" and observed that all branches of intellectual activity have in common one fundamental group of functions, whereas the remaining or specific elements of the activity seem in every case to be wholly different from those in all the others.

DuBois (1970) explained that Spearman's theory of general intelligence involved " a statement about the nature of intelligence from which could be deduced types of tests that would measure it; and mathematical procedures that could be used to support or reject the theory" (p.46).

In 1921 the Journal of Education and Psychology published an article, "Intelligence and its Measurement: A Symposium" in which fourteen educational psychologists attempted to define the word 'intelligence'. The widespread construction and use of group and individual intelligence tests and the lack of a consensus on the definition of what was tested, made clarification a difficult task. Responses from those participating in the symposium fell into two major categories: intelligence orientations and behavioural orientations. The intellectual focus was on aspects of mental processes and knowledge.

Thorndike (1921) defined intelligence as "the power of good responses from the point of view of truth or fact," (p. 124), and he felt that intellectual capacity was general, while Terman (1921, p.128) believed that" an individual is intelligent in proportion as he is able to carry on abstract thinking". For Terman, intelligence was not solely the ability to pass a given test but skills which extended into the real life of each person. For Henmon (1921, p.196) intelligence was "indicated by" the capacity to appropriate truth and fact as well as the capacity to discover them" . Woodrow (1921, p.208) acknowledged it as a "growing thing...fixed partly by heredity and partly by environment factors acting before the age of five," while to Haggerty, (1921) intelligence was just a "group of complex mental processes." (p. 212).

The paucity of statistical methods remained a problem in the psychometric movement until Leon Thurstone (1931) elevated psychological testing to the rank of a true science. His major contribution to the field was in the area of 'factor analysis' and correlation, for Thurstone believed that" a variable can be broken into a number of additive components representing underlying characteristics." (DuBois 1970, p.117).

Thurstone's view of intelligence was stated in the 1921 Symposium referred to earlier, and was further developed in The Nature of Intelligence (1927). In his opinion, intelligence was"a capacity to trial - and - error existence with alternatives, it is 'the capacity to acquire the capacity'.

Resnick (1976) concurred on the value of Thurstone's research on intelligence, and his achievements in the area of data analysis. Eysenck and Kamin (1981) devoted considerable space to the Spearman-Thurstone controversy wherein Thurstone questioned that Spearman's general ability factor really measured a number of primary abilities. In 1938 Thurstone published his monograph, Primary Mental Ability, and listed seven independent factors of intelligence: V-visual, P-perceptual, N-numerical, S-spatial, M-memory, I-individual, and D-deductive. Thurstone's meticulous collection of data, along with Spearman's further empirical studies, produced a paradigm of general and specific abilities that endures today. (Eysenck, 1981).

A new approach to the study of intelligence which reflects the ideas of Thurstone is seen in the work of Guilford (1964). Guilford proposed an extension of Thurstone's theory which divided further the primary mental abilities and increased their number from seven (visual, perceptual, numerical, memory, individual and deductive) to 120 independent measurable abilities which he belived were required to describe human capacity. (Sternberg and Powell, 1983). Eysenck (1979) explained that Guilford's 120 independent abilities were not hierarchical.

Guilford's system starts by classifying possible kinds of ability under three headings. Abilities may vary according to: (a) the basic psychological processes involved, which are cognition, memory, evaluation, divergent production and convergent production; (b) the kind of materials or processed, such as symbols (e.g. letters, numbers and words, when meaning as such is not considered), and (c) semantic (meaningful) materials such as classes, systems, relations or transformations.

The possible processes or operations according to Guilford and Hoepfner (1966) are five in number, the kinds of content four, and the kinds of product six. Since these are independent cross--classifications, this system yields a large number of possible different abilities, i.e. 5 X 4 X 6 = 120.

This way of classifying abilities according to logical principles is similar to the ideas suggested independently by Guttman (1965), although this writer would not necessarily agree with Guilford's actual bases of classification.

Throughout the psychometric movement, a recurrent theme in regard to intelligence has been the indications of a two-factor model with the factors inter-related yet independent of each other .It was Cattell who first presented his two-factor theory of intelligence (Cattell, 1971). Cattell called these sub-factors "fluid" and "crystallized", and defined them thus: "Fluid intelligence is that power unconnected with culture skills, which rises at its own rate and falls despite cultural stimulus, being adapted to almost any problem." (p. 80), while "Crystallized intelligence is invested in particular areas of crystallized skills which can be upset individually without affecting the others.The former represents the basic intellectual capacity of the organism while the latter represents the abilities acquired by learning and education."

Horn and Cattell (1966) explained the relationship between intelligence testing and "fluid" and "crystallized" intelligence. "Fluid" ability is best measured by subtests which require mental manipulation of symbols, such as series completion and figural, analogical and classification problems. Tests which measure crystallized ability include reading comprehension,vocabulary,and general information tests which reflect cultural milieu.

This development of fluid and crystallized intelligence appears to support the view of intelligence expressed by Jensen (1967). He explained fluid intelligence as the product of heredity and governed by the laws of biological maturation. and, he also believed that crystallized intelligence was primarily a product of the environment and socio-economic status. Stenbery (1983) and his colleague expressed strong agreement with the two sub--factors of fluid and crystallized intelligence and declared that they play an integral role in the hierarchical model of intelligence.

Jensen (1970) represented his own view of intelligence which distinguished between two types of intelligence: "Level I" ability was measured by standard memory tests, while "Level II" ability was measured by tests of comprehension.Jensen equated "Level I" with the capacity to receive, store, and later recognize material or outside stimuli. It appeared that "Level I" ability required no elaboration or transformation of the input for accurate remembering. On the other hand, Level II ability required more manipulation of the input received and involved the following two processes:

1-encoding and decoding of the stimulus in terms of

past experience,and

2-relating new learning to old learning (pp.156-157).

Jensen (1970) stated that the two types of intelligence increased at different rates."Level I" ability increased up to the age of four, and then decreased with age.On the other hand, the "Level-II" growing period occurred between the ages of four and eight. In assessing Jensen's conception of ability Sternbery and Powell (1983) believed that the Level--concept was useful for testing basic learning ability. However, "Level-I" correlated less well with IQ tests than "Level II" .

Piaget's intelligence theory deals primarily with the thinking, reasoning, and problem-solving processes of children from birth through adolescence. It postulates sequential development through four levels: sensory motor, preoperational, concrete, and formal operational. According to Piaget and Inhelder (1969) this sequence is invariant through the age of acquisition, and the rate of development varies from culture to culture. Piaget (1976) also elaborated upon the influence of four factors on cognitive development: maturation, equilibration, social and interpersonal co--ordination, and educational and cultural transmission. Since cognitive development takes place through an interaction between an individual and his environment, the influence of general as well as specific socio-cultural factors needs to be examined.

Some psychologists such as Neisser (1979) have studied people's conceptions of intelligence. According to Neisser, an intelligent person is:

> a prototypê--organized Roschian concept. Our concept that a person deserves to be called "intelligent" depends on that person's overall similarity to an imagined prototype, just as our confidence that some object is to be called "chair" depends on its similarity to prototypical chair. There are no definitive criteria of intelligence, just as there are none for chairiness; it is a fuzzy--edged concept to which many features are relevant. Two people may both be quite intelligent and yet have few traits in commonthey resemble the prototype along different dimensions. Thus, there is no such quality as intelligence, any more than there is such a thing as chairiness--resemblance is an external fact and not an internal essence. There can be no process--based definition of intelligence, because it is not a unitary quality, it is a resemblance between two individuals, one real and the other prototypical. (p. 185)

Neisser (1979) has noted that he is not the first to express such a view, which he has traced back to Thorndike (1920), who suggested:

> For a first approximation,let intellect be defined as that quality of mind (or brain or behaviour if one prefers) in respect to which Aristotle, Plato, Thucydides,and the like, differed most from Athenian ideas of their day, or in respect to which the law yers, physicians, scholars, and editors of reputed greatest ability at constant age, say a dozen of each, differ from idiots of that age in asylums. (p.126)

Neisser has suggested that tests such as the Stanford-Binet are reasonably successful because they consist of large numbers of items that assess resemblance to different aspects of the prototype. Individual items function like individual dimensions of a chair in the construction of a prototype.

Neisser (1979) collected informal data from Cornell undergraduates regarding their conceptions of what intelligence is. More formal studies were conducted by Cantor (1978), who asked adult subjects to list attributes of a bright person, and by Bruner, Shapiro, and Tagiuri (1958), who asked people how often intelligent people also display certain other personality traits. These authors found, for example, that intelligent people are likely to be characterized as clever, deliberate, efficient, and energetic, but not as apathetic, unreliable, dishonest, and dependent.

## 1.3 The Effects Of Race on Intelligence Testing.

The role of race in accounting for differences in IQ test scores continues to puzzle diagnosticians. When IQ tests are interpreted as absolutes, the variances in scores due to race can point to racial differences in aptitude. Ambiguity surrounds the question of race but it can be resolved through research. There are certainly striking differences between people of different countries, cultures and tribes in the way they tackle intellectual tasks and in their effectiveness in solving particular kinds of problems.

Scriven (1970) defended Jensen and his right to research racial differences and emphasized that such questioning is valid and should not be assumed to derive from a racist attitude. This position (of Jensen's race differences) is not shared by Chomsky (1972), who stated that the question of race and intelligence has very little scientific importance. Chomsky (1972) saw the race-intelligence issue as having no value for society unless that society is a racist one.

Compounding the question of differences in racial aptitudes is the practical consideration of how these differences would be measured if they do exist. It has become clear that measurement in the terms of the core culture can present the best indicators of the cognitive abilities of the racially different child.

The lines demarcating the races are vague and overlapping; the definition of such lines has an arbitrary basis which eludes scientific validation. In research methods, racial differentiations cannot easily be separated from economic distinctions, because they are often overlapping variables.

In spite of such practical hindrances to research on racial differences, Jensen (1972, p.331) has asserted that IQ tests do present a valid construct which shows weaker intellectual capacities in specific groups and classes.

According to research by the National Institute of Health (U.S.A.), (1979) as many as 20 to 30 percent of the black children in some of the largest urban centres suffer severe psychological handicaps. However, the U.S. Government has not supported, does not and will not, support the idea that any genetic factors are involved in this differential rate of mental handicap.

Deutsch (1973, p. 248) suggested a method for researching the race issue and separating it from the question of socioeconomic class. Through collecting data not within racial groups, but across classes, a relationship could be defined within each group with respect to socioeconomic class. Another method prescribed by Deutsch (1973) is to sample within just one racial group and to generalize only on that group.

Deutsch's methods have been used and have produced data of interest. Gottesman's report (1972) to the U.S. Senate Select Committee on "Equal Education" Opportunity, on the relationship between IQ and Social Economic Status (SES) indicated that race is not an influential variable. A study by Gottesman (1972) had tested the performance of white and black seven--year--old children in Boston.

Gottesman (1972) also presented evidence from a study completed by Nichols, which tested the performance of white and black seven--year--old children from Boston (U.S.A.) and Philadelphia. The sample comprised 5256 white and 4613 black children from these cities. The differences found in IQ scores were between cities, rather than between races.

In an early study of the racial variable, Reschly (1979) tested 40 southern black children who were not retarded, in the judgment of observers, but whose mean Full Scale WISC IQ was 69.8. Reschly followed up these results by testing a larger group of southern black, rural children. His researches found a markedly low mean WISC Full Scale IQ score of 67.7 and be concluded that "We must question whether the WISC is a suitable test for the southern Negro child".

Pursuing the within--race method, Young, (1984) reported the results of a longitudinal study of 89 black children from different social classes. It found no significant class differences on the Cattell Infant Intelligence Scale at 18 and 24 months of age, but a highly significant 23 point IQ difference on the Stanford--Binet at three years of age in children from welfare and middle--class black families. This study suggests that by three years of age,class differences have to be considered in interpreting the results of IQ tests.

When the effects of social class were removed from the studies reported in this review,the significance of the racial issue in testing lost import. Regional differences, according to Kaufman (1981) do complicate the performance of blacks on the WISC-R.

> Whereas the IQs of Whites are not a function of geographic region,the 10 point discrepancy in the FSIQs for Blacks from the Northeast vs. Blacks from the south is striking....the mean IQ for Northeastern Blacks,for example,is only slightly below the mean of 100 and close to the mean IQ of Whites from the North Central and South.

The inclusion of 'minority' children in the WISC-R standardization did not resolve the question of test bias. When the same groups of 'minority' children were administered the WISC and WISC-R. (The first Intelligence test published by David Wechsler was in 1939 and known as the Wechsler-Bellevue. It was designed primarily as a test for adults. In 1949 Wechsler produced the Wechsler Intelligence Scale for Children (WISC). It consists of 10 basic subtests, five "Verbal" and five "Performance". The WISC test covers the age range 5-16 years). A comparison between the two tests showed that the WISC-R scores were lower. Weiner and

Kaufman (1979) explained that the lower IQs on the WISC-R reflected the more advanced intellectual status of children who make up the new WISC-R norms when compared with the status of children a generation ago.

Wechsler's co-worker, Matarazo, (1972) noted that

> The problems of "Whites vs. non-Whites" were circumvented in the Wechsler Bellevue, the test from which the Wechsler WISC-R and the Wechsler Adult Intelligence Scale-Revised (WAIS-R) were devised.

Non-white subjects were omitted because it was felt at the time that norms derived from a mixed population could not be interpreted without special provisos and reservations.This appears now to have been an unnecessary concern, first because the admission of the non-white subjects in the standardization would,in view of their number, have only negligibly altered the norms; and second, because certain groups whose inclusion might similarly have been questioned were nevertheless used. (Matarazzo, 1972) p.72

In an effort to explain his views and those of Wechsler, he said:

> Here, again, our view is that in an ideal standardization there ought to be separate norms for each of these categories,to make allowances for their respective influences We refer to the factors of race, social milieu and economic status.

However, Wechsler's associates Weiner and Kaufman (1979) claimed that it was not possible to have separate norms at the time, particularly when "those to whom we might look for the facts are at such odds among themselves as to what the facts are".

Since this insight was offered, Mercer and Lewis (1981) have established norms for the WISC-R for Hispanics and black children in California. Kaufman (1981) argued that since these norms are specific to California, they cannot as yet be used with confidence with other populations. He does say that the relationship between parent occupation and IQ for both blacks and whites leads to the importance of supplementary norms which take socioeconomic background into account.

The issue of race as it affects test scores is subsumed in confounding variables. The controversy inherent in this question cannot be researched without some understanding of the sociopolitical perspective from which this issue arose in the United States.The details necessary to an understanding of this perspective go beyond the scope of this literature review.

## 1.4 The Effect Of Culture on Intelligence Testing

The role of cultural factors on mental test performance of different groups received recognition as early as 1910. Studies of the intellectual development of what were considered as culturally deprived groups, including children raised in gipsy camps, isolated mountain communities and city slums brought attention to the effect of cultural experience upon intelligence. More often, these comparisons resulted in discussion of the inferior performances of such groups rather than questioning the instruments used to measure what these children had learned from their environments.Since tests are inventories of behaviour, how well they evaluate a child's learning potential depends on how well they sample the experiences of the child.

Valentine (1968), a sociologist, defines culture as "the organization of experiences shared by members of community, including their standards for perceiving, predicting, judging, and acting. Through culture people adapt themselves, collectively to environmental and historical circumstances". From Valentine's viewpoint, "It is neither intellectually nor ethnically acceptable to portray another way of life merely in terms of comparison, invidious or otherwise, with one's own cultural standards." An anthropologist, Maretzki (1973), observed that "the way people construct reality often affects performance on psychological constructs". (p.IX).

Andrade (1966, pp. 115-127), also an anthropologist, wrote that

> constructions of reality are influenced by cultural traditions, by physical interaction with the surrounding environment, by statuses and roles held in social institutions, and by the individual's personalty characteristics. (pp. 115-127).

He used a modified version of the Kohs Block Test (from which the WISC-R Block Design subtest derives). Andrade found this test to be culturally inppropriate for use with Hausa children (in Nigeria). They were able to make copies of the actual block formed in a design but had difficulty in interpreting a picture of the block. Andrade's results clearly demonstrated the role of culture in interpreting two-and three-dimensional fields.

Cole (1975) similarly arrived at the conclusion that differences in performance reflect unique cultural circumstances as much as differences in the potentials of individual children. In reference to the conservation principle of Piaget, Cole noted the fact that because an African child does not perform a laboratory task this does not mean that in daily activities the child may not show an understanding of this principle of conservation. (p. 171)

In a decade of research on the topics of cognition and culture, Cole found that the very question he pursued in his research had to change. He became increasingly sceptical of the artificial conditions traditionally involved in measurements used as indicators of innate abilities. Cole (1975, p.174) discussed the way in which ethnographic descriptions of a child's adaptation to his/her own world may give a better estimate of what the child knows or can do.

Reschly (1979) in reviewing WISC and WISC-R records of boys from the San Francisco Unified School District, found that subtest averages differed according to the testees' identity as Chinese, Hispanic, or white. Chinese and Hispanic boys scored significantly lower than white boys on the verbal subtests (Information, Vocabulary, and Comprehension). A significant difference occurred in the performance subtests, with the exception of Coding, in which the Chinese sample scored highest. Using the test protocols of reading-disabled students from the same school district, Mercer (1980) compared WISC and WISC-R subtest results for Anglos, Chinese, and Hispanics. Anglos were found to score significantly higher than Chinese in Vocabulary, Similarities,and Information and higher than Hispanics in Information, Similarities, Vocabulary, and Comprehension

In contrast to this finding that language tests may be more culturally biased than non-language tests, Anastasi (1976) cautioned that the latter may be as culturally loaded as the former. Investigators using a wide variety of cultural groups in many countries have confirmed larger group differences in nonverbal tests than in verbal tests. Sattler (1982, p.32) quotes Jensen as noting that Block Design has little bias, but Sattler himself found the performance scale of the WISC to be as difficult as, or more difficult than, the verbal scale for black children of lower status.

Mercer and Lewis's (1971) data from The System Of Multi--cultural Pluralistic Assessment (SOMPA) sample confirmed Sattler's conclusion by showing that the verbal and performance IQs of black children: 88.7 , 90.1; Hispanic: 87.7 , 97.9; and whites: 102 , 103.8; were significantlly different. (p.129).

Mercer (1981) when looking at learning disabilities and mental retardation placement, demonstrated that more Spanish--speaking children than Anglos were placed in the mental retardation classes. Identification of retardation may have been a consequence of limited English verbal abilities.

These results can be understood in terms of a mental structure climate of "phenomenal absolutes" in which people unreflectively take their own cultural values as objective reality and automatically use them as a context within which to judge the values or skills of others. Judgments of ability based on culturally particular standards can be rationalized, according to Sattler (1982), under these circumstances:

> If the intent is to predict the individual's ability to learn the content of the more general culture, tests designed for the subculture will be less reliable than those which sample from the general culture.

Sattler (1982) justifies this rationalization by stating that "conventional" verbal abilities are needed for the individual to progress in the dominant culture and its educational system.

This argument can be countered with the theory that a child's success in adapting to and learning his/her own culture would seem a wiser test of this child's potential to learn in the dominant culture. Yet WISC-R results may be valuable information for the diagnostician to have in order to assist the child effectively . Specifics from WISC-R sub-tests can be used to plan what a child needs to learn so that a comfortable adjustment can be made to a dominant- culture school.

## 1.5 The Effect of Socioeconomic Level on Intelligence Testing

In a review of the literature on IQ and socioeconomic level, Deutsch (1973, p.33) concluded that "Recent studies have yielded data to contradict the consistent findings of Social Economic Status (SES)--IQ relationships in the direction of higher IQ for subjects of higher SES."

The same writer (1973, p.25) points out that socioeconomic status refers to the broad grouping of people essentially in terms of the amount of income a family receives and the way it is received. Considered under the mode of acquisition is the general esteem in which the particular occupation is held. Related to occupation, and also related independently to socioeconomic status, is the amount of formal education the parents have acquired. Deutsch states that:

> Although numerous other variables have been included in various socioeconomic indexes which have been constructed, economic level and prestige attribution are the major characteristics of the many definitions of social class.(p.39).

A provocative argument on definition of social class is given by Ogha (1978), who holds that where class and caste coexist, as he contends they do in the United States, the controlling factor tends to be the caste system.Each caste, black and white, has its own economic classes, but the two caste systems are not equal in terms of access to education, occupation, income, and other attributes that determine social class membership for the individual.

Eells and others (1951) researched the extent to which Intelligence Quotients, as determined from nine different group IQ tests, were related to socioeconomic status. The pattern of wrong responses on 315 items was compared for students from high and low status homes. For 136 of the items the pattern was closely similar, but for 75 items there were significant differences. It was shown that when an item was expressed in terms of strange, academic, or bookish words, the status differences were much greater than when the items were expressed in simple everyday words.Other factors contributing to status group differences in testing included familiarity with certain objects and the relationship of items to formal school learning. Eells (1951) and his associates stressed the need for careful labelling of what a test measures, and they emphasized the dangers of regarding a test as a measure of general intelligence when it is composed of items which have social status bias. The results of this comprehensive study were possibly compromised because a number of the tests required reading skills.

Problems in reading skills, as noted by Deutsch (1973), have a higher incidence among lower socioeconomic groups than among middle-class groups. Once such a descriptive statement is made and it is verified by further descriptive studies , the behavioural scientist who wants to gain greater insight into the relationship ,or the educator who wants to ameliorate the reading disabilities in lower class groups, must delve more deeply into the variables involved relating specifics to specifics. (Deutsch, 1973)

This connecting of particular reading skills to particular learning tasks requires a high level of sophistication in reading diagnostics and knowledge of class differences in language. The documentation of class differences in reading and other verbal skills has yet to be delineated in any specific way. In using an intelligence test to isolate a child's learning strengths and weaknesses, the diagnostician needs to know the relationship between test content and the child's experience.

In a clearer example of culture specific test items Williams (1972) developed the Black Intelligence Test Of Culture Homogeneity (BITCH). This test was used to confirm that culture--specific items could be constructed so that the average and bright members of a white middle--class group would not succeed as well as blacks in their responses.

Komm (1978), perhaps missing the point of Williams' (1972) instrument, compared the WISC-R and BITCH in order to assess the suitability of the WISC-R for testing the intelligence of black lower--middle--class students who would be more familiar with the street language used in the BITCH. He predicted that middle--class black students would score lower on the BITCH than would the lower--class black students. Komm found that this did not occur; in contrast to what he had predicted the middle--class black students were more successful on both the WISC-R and the BITCH. The mixed results of his study suggest that not just the language of the test, but the testing context also influenced the effects of socioeconomic level. Significant class differences did exist in the subtest results on the WISC-R. Those subtests which were a Standard Deviation higher in the black middle--class results included Information, Similarities, Comprehension, Block Design, and Coding.

The frequencies within which an upper socioeconomic group, range on the WISC-R are seen in a study completed by Kaufman (1981). He reported on the WISC-R scores of a group of upper socioeconomic class children. Psychologists would call these children "gifted children" because they obtained a mean verbal IQ of 127.6 with scale score of 130, qualifying these students for gifted class placements. The authors accounted for the results of their research by explaining that the high verbal scores reflected the children's cultural and home environments.

Vandergriff (1978) indicated that "Highly gifted Anglo state school children score higher verbal and full scale IQs on the WISC-R (1974) than WISC (1958), with the Information, Similarities, Comprehension, and Coding scores particularly higher on the WISC-R". (1974). The results of studies which compare the WISC and WISC-R show that the WISC-R is more difficult, but not, apparently, for the upper socioeconomic class children. This could indicate that whatever class bias exists in the WISC has been increased with its revision in 1974.

The study completed by Tuma, Applebaum, and Bee (1978) called attention to a widening of the gap, for any one class, between scores for the verbal and the Full Scale IQs. These researchers included in their sample, children from a low social class whose Full Scale IQs on the WISC and WISC-R ranged from 67 to 118, and children from an upper social class whose IQs ranged from 108 to 152. Significant differences existed between the two groups in the Verbal, Performance and Full Scale IQs for both the WISC and the WISC-R .

When studying the influence of socioeconomic level on IQ scores, it is important for investigators not to select their subjects from a narrow range, as this will lead to a homogeneity of IQ scores. The homogeneity of IQ scores in the sample is likely to obscure the operation of the socioeconomic variable intended to be studied. For example, a study by Dean (1977) limited the range of IQ in a study of reading achievement, social class, and subtest patterns on the WISC. After excluding children from their study who had IQs above 110, they argued that socioeconomic level had no effect on subtest patterns or IQ scores.

It is not uncommon for a comparative study to take no account of the socioeconomic levels in a sample. This occurred in Dean's study (1979) of 60 Anglo and 60 Mexican American children. There were significant differences in two Verbal subtests, Similarities and Arithmetic, and one Performance subtest-Picture Completion. The Mexican-American children scored lower than the Anglo children .

As shown in the WISC-R literature, children who possessed Spanish surnames and black children from low economic levels were more likely to be placed in special- education classes for the mentally retarded. These findings were reported by Mercer (1981), who questioned the use of the test as a sole determinant of retardation. She decried the fact that 99 percent of the persons she studied, who were selected by the schools as mentally retarded, had been given an IQ test, but only 13 percent had received a medical evaluation. Mercer declared that persons from lower socioeconomic backgrounds were "overlabelled" and persons from higher socioeconomic background were "underlabelled" when compared with their percentages in the general population. The children labelled as mentally retarded from high socioeconomic levels were more likely to have physical disabilities than children thus labelled who were from low socioeconomic levels.

Misinterpretation of tests may occur if consideration is not given to the socioeconomic level of the child with learning needs. An even greater score error may result if the test results are looked on as representing the child's actual cognitive abilities without acknowledging the relationship of testing context, test content, and socioeconomic level of the family.

Evidence exists that the predominant cause of educational underachievement, if viewed historically, is social class injustice resulting from racial and ethnic discrimination. In a scathing critique of the field, Cole (1981) wrote that a concern for the aetiology of learning and disabilities should focus on children rather than environments.

Educators concerned with such aetiology might better turn their attention to such primary social factors as excellent educational practice, family relations, and the child's sense of self, and if problems in these aspects were eliminated, their positive influence would produce a noticeable degree of educational improvement.

## 1.6 The Psychometric Approach To Intelligence.

Testing has played an invaluable role in the forging of a definition of intelligence.In fact, Resnick (1976) considered "the measurement of intelligence for the purposes of prediction to be one of psychology's major success stories". Intelligence testing was first used in the latter part of the nineteenth century. The same social conscience which prompted humane treatment for the mentally retarded also established special institutions for their care. A need arose for some way to differentiate between individuals and to prescribe treatments for them. Although tests had been used since Galton's time, no large scale method of assessment existed until 1905, when Binet and Simon (1916/1973) established their testing procedures for the Minister of Public Education in Paris. Binet chose to construct a large series of short tasks which did not directly depend upon formal education, and from which was extracted a single score representative of a child's intellectual potential (Gould, 1981; Butcher, 1968).

Anastasi (1976) stated that the function of a psychological test is to "measure the differences between individuals "and that it is" essentially an objective and standardized measure of a sample of behaviour." (p.23). Intelligence tests measure samples of behaviour and correlation is a technique for determining "the extent to which two psychological traits or qualities vary together" Butcher (1968) in discussing the features of psychological tests ,considered standardization, reliability and validity as essential elements. 'Standardization' means that procedures used in both scoring and administering psychological tests are uniform (Anastasi, 1976). In establishing norms,the test is administered to a large, representative sample of people and 'norms' are set in terms of the average score and performance.

'Reliability' is equated with the consistency of scores obtained by an individual when tested with equivalent items or when re-tested with the same test after a period of time (Anastasi, 1976). Butcher (1968) identified two types of reliability: 'consistency', which refers to the correlation of the past and present results; and similarity of results over time (p.204). 'Reliability' refers to the precision of the test itself, while 'validity' concerns what the test measures. Anastasi (1976) and Butcher (1968) divided validity into four categories:

1- content validity: how well the test measures or samples the curriculum or area it is to measure;

2- concurrent validity: how well the test diagnoses the existing situation;

3- predictive validity: the capacity of the test for predicting future behaviour;

4- construct validity: how well a test measures the trait it claims to measure.

In every measuring instrument, validity is an important consideration.Eysenck and Kamin (1981) considered internal validity most important in testing and considered it as measuring "a factor objectively, with a degree of error which can itself be measured", and measuring also "to what extent it correlates positively with other tests of the same factor".

Intelligence tests are either individual assessment tools or group tests, and although a proliferation of reliable instruments exists today,this chapter will consider the Wechsler Scales.

The Wechsler Intelligence Scales are based on Wechsler's (1958) definition of intelligence as "the aggregate or global capacity of the individual to act purposefully, to think rationally and to deal effectively with his environment." (p.7). Because intelligence is not a mere amalgam of intellectual abilities for Wechsler, quantification of different aspects of intelligence was the obvious means of measurement. Wechsler believed that an intelligence test was valuable in so far as it reflected intelligent behaviour. The Wechsler Intelligence Scales, like the Stanford--Binet Scales, are wide--ranging in their content.

The first test published by David Wechsler in 1939 was the Wechsler-Bellvue , an intelligence test designed primarily for adults and created in reaction to the lack of suitable tests for that population (Wechsler, 1958). He believed that questions formulated to tap a subject's knowledge of information were a good measure of intelligence. In 1955 the Wechsler Adult Intelligence Scale (WISC) was developed. It contained a Verbal Scale with six subtests: Information, Comprehension, Arithmetic, Similarities, Digit-Span, and Vocabulary; and a Performance Scale with five subtests: Digit Symbol, Picture Completion, Block Design,Picture Arrangement and Object Assembly (Anastasi, 1974). Speed and accuracy were taken into account in test administration and scoring.

The WISC was developed originally as a norm-scaling of the Wechsler Bellvue Scale (Anastasi, 1976; Sattler, 1974). It contained twelve subtests,six of the Performance Scale and the others which formed the Verbal Scale. Sattler (1974) believed that the chief characteristics of the WISC were the following: the IQ tables were based on the administration of 10 subtests and had to be adjusted if less were given. Originally, the Wechsler had included only a deviation IQ, but added a mental age scale in the subsequent editions of the test. A deviation IQ was a standard score obtained by comparing the subject's score with the scores earned by a representative sample of his/her own age . In 1974 a revised edition of the WISC was developed and was then referred to as the WISC-R.

## 1.7 Reliability And Validity

The WISC-R has excellent reliability and validity. (Sattler, 1974). The WISC-R manual (Wechsler, 1974) follows the procedure of the WISC manual (Wechsler, 1949) by reporting split--half reliability coefficients, with the exceptions of the Digit Span and Coding subtests which utilized test--retest procedures. The

Verbal, Performance, and Full Scale IQ have "high reliabilities across the entire age range, the average coefficient being 0.93, 0.90, and 0.94 respectively." (Wechsler, 1974).

The reliability for the subtests ranges from a low--figure score of 0.57 for Mazes at the 16 1/2 age-level to a high figure score of 0.92 for the Vocabulary subtest at the same age level. The highest reliabilities are generally found with the Verbal Scale subtests, except for the Block Design Scale.

The subtests of its predecessor, the WISC, are subject to large error variances and limited subtest specificality, thereby limiting the interpretation of specific subtest functions. For the WISC-R the picture is quite different. (Sattler, 1974). For most WISC-R subtests there is enough reliable specific variance to justify some degree of interpretation of an individual's strengths or weaknesses in the abilities hypothesized for a particular subtest (Kaufman,1975).

Although these findings provide a firmer ground for profile analyses of subtest scores, Hirshoren and Karale (1975) warn against placing too much credence on such analyses. Wechsler (1974) and Sattler (1974) have also cautioned against this practice, due to the reliabilities of the individual subtests. The lower the reliability of a subtest, the less confidence there can be concerning the ability purported to be measured by that particular subtest.

The standard error of measurement in IQ points based on the average of 11 age--groups, are 3.19 for the Full Scale, 3.60 for the Verbal Scale, and 4.66 for the Performance Scale. Therefore, more confidence can be placed in the IQ based on the Full Scale than in the IQ based on either the Verbal or the Performance Scale.

In an item analysis of the WISC-R, evidence was found to indicate that the increase in the number of items ''s helped to increase its internal validity. Analysis of the data regarding the internal consistency of this test reveals that the majority of the items operate as significant discriminates (Vance, Gaynor and Coleman, 1977).

## 1.8 Description

The WISC, originally published in 1949, was revised in 1974, 25 years after publication. The WISC-R has now replaced its predecessor as the major instrument for assessing the intellectual functioning of school-age children (Anderson, 1976).In the preface to the WISC-R manual, Wechsler (1974) states:

> The revised WISC, like the scale it succeeds, has been designed and organized as a test of general intelligence. Its author believes that general intelligence exists; that it is possible to measure it objectively; and that by so doing, one can obtain a meaningful and useful index of a subject's mental capacity.

He also believes that the much--challenged and berated IQ, in spite of its liability to misinterpretation and misuse, is a scientifically sound and useful measure, and for this reason he has retained the IQ construct as an essential aspect of the revised scale (Wechsler, 1974).

The WISC-R presents a number of improvements on the WISC. The principal improvements include lengthening several subtests to enhance reliability; revision of items believed to be out--of--date or culturally biased; changes in administration of Verbal and Performance subtests; inclusion of non--white children in the normative sample; and updating of norms. The inclusion of dark--skinned people was another new feature of the WISC-R. There were major changes in all of the subtests except Digit Span. Wechsler reported no change in the Coding subtest, but a change from achromatic to chromatic (colour) stimuli may have affected performance due to perceptual differences.

The Coding tests consist of two separate sets: Coding A is administered to children under the age of 8 years, and Coding B is administered to children who are 8 years old and over. Research indicated that Coding A (for those under the age of 8) and Coding B (age 8 and over) may not be parallel forms of the same basic construct. Coding A tended to be associated with a perceptual-performance factor, while Coding B was unidentified. The inference was made that Coding A was a non--language task, while Coding B was a symbolic language task.

A total of 73 percent of the WISC items were retained, making the WISC-R a basically similar instrument. (Sattler, 1974). The standardization procedures were excellent (Sattler, 1974) sampling four geographical regions, both sexes, white and non-white populations, urban and rural residents, and the entire range of socioeconomic classes.

The WISC-R was standardized on 2200 white and non-white American children selected to be representative on the basis of the U.S. census of 1970. In the standardization sample there were 11 different age-groups, ranging from 6 1/5 to 16 1/5 years, with 200 subjects in each group. Unlike the WISC, which did not include non--whites in its standardization sample, the WISC-R included 330 non--white subjects (305 blacks and 25 altogether from the American Indian, Oriental, and Mexican-American groups, in approximately the same proportion as that which they represent in the U.S. population.

Oakland (1977) noted that non-white children constituted 15 percent of the sample, appropriately similar to the non--whites in the general population. While such sampling procedures are accurate, the practical effect of including so few minority children is that their performance has little effect upon the norms. The performance of the majority group clearly dominates. More research studies are needed to evaluate the validity of the scale for minority children.

The IQ tables in the WISC-R manual are based on ten of the twelve subtests. The two supplementary scales, Digit Span and Mazes, are excluded. Even when Digit Span and Mazes are administered the manual recommends that they not be included in the tabulation of the IQ. The Twelve WISC-R Subtests are:

## 1- Information

The Information subtest contains thirty questions which sample a broad range of general knowledge. This subtest is started at different points, depending on the child's age. For a sample of children aged 6 to 7 years, begin with Item 1; and for children, aged 8 to 10 years, begin with Item 11. All items are scored 1 or 0 (pass-fail) and the subtest is discontinued after five consecutive failures.

The questions can usually be answered with a simple stated fact; correct answers may be brief; children simply demonstrate whether or not they have these facts at their command; they need not find relationships between facts.

On Wechsler's own admission, his Information test was a task he had found as part of ordinary clinical examinations and also a part of the Army Alpha Test used in the first world war. (Yoakum and Yerkes, 1920). In fact, an information test had been developed by Whipple (1909) and included by him in what was called the "National Intelligence Tests". (Whipple, 1921). An information test had also been included in the series of tests put together by Healy and Fernald (1911), Knox (1914), Terman and Chamberlain (1918), Thorndike (1920), Thurstone (1921, 1931) and Wells and Martin (1923)

## 2- Similarities

The Similarities subtest contains seventeen pairs of words, and the children must explain the similarity within each pair. All the children begin with the first item. The first four items are scored 1 or 0 (pass fail); eight items--5 through--17 are scored 2 , 1, or 0, depending on the conceptual level of the response. The subtest is discontinued after three conceptual failures.

## 3- Arithmetic

The Arithmetic subtest contains eighteen orally presented problems, many of which are similar to ones often faced by children. The children are required to give their answers without using paper and pencil. The child's age determines where the subtest is started. Children aged 6 to 7 years (and older children suspected of mental retardation) are started from Item 1; those from 8 to 10 years are started from Item 5; 11 to 13 years, Item 8; and 14 to 16 years, Item 10. All of the problems are timed, with the first thirteen items having a thirty--second time limit; Items 14 and 15, a forty--five second time limit; and Items 16 to 18, a seventy--five second time limit. All items are scored 1 or 0, with the exception of Items 2 and 3, which can be given an additional 1/2 point. The subtest is discontinued after three consecutive failures.

An Arithmetic subtest was included in the Binet and the Army Alpha Test (Yoakum and Yerkes, 1920) and in the tests constructed by Healy and Fernald (1911), Terman and Chamberlain, (1918), Thorndike (1920), Manson (1925) and the National Intelligence Tests (Whipple, 1921).

**4- Vocabulary**

In the Vocabulary subtest, which consists of thirty-two words arranged in order of increasing difficulty, the child is asked to explain orally the meaning of each word (e.g. "What is a _______?" or "What does _______ mean?"). The subtest is started at different points, depending on the child's age. Children aged 6 to 7 years (and older children suspected of mental retardation) are started from Item 1; 8 to 10 years, from Item 4; 11 to 14 years, Item 6; and 14 to 16 years, from Item 8 . All words are scored 2, 1, or 0. The subtest is discontinued after five consecutive failures.

This type of test was a standard measure of intelligence used by other test developers, such as Binet, Terman, and Chamberlain (1918) and in the National Intelligence Tests (Whipple, 1921).

## 5- Comprehension

The Comprehension subtest consists of seventeen questions which deal with a variety of problem situations, involving subjects such as one's body, interpersonal relations, and social mores. All children begin the subtest with the first item, and all items are scored 2, 1, or 0. The subtest is discontinued after four consecutive failures.

Comprehension questions were also part of the traditional clinical examinations for intelligence and were also used in the Army Alpha Test (Yoakum and Yerkes, 1920). A comprehension subtest was to be found in the Binet (1905) and in the tests of Terman and Chamberlain (1918) and Whipple (1921).

## 6- Digit Span

On the Digit Span, a supplementary subtest, the child listens to a series of digits given orally by the examiner and is then required to repeat the digits. There are two parts to the subtest. The Digits Forward part contains series ranging in length from three to nine digits, while the Digits Backward contains series ranging in length from two to eight digits. There are two series of digits for each sequence length. Digits Forward is administered first, followed by Digits Backward. This subtest is not counted in obtaining the IQ when the five standard Verbal scale subtests are administered. All items are scored 2, 1, or 0. The subtest is discontinued after failure on both trials of any item for both parts of the Forward and Backward of the Digits Span.

The Digit Span subtest was used by Healy and Fernold (1911), Woodworth and Wells (1911), Pintner (1923), the Army Beta Test of Yoakum and Yerkes (1920), and the National Intelligence Tests (Whipple, 1921).

## 7- Picture Completion

The Picture Completion subtest consists of twenty-six drawing of objects from everyday life. The pictures, which are lacking a single important element, are shown one at a time. The child's task is to discover and name (or point to) the essential missing portion of the incompletely drawn picture within the twenty-second time limit.

Healy (1911) had previously developed a picture completion test, and this task was to be found not in the Binet Scale Test but in tests put together by Knox (1914), Pintner (1923), Thorndike (1920), (Yoakum and Yerkes ) the Army Beta Test (1920) and Grace Arthur (Arthur, 1925 1930).

## 8- Picture Arrangement

The Picture Arrangement subtest requires the child to place a series of pictures in a logical sequence. The twelve series (or items) are similar to short comic strips. The examiner places the individual pictures (or cards) in a specified disarranged order, and the child is asked to rearrange the pictures in the "right" order to tell a story that makes sense. One set of cards is presented at a time. The motor action required to solve the problems is simply to change the position of the pictures so that they make a meaningful story.

A Picture Arrangement subtest was originally found in the Army Beta (Yoakum and Yerkes, 1920), and the Pintner (1923).

## 9- Block Design

In the Block Design subtest, the child is shown two-dimensional red and white pictures of abstract designs. The task requires using blocks to assemble a design that is identical to the design on each picture. There are eleven items on the subtest, which is started at different points depending on the child's age.

Healy and Fernald (1911) developed a Block Design test, the use of which was amplified by Yoakum and Yerkes (1920) and which was contained in the scales constructed by Knox (1914), Pintner (1923), and Arthur (1925, 1930).

## 10- Object Assembly

The Object Assembly subtest involves the presentation of four jigsaw problems. The task consists of assembling the pieces correctly to form common objects; a girl (seven pieces), a horse (six pieces), a car (nine pieces), and a face (eight pieces). The items are given one at a time, with the pieces presented in a specified disarranged pattern. There is also one sample item--an apple, which consists of four pieces. All the children receive all the items, beginning with the example and continuing with items 1 through 4.

Healy and Fernold (1911) developed the object assembly test, and such a scale was found also in the scales constructed by Knox (1914), Pintner (1911), Yoakum and Yerkes, the Army Beta Test (1920) and Arthur (1925, 1930).

## 11- Coding

The Coding subtest requires the coding of symbols that are paired with other symbols. The speed and accuracy with which the task is performed are measures of the child's intelligence. This subtest consists of two separate and distinct parts. Coding A is administered to children under the age of 8 years, and Coding B to those who are 8 years of age and over.

## 12- Mazes

On the Mazes subtest children are required to draw a line showing how to find their way out of a series of mazes without becoming blocked (e.g., not going through a line that represents a wall). Each maze is presented separately.

This is a supplementary test that consists of nine test mazes and one sample maze. The sample subtest is not counted in obtaining the IQ when the five standard Performance scale subtests are administered.

Items for the WISC and WISC-R are selected differently from those for the Stanford-Binet. Siegler and Richards (1982) stated that "The primary criterion is the item's correlation with other accepted measures of intelligence." (p. 902). Wechsler (1958) explained his criterion in much the same way . "Selection of the items was then made on the basis of the incidence of successes and failures among the various groups. A question was held to be a 'good' one if it showed increasing frequency of success with higher intellectual level." (p. 66). In defending the choice of subtests, Wechsler continued: "Each test is necessary for the comprehensive measurement of general intelligence." (p.93).

## 1.9 WISC-R and the Short Form

The WISC-R is an expensive device because of the time necessary for its administration, scoring, and interpretation. Because of the breadth of information that the complete WISC-R provides, it is used whenever possible in initial evaluations, but for screening, research, and revaluation purposes, the Full Scale IQ (FS IQ) often provides the most essential piece of information. The FS IQ is then combined with knowledge obtained from other tests, interviews, and observations. In these instances, an abbreviated form that accurately predicts the FS IQ would be more cost-effective and, therefore an invaluable instrument.

With the Wechsler, a variety of abbreviated forms of this test have been offered; Wechsler himself discussed the issue of abbreviated forms (Wechsler, 1944 p. 145). He thought that the Scale could be shortened to five subtests, "provided a judicious choice of subtests is made". Wechsler suggested that since the Verbal IQ correlated so highly with the Full Scale IQ, the result would seem meaningful if an abbreviated form of the test was warranted; the full test could be reduced to the verbal subtests.

Basically, there are two kinds of the WISC-R short form: item reduction, where each subtest is administered, but items on most subtests are limited to every second or third; and subtest reduction, where a specific number of subtests are omitted. Each form will be discussed in turn below, with particular attention to the issues of validity, reliability, and application. Factors to be covered are the following: the correlation between Full Scale and short form IQ, the difference between Full Scale and short form IQ, the percentage of IQ classification changes resulting from the use of the short form, and finally, the applicability of the short form for both normal and retarded children.

## 1.10 Item Reduction Short Form

This method of shortening the entire Wechsler Scale originated with Sataz and Mogel (1962), who initially applied it to the Wechsler Adult Intelligence Scale (WAIS). The idea is to retain the clinical features provided by making use of all the subtests while still reducing the administration time. Such forms typically utilize 45 percent of the total test, and the testing time is cut by a half.

Yudin (1966) and Sataz et al. (1962) adapted the item reduction method to the WISC, and were shortly followed by the Sataz group (Sataz, Vande Reit and Mogels, 1967). Differences between the item reduction short forms and subtest reduction short forms are based on:(1) the subtest shortened, (2) the specific items to be deleted, and (3) the method used to obtain full scores. Yudin's procedure is to utilize only the odd items on five of the subtests, only even items on one subtest, and every third item on three subtests. Digit Span and Coding remain unchanged from the original version.. Other variations include that of Finch, Ollendick, and Ginn (1973), who also used every third item for three subtests, but began with item number 1. Silverstein (1968) used a format differing from Yudin's by using odd items on six rather than five subtests. Also, since the Digit Span subtest was not used in the IQ computation in the standardization of the WISC or WISC-R, Silverstein omitted it, as have other researchers since. Hobb (1980) used a split--half approach, administering alternate items for all subtests but Coding--an approach also used by Hackett (1979).

Goh (1979) reported yet another variation. While conforming to the Yudin method of item reduction, he argued for the selection of even items, or every third item. He also pleaded the case for local norms, to be obtained by basing selection on results for different samples of the population to be studied. He reasoned that a short form devised for a specific population would be more accurate then one derived for a general population, and then applied to a specific population.

The validity of the Yudin--type short form was estimated for normal children by Silverstein (1982), who based his study on data from the WISC-R standardization sample. As noted above, his formula utilized every other item on six of the subtests, every third on three others, and all items on the Coding scale. This formula is comparable with most versions of the item reduction short form. He reported the validity of the short form FS IQ with the long form as .92, with a standard error estimate of 5.8.

This encouragingly high correlation is confirmed in a series of research studies covering the full spectrum of bright to normal to referred borderline subjects. Specifically, Rasbury, Falgut, and Perry (1978), utilizing 70 bright 7 1/2 year-olds, found the validity of the short form FS IQ to be .95, with S.D. 3.7. The mean IQ differences were insignificant (121 versus 125). However, classification changes were not systematically explored.

Goh (1979) also utilized a sample of 142 bright youngsters, using both the Yudin type and his own population--specific short forms. He revealed equivalent validities for these two short forms of 0.95 and 0.94. However, the mean IQ based on the Yudin format dropped significantly (from a Full Scale IQ of 110 to 104), while with Goh's short form the mean IQ was similar to the Full Scale IQ. When changes in

IQ classification were determined, these occurred for 67 percent of the Yudin scores, and only 55 percent of the Goh scores.

For 100 normal children, Dean (1977) reported a (corrected) correlation of 0.92 if the Yudin scoring procedures were used, and 0.91 if Silverstein's versions were employed. He did not report whether there were any mean differences or classification changes. Learning-disabled children (N=76) were utilized by Resnick (1977), who reported a validity coefficient of 0.98. Only 34 percent of the cases changed classification, but the mean IQ was higher in the short form by three points.

Intelligence classification as measured by Full Scale IQ with borderline children(A person with Full Scale IQ of 70-79 is considered borderline) was examined by Finch, Kendall, Spirito, Montgomery and Schwartz (1979), who found a correlation of 0.94 between full scale IQ long form and the Yudin short form (N=100). However, the reliability of the Full Scale IQ dropped significantly from 0.79 to 0.71, with short form IQ and 50 percent of the cases changed intellectual classification. Using a similar procedure, Dirks (1980) tried out three separate item reduction short forms: those of Yudin, Sataz, and Finch. Correlations with the Full Scale IQ long form were identical and high at 0.97, but IQs varied according to the form used. The Full Scale IQ was closely matched to the Sataz and Finch IQs (0.78 and 0.84 respectively), while the Yudin IQ short form was 7 points lower than the Full Scale IQ.

Since each of the above studies compared a form extracted from a standardised administered test, Hackett (1979) explored a separately administered short form with the full test, using a mentally--retarded sample. The IQs were identical for the short form and the Full Scale IQ; the correlation of the short form and the Full Scale IQ was 0.96.

It was Wolfson and Bacheleis (1960) who first developed an item reduction test. This short form was composed of: every third item of the Information and Vocabulary subtests; the odd--numbered items of Comprehension, Arithmetic and

Similarities; and all of the Digits Span subtest. Sataz and Mogel (1962) and Marsh (1977) followed this pattern.

Resnick and Entin (1971) have suggested three criteria that a short form of the WISC should satisfy in order to be considered an appropriate measure of IQ. They are: (a) the correlations between shortened and standard forms should be highly significant; (b) the t-tests that compare means from abbreviated and long forms should be nonsignificant;and (c) the percentage of IQ misclassifications should not be so great that the abbreviated form is ineffective.

An empirical approach was devised by Kennedy and Elder (1982), whose abbreviated form used a stepwise, multiple regression to produce a prediction equation based on five of the ten subtests. The criteria for selecting the terminal step in their regression equation were: (1) a maximum standard error of estimate (SE) of 4.0, chosen because it closely approximates the WISC-R average SE of 3.19; and (2) a minimum correlation coefficient of 0.95, arbitrarily set for a subtest to be acceptable in multiple regression equation. Because they did not report a comparison of actual Verbal predicted FS IQs, there is no way to determine how well their model fits Resnick and Entin's (1971) criteria.

Beck and his associates (Beck, Ray, Seidenberg, Young, and Gamache, 1983) used both Kaufman's (1976) linear equation model and Kennedy and Elder's (1982) regression model on a rural psychiatric sample of 600. Correlations between short and long forms were highly significant in both cases. Comparisons were made between actual and predicted mean IQ scores and all means were within one IQ point of each other.

Wikoff and Parolini (1978) made a comparison of two WISC-R short form models to determine the adequacy of each, and the possible superiority of one. Data from 192 psychiatric and 200 special-education subjects were used to compare Kennedy and Elder's (1982) regression model with Kaufman's (1976) linear equating model. In addition, a regression-derived prediction formula determined from each sample was used to predict FS IQs for the other sample.

Sattler (1982) pointed out that "Several methodological problems exist with short forms, including adequacy of sampling, validity, reliability, equivalence of means and agreement in intellectual classification for conversion of scores into IQs. Even short forms with high validities may misclassifiy individually."

## 1.11 Subtest Reduction Short Forms

This method of abbreviating the WISC and WISC-R by omitting subtests, was initially the prototype of short forms. Many forms proliferated for the WISC, developed more or less by Wechsler (1958). However, soon after the introduction of the WISC-R, a systematic approach to determining the best subtest reduction short forms was presented by Silverstein (1975), using a formula that takes subtest intercorrelations and reliabilities for the 2200 children in the WISC-R standardization sample. The best combinations of two to five subtests (dyads to pentads) as determined by the extent that each correlated with the Full Scale, were determined. Sattler (1980) presented similar findings using a slightly different formula. Goh (1979) considers the two sets of Sattler and Silverstien results as relatively interchangeable. Clearly, the relationship to the Full Scale IQ increases with the number of subtests used. For example, with the best dyad (Vocabulary, Block Design) the correlation rises to 0.93 . The relationship between the size of the correlation and prediction of the Full Scale score is revealed by the standard errors of estimate, which, as expected, decrease with the magnitude of the correlation. As calculated by Silverstein (1982), the standard error of estimate for the single best dyad, triad, tetrad and pentad are 7.1, 6.4, 6.0 and 5.5 IQ points respectively. With errors of estimate of this size, the theoretical percentage of agreement of short-form and Full Scale IQs for subjects (with Wechsler's seven category system) scoring from a superior intelligence to mentally deficient, are 61 percent for the dyad, 64 percent for the triad, 67 percent for the tetrad, and 69 percent for the pentad. Thus, even the best dyads to pentads would misclassifiy 39 to 31 percent of the subjects. However, when compared to the item -reduction short forms, the best tetrads and pentads appear to be as good, with no greater expenditure of time and with the advantage of using a 'standard' rather than an altered form of a subtest.

Only one study, to date, has systematically explored the subtest reduction short forms which Silverstien (1982), generated for a population. Sataz and Mogel (1962), using a sample of 210 children certified as perceptually impaired, determined validity scores for each of Silverstien's best combinations. Coefficients were systematically lower than those from the standardization sample, particularly for dyads, which dropped an average of 0.09. However, pentads were encouragingly close, with less than a 0.01 difference.

Many studies also explored the Silverstien combination using the best 'short-form' Vocabulary and Block Design subtests (reliability of 0.88). For large samples of bright children, Brown and Otts (1983) reported a validity coefficient of 0.92 with Full Scale long form, while Dirks, Wessels, Quarforth, and Quenon (1980) found a correlation of 0.79 with the Full Scale IQ long form. They attributed the weakness of Dirks' combination to the erratic inflating effects on scores. Finch et al. (1973), using the short form subtest Vocabulary and Block Design, reported a correlation of 0.88 with the Full Scale IQ long form of WISC-R, and identical mean IQs. However, 44 percent of the children were misclassified by the short form. (Pearson's correlation for this combination was 0.80).

Dirks (1980) explored one of Silverstien's best quartets (Information, Vocabulary, Picture Arrangement, Block Design) for a borderline population, and found a validity of 0.93, or slightly better than Silverstien's 0.91, while IQs were understimated by two points; 47 percent of the sample were within three points of the Full Scale IQ.

Most other combinations studied were either arbitrary in selection or drawn from early WISC research. Dirks (1980), reported various dyads ( ranging from 0.62 to 0.82), and triads (r ranging from 0.67 to 0.85) which had been used to identify gifted children, with somewhat limited success. On the other hand, Finch et al. (1979), using a behaviour--problem sample, reported coefficients ranging from 0.79 to 0.86 for dyads, with 54 to 57 percent of them misclassified. Finch et al. (1973) found a correlation of 0.80 to 0.96 for the pentad for the WISC (Information, Picture Comprehension, Picture Arrangement, Block Design, and Comprehension). This combination was also checked by Finch et al. (1979), who reported a correlation of 0.96 with the Full Scale IQ long form and misclassification of only 32 percent for the pentad. Combination cross validity by Finch, (1973) gave correlations for various samples ranging from .91 to .96 using Information, Picture Arrangement, Block Design and Comprehension.

In his approach, Kaufman (1976) developed a tetrad which was based on both rational and statistical criteria, using a combination of Arithmetic, Vocabulary, Picture Arrangement and Block Design systematically tapping both Verbal and Performance scales. The (corrected) validity coefficient was 0.91 with a reliability of 0.92, and a misclassification of 28 percent.

Still another issue to be solved is, How valid are the profiles derived from short form? In other words, how well do item reduction short forms assess the subjects themselves?

Dean (1977) stated: "In essence, subtest reduction short forms offer a maximum savings in time according to the number of subtests omitted, with quite adequate validity with the longer versions. Perhaps the best reason for using the subtest rather than item reduction short form is the ease of administration coupled with ability to complete the entire test if results are at all equivocal"

In spite of any warning as to their limitations, including those implicit in this review, short forms will be used by time-pressed examiners. However, knowledgeable selection of an appropriate short form for the task involved can enhance prediction.

# CHAPTER TWO

## Literature Review

## Chapter Two

## Chapter Two

### 2.0 Review of Literature on the WISC and WISC-R

In this chapter work is reviewed which has been carried out in countries other than the UK and USA on the adaptation and standardization of the WISC and WISC-R.

Adaptation is not equivalent to translation. The British or American tests based on their special environments render translation only of such tests as quite inadequate and undesirable. Even ordinary things like dress, domestic articles of everyday use, social customs, accepted social behaviour, ideals, normal school experiences and everyday experiences in general are different in other countries.

Every language has its own nuances and modes of expression which cannot be conveyed through another language by mere translation. Some words of one language may not have exact equivalents in another language; the sentence pattern of a particular language has specific characteristics which which may be peculiar to it.

Moreover, the content of items of intelligence tests is usually based on school experiences and school subjects. The curricula and courses of study in Iraq, differ from other countries and therefore a test based on the experiences of foreign children cannot be adequately translated for Iraqi children.

The difficulty value of a test in one environment is unlikely to be the same in another. Therefore such an adapted test must pass through the whole process of standardization involving much statistical work, including the establishment of norms, reliability and validity . Thus, to adapt a test means re-standardizing it after making suitable changes according to cultural differences.

In India, Bhatt (1972) adapted and standardized the Wechsler Intelligence Scale for Children WISC on 440 children (220 boys and 220 girls) of Ahmed-abad city, aged 5 to 15 and studying in Grade One to Grade Eleven. Forty children in each grade, 20 boys and 20 girls, were equally distributed among the eleven age groups. Each child was tested within one and a half months above or below the age range of the test level. Parental occupations of the subjects were taken into consideration. (The national classification gives ten categories, but in Bhatt's study she reduced it to six). Moreover, no time-limit was imposed; all the testees were allowed to take as much time as they needed to work on any item. The results of the study are shown as follows in Table 2.1 .

Khayyer and Mojdehi (1979) in Iran used 40 delinquent children who had a median age of 13.11 years and education of four years, and 40 normal children who had a median age of 14.01 years and education of four years.

Six subtests, three from the Verbal Scale (Comprehension, Arithmetic, and Digit Span) and three from the Performance Scale (Picture Compilation, Picture Arrangement, and Block Design) of WISC were used in this study. Table 2.2 presents the results.

Table 2.1

Means, S.Ds and Reliabilities of Bhatt's Study Showing Verbal, Performance and Full Scale IQ of WISC.

| Test | Mean | S.D. | Reliability |
|---|---|---|---|
| Verbal | 60.05 | 13.30 | 0.98 |
| Performance | 60.01 | 35 | 0.97 |
| Full Scale | 119.89 | 23.64 | 0.98 |

Table 2.2

WISC IQ Scale Score Comparison of Means and S.Ds of Delinquents and Non--Delinquents of Khayyer's Study.

| WISC Scale* | Delinquents | S.D. | Non-Delinquent | S.D. |
|---|---|---|---|---|
| Full Scale | 73.40 | 9.92 | 71.12 | 9.84 |
| Verbal | 82.80 | 9.95 | 85.72 | 11.55 |
| Performance | 68.02 | 10.01 | 60.27 | 9.87 |

* Scaling procedures different from those of WISC (1974)

As shown in the Table 2.2, no significant differences were observed between the means and Standard Deviations of the IQs of the delinquents and those of the normal children on the Full and Verbal Scales, but the delinquents showed a higher mean Performance IQ than the normal children. The findings in this study show low IQs of subjects, attributed to the low socio-economic and educational status of the lowest stratum of Iranian society.

Malin (1954) adapted and standardized the WISC for use in India, with samples of 50 children (20 boys and 30 girls) for each age--level. Altogether, a total of 600 children was tested. He made some changes, converting raw scores directly into a modified IQ or a test quotient rather than scaling the scores, and he also excluded the Picture Arrangement test and reorganized the Vocabulary test.

Delvaux (1969) used a modified form of the verbal section of the WISC to assess the intelligence of Congolese children aged 7 to 15 years. The items were either new or taken from the Wechsler Scale and were double the number essential for the final scale. The initial sample consisted of 450 subjects aged 9 and 11 years. After item analysis the best items were selected according to difficulty level and discriminative qualities. The final Verbal Scale was applied to a second representative sample of 450 subjects, which contained 50 subjects for each age--level from 7 to 15 years. Reliability Coefficients of the Verbal Scale were equivalent to those of the WISC. The validity of the Verbal Scale was estimated on the basis of the correlation between IQ test scores and school achievement, forming a value of 0.73.

Wilson (1982) in his research established local rural Alaskan native norms for the WISC-R, from a total subject population of 320 students picked randomly. The results tend to show that bicultural bilinguals in rural Alaska seem to perform within what is the normal range for a 'norm' sample on the Performance items, which require only a simple knowledge of the English language.

However, in these students there is a large discrepancy between the scores of the Verbal subtests and those of the Performance subtests, the Verbal scores being significantly lower, as shown in Table 2.3 .

Table 2.3

The Means and S.Ds of IQ of Alaskan Rural Native Students.

| WISC-R Scale | Mean | S.D. |
|---|---|---|
| Verbal IQ | 75.4 | 12.1 |
| Performance IQ | 95.2 | 10.4 |
| Full Scale IQ | 84.1 | 12.3 |

Dastoor and Emovon (1972) used the Block Design of WISC only to investigate the effects of educational status and socioeconomic background on test scores. The subjects for the study were 45 male and female 9 year--old Nigerian children. The children were divided into three groups of 15 children each, the groups categorized as follows:

Group A: no formal education, from poor socio-economic
background;

Group B: four years' formal education, from poor socioeconomic background, attending a free primary state school.

Group C: four years' formal education, from higher socio-economic background, drawn from a fee-paying private school.

In Group C, 9 out of the 15 subjects earned time "credits" and made the correct responses to the Block Design before the limited time expired. Groups A and B rotated each Block a number of times, and still could not get the correct position. Several subjects in Groups A and B were unable to complete the design. The results are shown in Table 2.4 .

Table 2.4

The Means and S.Ds of the Block Design Three Groups on Dastoor's Study.

| Group | Mean | S.D. |
|---|---|---|
| Group A | 6.27 | 4.65 |
| Group B | 7.87 | 5.23 |
| Group C | 24.70 | 8.91 |

McFie (1970) compared the educational and environmental influences on the test scores of 12 subjects. (The sample was too small as a basis for interpretation). A comparison between 13 year-olds in Nigeria and the USA showed that for WISC tests the Nigerian group had a considerably lower score on most subtests, especially the Block Design, and Picture Arrangement. Their speed of construction was slower and less accurately oriented than that of the American subjects. However, the Nigerian subjects achieved the same Mean score as the American children on Memory Design.

Haritos-Fatouros (1963) in Greece, in her Master's thesis set out to establish whether the WISC could be modified to make it suitable for Greek school children. The sample was taken from Salonika city. The criteria for selecting the sample were sex and grade only: two groups of 20 boys and 20 girls for each of the age -groups 9 and 10, (80 children altogether) were tested. Comparing her findings with the results of American investigators, she found that the means IQs of the Americans were regularly above what she obtained. Table 2.5 shows the results.

Table 2.5

Mean IQs and S.Ds for Age-Group 9-10 Obtained in Haritos-Fatouros's Study.

| Age | cases | Verbal | | Performance | | Full scale | |
|---|---|---|---|---|---|---|---|
| | | M. | SD. | M. | SD. | M. | SD |
| 9 | 80 | 95.16 | 14.11 | 87.92 | 13.23 | 90.89 | 13.72 |
| 10 | 80 | 93.90 | 14.10 | 86.03 | 11.77 | 89.34 | 12.44 |

Haritos-Fatouros (1963) concluded that further research was needed and that the use of larger samples was required. Moreover, on sex differences, she stated that her evidence suggested that the test was easier for boys than for girls. Also, she concluded that the modifed Vocabulary subtest was more appropriate for use with Greek children, but there was no reason to change the Performance part of the WISC.

Alexopoulos (1979), in Greece, standardized the WISC-R for school children aged 13 through 15 years. The total number of subjects was 100 (50 boys, 50 girls) for each age-group (13, 14 and 15 years). The reliabilities of the scales varied from 0.89 to 0.93 for the Verbal Scale, to 0.83-0.94 for the Performance Scale, to 0.92-0.96 for the Full Scale. Table 2.6 presents the Means and S.Ds for the Verbal, Performance and Full Scale IQs of the study.

Table 2.6.

Means and S.Ds of Verbal, Performance and Full Scale IQs. by Age for Alexopoulos's Study.

| Age | Verbal | | Performance | | Full Scale | |
|---|---|---|---|---|---|---|
| | M. | SD. | M. | SD. | M. | SD |
| 13 | 49.56 | 11.34 | 49.29 | 10.47 | 98.52 | 19.52 |
| 14 | 49.88 | 11.65 | 50.44 | 10.49 | 100.41 | 19.78 |
| 15 | 50.56 | 11.58 | 50.18 | 10.76 | 100.72 | 19.63 |

The researcher concluded that the standardization of the WISC-R for Greek children aged 13, 14, and 15 was satisfactory because in most respects it was equivalent to its American prototype.

The Wechsler Intelligence Scale for Children (WISC) has been revised and standardized in several Latin -American countries. As early as 1951 a Spanish translation of WISC was used in Puerto Rico. Based on the results of a study by Glasser (1977) using WISC with 128 children, small changes were made in all the Verbal subtests, except Digit Span.

Wandemburg (1966) made use of the Puerto Rican revision of the WISC in her adoption of the Verbal subtests in Ecuador and found that it was necessary to re-order the questions based on their difficulty and to modify some items.

Ahumada (1966) encountered problems attempting to obtain a representative sample of children in Mexico because of the very large numbers of children who were not represented in the school population. Test results as between males and females in the Mexican sample of 444 school children were significantly different, with the males scoring consistently higher. Ahumada concluded that these differences were caused by the different subcultures for men and women existing in Mexico.

In 1970 Ramos Lopez investigated the use of the Puerto Rican version of the WISC Vocabulary subtest in Peru. The sample he used consisted of 301 children between the ages of 6 and 15. Lopez concluded that the Puerto Rican revision of the Vocabulary subtest should be used with caution for Peruvian children, since the differences between the results of the children sampled and the results predicted by the norms of the subtest were statistically significant. A new order of difficulty for the Vocabulary items was found and provisional norms established.

Padilla and Roll (1982) administered the WISC-R to 1100 children between the ages of 6 and 16 from randomly selected schools in the District Federal of Mexico. The mean Full Scale IQ was 87.3, with a Standard Deviation of 13.7 . The average mean for all ages in Verbal IQ was 89.2 and the mean Performance IQ was 88.0. Table 2.7 presents the Full Scale IQ obtained by each age--group in the sample.

In Brazil, Paine and Garcia (1974) adapted the Verbal Scale of WISC by modifying several of the items before using them with a total of 640 children from the ages of 6 through 15 years old. The sample consisted of 325 males and 315 females.

Table 2.8 presents the new mean and Standard Deviation in raw scores for each subtest at each age–level for Paine and Garcia's (1974) study.

Sans (1984) in her research adapted the WISC to Argentinian children with a sample consisting of 100 subjects, belonging to different social and cultural levels, attending the highest grade in primary school, and aged twelve and thirteen years. These subjects were given a translated and adapted version of the WISC.

The difficulty presented by each of the items of the various subtests was analyzed and it was observed that the order of some of them did not answer the increasing difficulty criterion which had been attached to them. In the Information and Vocabulary subtests, for example, some elements turned out to be too easy in the order in which they had been placed. Besides, there were

Table 2.7

The Full Scale IQ Means and S.Ds by Age, of Padilla and Roll's Study.

| Age | Mean | S.D. |
|---|---|---|
| 6 | 92.7 | 12.4 |
| 7 | 91.3 | 15.3 |
| 8 | 87.8 | 17.0 |
| 9 | 87.3 | 13.0 |
| 10 | 84.4 | 13.0 |
| 11 | 86.8 | 13.3 |
| 12 | 85.8 | 14.1 |
| 13 | 89.9 | 12.6 |
| 14 | 88.5 | 13.1 |
| 15 | 83.8 | 09.8 |

some words that were too sensitive in terms of the subjects' cultural differences. In the Comprehension subscale some questions turned out to be too difficult, due to the fact that the required information was less frequent in the children's environment. In the Arithmetic subtest one of the items did not meet the difficulty

Table 2.8

The Subtests' Means and S.Ds by Age, of the Paine and Garcia Study.

| | IN. | | CO. | | AR. | | SI. | | VO. | | D.S. | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Age | M. | SD. | M. | SD. | M. | SD. | M. | SD. | M. | SD. | M. | SD |
| 6 | 6 | 2.27 | 7 | 2.10 | 5 | 1.87 | 5 | 2.40 | 13 | 5.26 | 6 | 2.04 |
| 7 | 7 | 2.22 | 7 | 2.74 | 5 | 1.76 | 6 | 2.33 | 17 | 8.80 | 6 | 2.15 |
| 8 | 9 | 3.07 | 8 | 3.02 | 7 | 1.93 | 6 | 2.53 | 23 | 10.92 | 7 | 1.17 |
| 9 | 10 | 2.85 | 9 | 3.36 | 8 | 1.90 | 7 | 3.02 | 25 | 11.61 | 7 | 2.40 |
| 10 | 10 | 2.62 | 9 | 3.07 | 8 | 2.05 | 7 | 2.62 | 28 | 12.65 | 8 | 2.12 |
| 11 | 12 | 3.47 | 10 | 3.40 | 8 | 2.24 | 8 | 2.86 | 34 | 13.81 | 8 | 2.26 |
| 12 | 12 | 3.31 | 10 | 2.93 | 9 | 2.53 | 9 | 2.71 | 32 | 13.84 | 8 | 1.86 |
| 13 | 13 | 3.80 | 11 | 2.79 | 10 | 2.87 | 9 | 3.81 | 35 | 13.65 | 9 | 2.26 |
| 14 | 14 | 4.40 | 11 | 3.52 | 10 | 2.50 | 9 | 4.12 | 36 | 13.58 | 9 | 2.05 |
| 15 | 16 | 4.91 | 11 | 3.18 | 10 | 1.92 | 10 | 2.35 | 43 | 13.12 | 9 | 1.44 |

IN.= Information CO.=Comprehension AR.=Arithmetic

SI.=Similarities VO.=Vocabulary DS=Digit Span

grading shown, while the last two items offered very little possibility of success for 98 percent of the subjects. In the Performance subtest objection was made to the difficulty grading of the subtests of the Block Design and Object Assembly. Though Sans's research was only descriptive, it showed the need for a better adaptation of this test in order to achieve reliable results with these subjects.

Padilla and Roll (1982) indicated that translations of the Wechsler Scales have been used in Mexico for decades and that recent translations of the WISC-R and other Wechsler tests have been published for use in Mexico and Latin America with the approval of the Psychological Corporation. In order to enable Mexicans to avoid relying on U.S. norms for scoring and interpretation, which until recent times has been the case, Padilla and Roll (1982) administered the Spanish version of the WISC-R to a standardization sample of 1100 children between the ages of 6 and 16 from randomly selected schools in Mexico City. The tested Mexican children obtained a mean Verbal IQ of 89.2 for the whole sample and a mean Full Scale IQ of 87.3 (SD=13.7). The mean Full Scale IQ for the 6 year--olds fell about one half a Standard Deviation below the U.S. mean, and dropped to a level more than a Standard Deviation below the U.S. mean for the older adolescents.

McFie (1970) administered the WISC and a Scheme II Reading Test to 61 West Indian children and 61 white controls referred to the London Child Guidance Clinic. The Full Scale IQ mean score of the West Indian sample was 87 as against a mean value of 98 for the English sample.

Bhatt (1972) calculated the reliability and the Standard Errors of measurement of her Scale by the Split-half technique and found for the ages 7--10 the values shown in Table 2.9 .

Table 2.9

Reliability Coefficients and Standard Errors of IQ Measurement of Bhatt's study.

| Test | Reliability | SEm |
|---|---|---|
| Verbal | 0.90 | 5.08 |
| Performance | 0.91 | 4.66 |
| Full Scale | 0.95 | 3.62 |

## 2.1 WISC / WISC-R and Minority Groups

Within any nation there are many distinguishable subgroups or cultures which differ clearly not only in dress, habits, speech, types of occupation, moral and political attitudes, but also in abilities and social class differences. These groups are widespread in Western societies, and even exist in such differently organised political entities as the USSR and other 'Eastern' communities. Iraq as well has its minorities: people of different race, colour or language, who are incompletely assimilated into the community. Obviously too, there are wide differences within these classes and subgroups. Therefore, whenever the average IQs of different ethnic groups are compared, considerable differences are usually found to exist although there is also much overlap, so that group differences only account for a small proportion of the overall variability among individuals.

The existence of differential performance in Intelligence tests, specifically in WISC-R, by minorities has been well documented (Anastasi, 1976; Kaufman, 1979; and Sattler, 1982). Several researchers have investigated this differential performance of ethnic groups.

In rather straightforward comparisons, Mexican American children have been found to score significantly lower than Anglo-American children on all WISC-R IQ scores (Dean, 1977, 1980; Hays and Smith, 1980). Puerto Ricans were found to score significantly lower than Anglo and black groups on all WISC scores (Marmoral and Brown, 1975). Altus (1953) found a differential depression of Verbal IQs in the WISC performance of Mexican--Americans. Reschly and Jipson (1976) found that Mexican--Americans were over--represented in a retarded classification when a cut--off of 69 was applied to their Verbal or Full Scale IQ, but were more normally represented when this cut--off was applied to their Performance IQs. A disproportionate number of Mexican--Americans were classified as retarded by all IQ measures when a criterion of 75 was applied.

Investigating distinguishing patterns of Mexican--American performance on the WISC-R, Dean (1979) found that Mexican--Americans scored significantly higher on Performance than on Verbal IQs, a finding not replicated for a comparable group of Anglos. Hays and Smith (1980) found that the Verbal IQ was the lowest IQ for three ethnic groups of a low--income referral population: black, Mexican--American and Anglos. However, the Performance--Verbal discrepancy was the greatest for the Mexican--American group, approaching Kaufman's (1979) criterion for significance of 15 points for this group only. He also found a Performance--Verbal discrepancy of over 12 points in Hispanic performance on both the WISC and the WISC-R.

Other studies, although not primarily addressing this issue, have also shown Performance--Verbal discrepancies in their data (Marmoral and Brown, 1975; Reynolds and Gutkin, 1980).

A language handicap related to bilingualism is an obvious hypothesis to explain the Performance-Verbal discrepancy characteristic of the Hispanic and Indian-American children's WISC and WISC-R Performance.

In one of the first reviews of the literature on this topic, Arsenian (1945) noted that the bilingual child's performance on non--verbal measures of intelligence was not lower than the monolingual child's performance, especially if the children were matched on environmental background. However, he noted that bilingual children tended to be penalized in their performance on verbal tasks relative to their monolingual peers. Thus, Arsenian isolated a language handicap as the most influential factor in the lower test performance of bilinguals.

Altus (1953) conducted a study typical of the investigations of this era, on the effects of bilingualism on intelligence performance. The Verbal scale of the WISC was compared for two groups, a monolingual Anglo group (N=52) and a bilingual Mexican-American group (N=67); the groups were matched for age. The 11 year--olds' mean Performance IQ for the bilinguals was below average (average 86.43; bilinguals, 84.01). Bilingualism was loosely defined, assuming that there would be much variation in the degree of bilingualism among the children. Altus himself examined all the children, presumably in English. The results were significant, with the monolinguals attaining a Verbal IQ of an average 17 points higher than that of the bilingual group. In addition, a significant Performance-Verbal discrepancy of 12 points was found for the bilingual group, while for the monolingual group, Verbal and Performance IQs were approximately equal, with the mean Verbal IQ exceeding the mean Performance IQ by two points.

Reschly and Sabers (1979) studied a sample of 555 children aged 6 to 11 (274 males and 281 females) from three ethnic groups: black (N=187), Hispanic (N=184), and white (N=184). They found that the overall mean Verbal–Performance was 10.74 (SD=8.84). The Hispanic subjects had the greatest Verbal-Performance disparity followed by white subjects and then black subjects.

Goody (1981) collected data from three samples of 30 children each (Anglo, Chinese, Hispanic) tested on the WISC from 1966 to 1974, and from similar samples of children tested on the WISC-R tested from 1974 to 1979. All 180 subjects were male disabled readers (aged 7 years 6 months to 12 years 3 months). The monolingual Anglos scored significantly higher on most of the Verbal subtests. The Chinese scored significantly higher on WISC-R Coding than Hispanics. The Chinese also scored significantly higher on most Performance subtests than the Hispanics and Anglos. A majority of the WISC-R profiles demonstrated Performance IQ equalling Verbal IQ (less than 15 points difference) for all subjects. The Chinese and Hispanics showed greater Performance-Verbal IQ differences than the Anglos. The main order of Performance found among the Anglos was Conceptual > Spatial > Sequential, and for Chinese and Hispanics, Spatial > Sequential > Conceptual.

Maltzman (1983) performed analyses on the same 180 male disabled readers studied by Goody (1982). In addition to confirming Goody's findings, she found that several WISC and WISC-R subtests were predictive of the types of Bender errors (Bender Gestalt Test) made by the Hispanic–American sample.

Swerdlike (1976) compared the WISC and WISC-R scores of 164 black, white, and Latino children aged 7 years to 15 years 11 months, who had been referred to school psychologists in the United States due to suspected mental deficiency. All the psychologists administered the WISC and WISC-R to the subjects with a test–retest interval of not less than a week or more than a month. All the children obtained significantly higher subtest Verbal IQ, Performance IQ and Full Scale IQ scores on the WISC than on the WISC-R. The WISC and WISC-R differences increased as the ability of the students decreased. The WISC and WISC-R score differences tended to vary significantly for blacks, whites, and Latinos. There were no significant subtest differences between these ethnic groups. The mean scores for the whole group were: Verbal IQ, 81; Performance IQ, 101; and Full Scale IQ, 91.

In summary, on these tests Hispanic children tended to score less than Anglos, but better than blacks on the WISC-R. Their Verbal IQ was found to be generally lower than their Performance IQ, and lower than that achieved by blacks.

In Britain many immigrant communities have been under- achieving, particularly among minority children in schools and in their later employment. Rutter and Madge (1976), in a review of this literature said that "Children from immigrant families, were more likely to be in lower streams in school and less likely to be selected for academic programmes in secondary school."

Yule et al. (1975) tested a sample of West Indian and indigenous British school children between 10 and 12 years old in London. He found that with the WISC, the West Indian children scored lower than the indigenous children. He found also, that West Indian children born in the U.K. scored higher than those who were born abroad.

Lunemann (1974) questioned the use of IQ for minorities; he found the correlation of IQ with school performance to be significantly lower for minority students. Peel and Lambert (1979) concluded from their research that there was no difference in IQ between English and Spanish administration of the WISC-R for students deemed equally proficient in both languages, on a preselection measure. They also found both English and Spanish Verbal IQs to be one Standard Deviation below Performance IQ and concluded that low SES (socioeconomic status) bilingual children are linguistically deprived.

Sandoval and Mille (1980) and Oakland (1979) found significant differences between Anglo, black, and Mexican–American children in relation to the order of difficulty of WISC-R items. Gutkin and Reynolds (1980) compared 2- and 3-factor analytical solutions to the WISC-R for Anglo, black, and Chicano children and concluded that the three groups data were factorially congruent. Reschly and Reschly (1978) found no differences in WISC-R factor–score prediction of teacher ratings of achievement and attention as among Anglos, blacks, Chicanos, and Papagos.

In a study of WISC-R prediction of the results of the Metropolitan Achievement Test, Reschly and Sabers (1979) found significant differences in slope between Anglo, black, and Chicano samples with under–prediction of achievement in some ranges and over–prediction in other ranges. Papago American Indian achievement was consistently over-predicted. The authors concluded that the WISC-R was a valid and non-biased predictor of achievement of the groups studied (N=910).

## 2.2 WISC-R and the American Indian

Many American Indian children experience education failure; achievement lags are common across many academic subjects and Indian drop-out rates at the high school level tend to exceed 50 percent. (Comptroller General, 1972; Weinberg, 1977). For these reasons, the use of intellectual assessment instruments with Indian children is quite prevalent, for both diagnostic and placement purposes. The Wechsler instruments are often used for assessment of Indian intellectual abilities and the contexts in which they are employed and interpreted seem to differ from those associated with non-Indian groups. For example, McShane (1979) compared records of psychological evaluations of 414 Indian and non–Indian students, concluding that Indian children's reasons for referral were different; that a narrower range of assessment devices were employed and that a restricted number of recommendations were made; that parents were involved less frequently; that their culture was not identified as a relevant variable; and that with them more frequent non--standard administration of Wechsler tests was reported. In a significant proportion of the reports, doubt was expressed concerning the adequacy of the assessment results. Mc Shane recorded that such doubts, concerning the valid and reliable interpretation of WISC-R results produced by Indian children, were common and might relate to a lack of examiner information concerning the Indian child's characteristic Wechsler Performance style.

Despite recent heavy practitioner use of the WISC-R for assessment of intellectual abilities in American Indian children, relatively few studies have been reported that have used this instrument for research purposes with American Indian groups.

Hynd and Garcia (1979) administered the WISC-R to 44 primary grade children for whom Navajo American Indian rather than English was the primary language; all attended a reservation boarding school. The means obtained were: Verbal IQ, 64; Performance IQ, 95; and Full Scale IQ, 77. The authors indicated that the WISC-R Verbal subtest scores tapping receptive and expressive English skills were lowest and that some of the Performance subtests (Picture Completion, Block Design, Object Assembly, Coding) seemed to provide an adequate and generally non-biased estimate of potential. Ertz (1985) administered the WISC-R to 24 subjects (age–and sex–matched pairs) of Indian and non-Indian learning-impaired students with a mean age of 10 1/2. (Coding and Digit Symbol tests were not administered). The Indian children scored lower on Information, Similarities, Vocabulary, and Comprehension, suggesting to the author a deficiency in Verbal skills for Indian learning impaired youngsters.

Results of the large SOMPA (System of Multicultural Pluralistic Assessment by Mercer, 1979) standardization study completed by the Indian Children's Village Program of Indian Health Service Mental Health (1985), show significant inter--tribal differences in Performance on the WISC-R. The study's Verbal, Performance and Full Scale mean IQs are shown in Table 2.10.

In a longitudinal study designed to identify specific intellectual and psycholingistic skills related to the achievement gains of Ojibwa American Indian reservation children, McShane and Plas (1982) found the mean Verbal Scale score for a group of 35 Indian children to be 90.5 (S.D.=14.1), while the mean Performance Scale score was 104.4 (S.D.= 14) and the mean Full Scale score was 96.5 (S.D.= 13.5). These performances resulted in a mean Verbal IQ - Performance IQ difference of 13.6 (S.D.=13.5).

Table 2.10

Means and Standard Deviations for Verbal, Performance and Full Scale IQs for the American Indian Children Studied with SOMPA.

| Number | Verbal | S.D. | Perfor. | S.D. | Full S. | S.D. |
|---|---|---|---|---|---|---|
| SOMPA 236 | 90.4 | 13.0 | 104.7 | 13.2 | 96.7 | 17.3 |
| WISC-R 241 | 76.0 | 13.8 | 98.8 | 12.6 | 95.5 | 12.2 |

Brittain (1976) applied the WISC-R to 177 native Indian children in south-western British Columbia. Significant main results (ANOVA) were demonstrated for the group. A large number of items, especially from Verbal subtests, were found to contribute significantly to the total test variance. Inconsistencies in item placement relative to difficulty level were found for several subtests. Because of the poor item validities obtained, the authors suggested that individual profile analysis as a tool of clinical inference should be avoided. On the other hand, McShane (1980), an area cordinator for special education services within the Bureau of Indian Affairs, reviewed 190 WISC-R and WAIS protocols administered to Indian students by various evaluators, and developed a weighted formula to minimize bias which raised Verbal IQ by 9 points, Performance IQ by 3 points and Full Scale IQ by 7 points.

McShane (1980) reported that random samples of American--Indian children showed a mean Verbal--Performance difference of 10 points. Children from traditional families had a 25–points difference while children from acculturated families displayed no Verbal–Performance difference whatever.

Several authors recommend the use of Performance scores to measure learning potential for normal American Indian children (Hynd and Garcia, 1979; Teeter, Moor, and Peterson, 1982). In a large--sample, WISC-R study of non--handicapped, educationally disadvantaged and learning–disabled Navajo children, Verbal scores constituted an index of proficiency in English-speaking classrooms while Full Scale IQ scores led to bias which might have been due to the pervasion of learning potential (Performance) by acculturation (Verbal) (Teeter, 1982). In addition, these learning-disabled children had lower Performance scores and the WISC-R may be unable to measure accurately learning potential for this population.

Bannatyne (1971, 1974) developed new categories of mean scaled subtest scores as a substitute for the Verbal and Performance subtest clustering, Spatial (Block Design, Picture Completion, Object Assembly), Conceptual (Comprehension, Similarities, Vocabulary) and Sequential (Digit Span, Coding, Arithmetic) categories.

## 2.3 Heredity and Environment

Hereditarians believe that little can be done to modify the intelligence with which a child is born, and that there is little reason to provide stimulation for children who are not innately intelligent. Eysenck (1973) Terman and Oden (1959) and Jensen (1972) have indicated that heredity, not environment, is the crucial factor in the variance in intelligence among individuals. The Louisville twin study by Wilson (1972) and Jensen's new (1972) analysis of the original data from the four largest twin studies by Newman, Freeman, and Burt (1955), indicated that statistically significant differences in intelligence are due mainly to heredity and not the environment.

The environmentalists, in contrast to the hereditarians, assume that intelligence is largely a product of the kind of environment in which one grows up; the more stimulating the environment, the higher the level of intelligence and the more intellectually impoverished and sterile the environment, the lower the intelligence. Environmentalists also believe that the child who enters school with a low IQ can, if stimulated adequately, develop a higher level of intellectual competence. MacNemar (1970) in a study of the effect of the environment upon IQ came to the conclusion that IQ changes can occur as a result of environmental changes.

The heredity or environmental influences which may be responsible for changes in the magnitude of IQ are debatable. In any case, whenever a group of children is tested, there is always some unexplained variation among their IQ scores, no matter what their inheritance or their environmental background.

Jensen (1969) concluded that genetic factors contribute 80 percent of the variance between blacks and whites on standard measures of intelligence, with blacks scoring lower than most whites on IQ tests, as well as lower than Mexican--American and Indian-American. "This results from the presence of some native difference between negroes and whites, as determined by intelligence tests". This attitude accounts for lowered social expectations for certain minority groups which, according to Jensen (1972), can eventually produce even lower mental test scores, in addition to the inferior inherited intelligence.

## 2.4 Classification and Test Bias

Test bias has been a major issue in the classification of minority children, specifically in cases of mildly retarded children and those undergoing special education. It has been variously claimed that most tests: 1) are culturally biased; 2) are designed according to a dominant linguistic style; 3) mirror cognitive styles directly opposed to those found in children from low income background and culturally diverse groups; and 4) base their scores on norms derived from samples not including minority children. Forceful allegations of cultural bias stand out prominently in the test bias controversy (Jensen, 1980). However, the concepts of 'culture bias' and 'culture loading' have sometimes been inappropriately used.

'Culture loading' refers to the degree of culture--specificity of test items and asks about the likelihood of success on the test for children of diverse cultural backgrounds. 'Cultural bias', according to Clarizio (1979) is a question which can be addressed empirically by studying a test's predictive validity. Furthermore, a test may be 'culturally' loaded but not culturally 'biased' (Reynolds, 1982).

The concern seems to be that most tests are culture-bound and that culture--specificity can be viewed on a continuum. Nevertheless, it can be seen in the voluminous literature concerning test bias, that 'culture--bound' and 'culture--specific' have been used synonymously, and for the most part Jensen A. (1980) used the term 'culture-bound fallacy' to refer to determination of bias based on casual inspection or judgment.

'Content--bias' has generated another bias definition which refers to the differences in mean levels of performance on cognitive tasks between white and ethnic groups. Those who recognize this form of bias state that there is no reason to believe that cognitive performance levels should differ across ethnic groups and that a fair test would yield identical distributions and means for all groups (Reschly, 1981). In his discussion of mean differences, Clarizio (1979) reported that research shows some support for the content bias hypothesis. For example, Mexican-American children score lower on the Peabody Picture Vocabulary Test (highly verbal test) than they do on the Raven Progressive Matrices (non-verbal test). However, Clarizio (1979) pointed to other data where evidence of cultural bias between blacks and whites was not demonstrated on the part of the WISC-R, the Stanford-Binet or the

Peabody Picture Vocabulary Test. The objective indices of cultural bias were internal consistency reliability, rank order of item difficulty, relative difficulty of adjacent items and relative frequencies of choice of error distractions.

Empirical analysis has been suggested by others (Cole, 1981; Jensen, 1980) to produce operational definitions of bias. Jensen's statistical statement regarding group differences leaves little room for ambiguity. To the extent that cultural allegations have been advanced with regard to standardized tests, criticisms regarding these tests' linguistic styles have also been raised in the literature (MacMillan and Meyers 1977; Olmedo, 1980; Reynolds, 1982). Assessment in the primary language is a logical, commonsense procedure. However, when language variables stand to influence the test's reliability and validity, simplistic interventions of Performance are not possible. Attempts to de-socialize tests have been a common strategy to reduce bias.

Olmedo (1981) outlined critical issues related to the testing of linguistic minorities. The first point has to do with the apparent assumption that speaking a second language naturally allows for testing in that language. The degree of similarity with respect to lexical, morphological, syntactical and phonological inferences from one language to the other is a variable to be considered. The second point has to do with the standard language forms used in most tests and whether scores reflect true ability or merely the child's linguistic acculturation.

Despite efforts to eliminate cultural and verbal components in evaluation instruments, there is reason to believe that low-income and minority children continue to score lower than white children on these tests (Bailey and Harbin, 1980). It has been suggested that unequal performance patterns are due to differences in cognitive styles (Figueroa, Delgado and Ruiz, 1984). Interest in learning styles has been shown (Ramirez and Cataneda, 1974), and there appears to be an emerging view that minority children have styles of learning and thinking different from children from the mainstream (Torrance, 1982). This realisation has been discussed from the perspective of test bias. According to MacMillan and Meyers (1977), non–verbal tests require covert verbalization. In addition, Figueroa, Delgado and Ruiz (1984) pointed to various empirical studies, in which two ethnic patterns of mental abilities have been documented. If non–verbal tests require perceptual skills having to do with abstracting, then these skills are characteristic of field–independence. If Hispanic children are field–dependent, then the performance of these children will be affected accordingly on these types of tests.

Another way to consider whither a test is biased is to refer to the standardization sample. The use of ethnic minorities within such samples has been rejected by Jensen (1980) and termed the "standardization fallacy". According to Jensen (1980), "renorming" an already established test and providing new norms that are more characteristic of the population of individuals being tested accomplishes little in the way of fairness. "Renorming" merely results in essentially different numerical values being assigned to the mean and Standard Deviation of the group's standardized scores and does not change the relative positions of individuals as between groups. How, then, should minority groups be taken into account in the standardization process? Jensen (1980) proposed that comparable item selection procedures (item intercorrelations and item rank order of difficulty) be performed separately within each subgroup prior to the computation of standardized scores for the entire sample.

By far the most organized search for bias in assessment has been generated from the literature on technical bias or determination of bias from a purely statistical calculation. For clarity's sake, the study of technical bias may be classified as of two types: external construct bias and internal–construct bias. External–construct bias has to do with whether a test predicts an external criterion equally well for all groups; internal–construct bias asks about the internal structure of the test (content and construct validity) and whether it measures the same construct for all groups.

Cleary's (1975) definitions concerning technical approaches for determining validity will serve as a starting point for further discussion. When a test user wishes to determine if a test score reflects test performance in a specific content domain, i.e. through a process of achievement testing, the content validity of a test of this kind is judged according to the extent to which it samples behaviours taught in a curriculum area. Construct validity is studied when the concern is to infer whether a test score represents a construct presumed to be reflected in test performance. Lastly, criterion related validity is simply the extent to which test scores are related to a criterion measure.

Reynolds (1982) cited an approach to identify bias in content validity. It involves an analysis of variance procedure. When "item X group" interactions are found, one may assume that the item is biased. However, Reynolds cited studies where "item X group" interactions accounted for a very small percentage of the variance in performance. Factor analysis has been utilized as a procedure to determine bias in construct validity. It allows for an examination of patterns of inter-relationships of performance within a group.

Clearly, there are many pressing issues in the assessment of minority children. Ethnic disproportion in special--education programmes has been the focus of many investigations. For example, Mercer (1980) found that in 1964, 58 percent of those labelled 'mentally retarded' were from low--status background, compared to 42 percent in 1968. These results come from a study which sought to evaluate the impact of school desegregation on special education classes between 1964 and 1968 in California, U. S. A. But in both 1964 and 1968, blacks and Chicanos continued to be disproportionaly labelled as mentally retarded in comparison to Anglo children. The researchers found further that Anglo children were more likely to be labelled in this way if they were physically disabled.

## 2.5 Cross-Cultural Research Concerning the Adaptation of Intelligence Tests to Other Cultures. .

Test bias has received a great deal of attention from many researchers who assume that background of subjects will play an important role in their performance on tasks. Most tests have been designed for those who speak English as their native language, which makes the tests unfair to minorities and groups from other cultures (Justman, 1967). According to Naglieri (1982), if a child's primary language is not English and he or she is tested for measured intelligence, the confounding variables of language preclude interpretation of performance of such items as a measurement of intelligence.

Cross–cultural research (Trimble, Lonner, and Boucher, 1983) has tried to minimize the parameters along which cultures may vary. One of these parameters is language. If cultural groups to be tested speak different languages, then the development of a test that requires no language on the part of the subjects is necessary. Another aspect of individual assessment that may contribute to the cultural variation is related to test construction. Most non–verbal tests contain information that is valid for certain cultures. Thus, a person raised in a certain culture may lack the background of experience to respond correctly to these items.

Several approaches have been suggested in an effort to resolve the issues raised by cross–cultural testing. Most common among these approaches (Samuda, 1973) has been the translation of existing tests into other languages. There are, for example, Spanish versions of the Stanford-Binet, WISC, and WISC-R,as well as Cattell's Culture–free Test. But according to Mercer (1971), the translation of the content of a test designed for a certain culture into the language of another culture does not eliminate the cultural differences. Another approach (Samuda, 1973) is to develop tests that are related to each major cultural group. With regard to this point, the comparison can be made only among individuals belonging to the same linguistic and cultural background. A limitation of this approach is related to the difficulty in delimiting the boundaries of each cultural group for which a test should be developed.

Mercer (1971, 1976) suggests a new trend in the evaluation of children who do not belong to the dominant culture. This approach is related to what Mercer calls "a pluralistic evaluation" by which each individual is compared with his own sociocultural background. This view includes the standardization of a test to a certain culture so that separate norms can be established.

According to Oakland (1982), a test is judged to be culture--fair if the following conditions are present: if the test is standardized on a representative sample of a country's population; if the mean scores and Standard Deviation for all racial--ethnic and social groups are the same; if the reliability and validity estimates are similar for various subgroups, and if the test minimizes language.

## 2.6 Test Construction

### Adaptation of Tasks to a Different Language

Selection of the appropriate instruments for use in cross--culture research is governed by specific criteria and rules. In addition to the requirement of acceptable levels of validity and reliability (Trimble, Lonner, and Boucher, 1983), the instrument must be translated with minimum loss of meaning. According to Trimble et al. (1983), the establishment of several types of equivalence is essential in the translation of an instrument to another culture. These types of equivalence include the following: first, functional equivalence in which the roles that are related to one culture must be comparable to those in another culture; second, translation equivalence or linguistic equivalence ; third, conceptual equivalence in which the meaning is associated with the terms; fourth, matric equivalence, which means that the same behavioural properties should be measured across cultures.

Sechrest, Fay and Zaidi (1972) have proposed several different aspects of equivalence which must be considered in the adaptation of an instrument to another culture. They point out that vocabulary, idiomatic equivalence and grammatical equivalence are very important aspects, but equivalence in terms of experience and concept is probably the most important. To achieve successful translation from one culture to another, the terms that are utilized should refer to real experiences which are familiar in both cultures. If a counterpart for one item does not exist in another culture, the item should be eliminated. In order to achieve an adequate translation of the item, it is necessary to ascertain out what the term means in one culture and then find an equivalent in the other culture.

With regard to conceptual equivalence, the problem arises when the same concept is interpreted in two different ways in two different cultures. This problem can be solved by finding terms or concepts equivalent to the terms or concepts in the original culture.

It is important to realize that the application of instrumentation to another culture requires some modification of the original instrument. This modification should not alter the characteristic of the original version of the test and should measure the same abilities as the original one (Bhusham,1974).

## 2.7 Item Selection

The major aim of item selection is to help improve the test by revising or discarding ineffective items (Aiken, 1985). Item evaluation tasks involve two forms of item worth: subjective judgment, and statistical judgment. The subjective judgment may be made by test specialists or by subject–matter specialists or by both. To judge an item's worth, they are looking for ambiguities of wording or special clues that lead to the answer and for item characteristics as determined through item difficulty and discrimination indices or item-total correlation.

The measure of item difficulty is very important in evaluating the appropriateness of an item for examinees taking a test. If the item difficulty is close to 0 or 1, then the item is either very easy or very difficult. Therefore, it is important to alter or discard the item because it gives no information about differences among the examinees' trait levels or abilities (Allen and Yen, 1979). Conversely, the closer the difficulty level approaches to .50, the more differentiation the item can make. For maximum differentiation, it would thus seem that one should choose all items at the .50 difficulty level. For most testing purposes, items close to this point are preferable (Anastasi, 1976). According to Allen and Yen (1979), an item difficulty of about .30 to .70 maximizes the information that the test provides about differences among examinees.

# CHAPTER THREE

## Pilot Study

CHAPTER THREE

## Chapter Three

### 3.0 Method and Procedures of Pilot Study.

Wechsler has emphasised the need for applying a number of considerations regarding their suitability when items are selected for inclusion in a test."Satisfying statistical criteria, items should also be considered in the light of actual experience, and must have common sense appeal, that is, must not be tricky or foolish or unfair to the examinee." (Wechsler, 1958).

Twenty-two alterations were made by the author to the Information, Similarities, Arithmetic and Comprehension sub-tests of the British WISC-R. The Vocabulary test was a new one in which the author followed the Wechsler steps when he made this scale the new Vocabulary test as shown in Appendix A . The Performance part of the British WISC-R remained untouched, because no differences were found in the Performance part of the British WISC-R between the U.K. norms and those of the Iraqi sample.

The bilingual judges (three faculty members from the College of Education, English Language Department, University of Baghdad), who had experience in testing and measurement, were involved in the translation process. These judges worked independently from each other and from the author of this study. The researcher explained to the judges (separately) the purpose of the study and asked them to (1) translate the instructions and items from English to Arabic and (2) assess the appropriateness of the items for Iraqi children.

**The Information Test.**

This subtest assesses general knowledge, and as such it contains many items of information which are specific to Western culture and the United States in particular. Therefore many alterations were required, such as changing the wording of some questions, and the content of others.

The Information test given to the present sample contained 30 items. Questions 5, 9, 16, 17, 19, 21, 24, 27 and 28 were altered. A list of old and new questions are given below.Thus, in the Information test 9 items were altered. All the alterations were suggested by the bilingual judges.

**Information**

| Old Questions | New Questions |
|---|---|
| 1-What do you call this? (show thumb.) | The same. |
| 2-How many ears do you have? | The same. |
| 3-How many legs has a dog ? | The same. |
| 4-What must you do to make | The same. |
| the water boiling ? | The same. |
| 5-How many pennies make a pound? | How many Fiels make a Dinar? |
| 6-What do we call a baby cow? | The same. |
| 7-How many days make a week? | The same. |
| 8-Name the month that comes after March. | The same. |
| 9-From what animal do we get bacon? | From what animal do we get wool? |
| 10-How many things make a dozen? | The same. |
| 11-What are the four seasons of the year? | The same. |
| 13-What does the stomach do? | The same. |
| 14-In what direction does the sun set? | The same. |
| 15-Which month has one extra day during leap year ? | The same. |
| 16-Who invented the telegraph ? | Who invented the blob ? |
| 17-From what country did Amerca become independent in 1776 ? | When do we celebrate the anniversary of the Iraqi Revolution ? |
| 18-Why does oil float on water? | The same. |
| 19-Name the two countries that border the U.S. | Name two countries that border Iraq. |
| 20-How many grammes make a kilogram ? | The same. |

| | |
|---|---|
| 21-In what continent is Chile ? | In what continent is Brazil ? |
| 22-What is the main material used to make glass ? | The same. |
| 23-What is the capital of Greece ? | The same. |
| 24-How tall is the average British (Irish) man? | How tall is the average Iraqi man ? |
| 25-What is a barometer ? | The same. |
| 26-What causes iron to rust? | The same. |
| 27-How far is it from London to Edinburgh ? (Republic of Ireland:How far is it from Dublin to Cork ?) | How far is it from Baghdad to Basra ? |
| 28-What are hieroglyphics ? | What are cuneiforms? |
| 29-Who was Charles Darwin ? | The same. |
| 30-What does aluminium come from ? | The same. |

**The Similarities Test**

Only one item was changed in the Similarity test: item 4 had to be changed to be appropriate for Iraqi children; thus, item Number 4 "4- Piano-Guitar" becomes "Lute-Violin".

**The Arithmetic Test**

Many items of the Arithmetic test had to be altered in accordance with the Iraqi monetary system. The questions altered were 7, 8, 10, 11, 13, 14, 15, 16, 17 and 18. Fortunately, the actual content of the Arithmetic Test was not changed as much as the Verbal tests like Information. The old questions and the new are as below.

| Old Questions | New Questions |
|---|---|
| 1- Place in front of the child the card showing 12 trees, and say, Count | The same. |

| | |
|---|---|
| these with your finger. Count them out loud so I can hear you. | |
| 2- With the tree still in front of the child, hand him the blank card and say, "Take this card (point) and cover up all of the trees except four.Leave four trees showing". | The same. |
| 3- With the trees still in front of the child say, "Now cover up all of the trees except nine. Leave nine trees showing". | The same. |
| 4- Remove the blank card. With the trees still in front of the child say, "If we were to add one tree at each end of the line, how many trees would there be altogether"? | The same. |
| 5- "If I cut an apple in half, how many pieces will I have" ? | The same. |
| 6- Barbara had 5 ribbons. She lost 1. How many did she have left ? | The same. |
| 7- John had 4 pennies and his mother gave him 2 more. How many pennies did he have altogether? | Ahmed had 4 Fiels and his mother gave him 2 more. How many Fiels did he have altogether? |
| 8- Jim had 8 marbles and had bought 6 more. How many marbles did he have altogether ? | Moustafa had 8 marbles and had bought 6 more.How many marbles did he have altogether ? |
| 9- A boy had 12 newspapers and sold 5. How many did he have left ? | The same. |
| 10-At 8 pence each,how much will 3 bars of chocolate cost ? | At 8 Fiels each, how much will 3 bars of chocolate cost ? |
| 11-Bill, Dave, and Tom each earned £9 working in a supermarket.How much did they earn altogether ? | Ahmed, Saif, and Moustafa each earned 9 Dinars working in a shop. How much did they earn altogether? |
| 12-A milkman had 25 bottles (cartons)of milk and had | The same. |

sold 14 of them.How many bottles (cartons) did he have left ?

13-A workman earned £36; he was paid £4 an hour. How many hours did he work ?

A workman earned 36 Dinars; he was paid 4 Dinars an hour. How many hours did he work ?

14-If you buy 2 dozen pencils at 45 pence a dozen how much change should you get back from 1 pound ?

If you buy 2 dozen pencile at 45 Fiels a dozen, how much change should you get from 1 Dinar ?

15-Four boys had 72 pence They divided them equally among themselves. How many Pence did each boy receive ?

Four boys had 72 Fiels. They divided them equally among themselves.How many Fiels did each boy receive ?

16-If 3 pieces of bubble gum cost 5 p. what will be the cost of 24 pieces ?

If 3 pieces of bubble gum cost 5 F.What will be the cost of 24 pieces ?

17-Tom bought a second-hand bicycle for 28. He paid 2/3 of what the bicycle cost new. How much did it cost new ?

Adnan bought a second-hand bicycle for 28 Dinars.He paid 2/3 of what the picycle cost new. How much did it cost new ?

18-A jacket that usually sells for 32 pounds was on sale for 1/4 less. When no one bought it, the shop owner reduced the sale price by 1/2. How much did the jacket sell for after the second nd reduction ?

A jacket that usually sells for 32 Dinars was on sale for 1/4 less.When no one bought it, the shop owner reduced the sale price by 1/2 How much did the jacket sell for after the second reduction ?

### The Vocabulary Test

When an intelligence test is translated into another language there is no certainty that the validity of any particular item will remain unaltered. It is a hazardous procedure to rely on the original statistical data as an indication of validity in the new situation.

Rees, (1952) seems to have done this in her adaptation into Welsh of the Watts Vocabulary Test for young children, on the assumption that the measure of a child's breadth of Vocabulary in Welsh would show a positive correlation with the measure of his linguistic background. She translated the items on the Watts Test into Welsh, meeting the difficulties of translation that have been described in her work. Despite having to replace a number of items with new ones she indicated that:

> The final form of the test could not be said to very fundamentally different to any extent from the English test. Had standardization of a Welsh test been possible, many other alterations which would appear desirable could have been made. But it was considered essential that the actual fabric of the test should be changed as little as possible so that the diminishing in its efficiency and validity would be reduced to a minimum.

In the Welsh version of the WISC no assumption is made regarding the validity of the test items in relation to the original scale. The Welsh Scale Vocabulary Test, for instance, is in no sense a translation of the WISC test, as the majority of the words selected have been taken from a dictionary of the Welsh language, whereas the items in the Performance Scale are to a large extent identical with the American WISC Scale.

Wechsler, in his attempt to construct the Vocabulary test for the Wechsler-Bellevue Scale selected 100 words (Wechsler, 1949). These words were given to over 500 children aged 12 through 15. The percentage of children passing each word at each age--level was plotted; words which did not discriminate between age levels were rejected. Approximately 40 words were selected which discriminated according to this criterion.

Haritos-Fatouros, (1963) in her standardization of the American WISC in Greece found that the reliability of the Vocabulary test would drop if she translated the test from English into the Greek language; so 100 words were selected from a well known Greek Dictionary (the Dictionary of N.P. Andiriotis). The random sampling of the words was done as follows: The average number of words per page was found. Then, if, for example, the dictionary contained 800 pages the researcher divided these 800 pages, by 100 (the number of words to be chosen) = 8. She then selected one word for every 8 pages. In other words, she took the words for the Vocabulary test from pages 8, 16, 24, 32, and so on.

These 100 words were included in the Vocabulary test given to the Greek sample she used. She hoped that in this manner a wide enough range of words would be given from which to choose 32, which is the target number of words required in the test to be used with Greek children of the particular age groups chosen by Haritos--Fatouros. It was felt that in this way, a better representation of the words in the dictionary was achieved (Haritos-Fatouros, 1963 pp. 47-48).

In the present study, for the Vocabulary subtest it was recognised from the beginning that the same words merely translated from English into Arabic language could not be used. Moreover, the degree of difficulty of each word would often be entirely different in the two languages and cultures. Having in mind the importance of the Vocabulary test in predicting a person's intelligence, a fact which, among others, Wechsler himself has pointed out, it was decided to use a new Vocabulary test, appropriate for the Iraqi population. This idea was supported by the faculty members of University of Baghdad College of Education because after translating the Vocabulary test they found that some of the words made no sense or were either very difficult or very easy.

The 100 words were selected randomly from a well known school dictionary called <u>Dictionary of the Modern Arabic Language</u> . The random sampling of the words was as follows: Beginning with an odd numbered page, the researcher chose every top word but one (omitting obsolete, technical or esoteric words,) from the left hand corner of every following fifth page, and continued until he reached 100 words. (The three judges from the University of Baghdad approved the omission of the obsolete, technical and esoteric words). Thus, the selection started with page 3, added five pages to make an even number (8) and continued thus with page numbers 13, 18, 23, and so on, until reaching 100 words. These words were subsequently given to 500 children aged between 12 and 15 years. The 500 students were chosen randomly from 20 different secondary schools in suburban and Central Baghdad consisting of boys', girls' and mixed--sex schools. Words which did not discriminate between age levels or reached too difficult a level for groups were discarded, and 40 words were selected at the end. The new Vocabulary test can be seen in Appendix A. Table 3.1 represents the number of students tested at each age--level.

The number of schools shown in this Table and the numbers of the students (boys and girls) were suggested by the Educational and Psychological Research Centre, University of Baghdad (Dr. Al-Ani, N.); the basis for the choice being to follow Wechsler procedures in standardizing his own test.

Table 3.1

Number of Students and Schools Involved in the Vocabulary Test Scheme in Baghdad.

| Age | Boys | Girls | Total | Numbers |
|---|---|---|---|---|
| 12 | 65 | 60 | 125 | 5 |
| 13 | 65 | 60 | 125 | 5 |
| 14 | 65 | 60 | 125 | 5 |
| 15 | 65 | 60 | 125 | 5 |
| Total | 260 | 240 | 500 | 20 |

**The Comprehension Test**

Due to cultural differences, two items in the Comprehension test required immediate alteration. These were question numbers 11, and 12, which may have presented some difficulties to Iraqi students. (There are no factories packing meat in Iraq; direct sale takes place from shops with fresh meat. Also, Iraq has no Charity groups or Charitable Societies, but poor people and beggars).

The old and new questions are listed below.

| Old Questions | New Questions |
|---|---|
| 1- What should you do when you cut your finger ? | The same. |
| 2- What are you supposed to do if you find someone's wallet or purse in a shop ? | The same. |
| 3- What should you do if you see thick smoke coming from the window of your neighbour's house ? | The same. |
| 4- What are some reasons why we need policemen ? | The same. |
| 5- What should you do if you lose a ball that belongs to one of your friends ? | The same. |
| 6- What should you do if a boy (girl) much smaller than yourself starts to fight with you ? | The same. |
| 7- In what ways is a house built of stone better than one built of wood ? | The same. |
| 8- Why is it important for cars to have number plates? | The same. |
| 9- Why are criminals locked up ? | The same. |
| 10-Why do we have to put stamps on letters ? | The same. |
| 11-Why is it important for the government to employ people to inspect the meat in meat-packing factories ? | Why is it important for the government to employ people to inspect the meat in butcher's shop ? |
| 12-Why is it usually better to give money to a well-known charity than to a street beggar ? | Why is it usually better to give money to a well-known poor person than to a street beggar ? |
| 13-Why is it good to hold an ele- | The same. |

| | |
|---|---|
| ction by secret ballot ? | |
| 14-In what ways are paperback books better than hard-cover books ? | The same. |
| 15-Why should a promise be kept ? | The same. |
| 16-Why is cotton often used in making cloth ? | The same. |
| 17-What are the advantages of having Members of Parliament | The same. |

The Performance Scales were left untouched; they were identical with the British WISC-R Scale and no alteration was made to any of the items, because this part of the WISC-R test is culture--free.

### 3.1 Pilot Testing.

During October 1987 a pilot test of the whole instrument took place in a typical secondary school in a suburb of Baghdad. Students of the second and fourth grades of this secondary school were selected as follows.

The entire populations of the 12 year-olds, amounting to a total of 274 (149 boys and 125 girls), and 296 (168 boys and 128 girls) aged 14 were randomised, starting from pupil number 5 in an alphabetical list and taking every seventh pupil i.e. 5, 12, 19 ...etc. for each population separately. If a subject's age was not 12 for the first group we added 1, e.g. 19 + 1 became 20; and if this procedure failed to get a pupil of appropriate age we subtracted 1 from 19, which in this case became 18; thus, the sixteenth pupil was selected. The final random sample of the pilot study was 64 students--32 subjects for each age--group.

Item analyses of the subtests were undertaken, 1) to ensure that the items within each subtest were in the correct order of difficulty so that students would experience greater difficulty as they proceeded through each subtest; 2) to check that the younger age group, as a whole, had greater difficulties with a considerable number of a subtest's items than the older age--group; and 3) to apply British WISC-R norms in scaling Iraqi sample scores to discover if these scores produced 'scaled' mean scores similar to a British sample. If not, new norms would have to be established for the Iraqi population. (See Tables 3.6 to 3.14).

Table 3.2 shows the details of the number of students at each age tested (12 &14) and the percentage of the students tested out of the total number of theStudents of that age within each grade tested.

The secondary school was a mixed one; most of the students' parental occupations were manual workers (skilled and unskilled), small shop owners, and civil servants. The criteria for the selection of the pilot study sample were gender, age and grade.

One and a half hours were spent, on average, with each student in administering the WISC-R's ten subtests. The conditions for administering the test were far from ideal. During the administration of the test many interruptions

**Table 3.2**

Numbers of the Students Aged 12 and 14 Years at the Secondary School used for the Pilot Study.

| School | Students' Sampling Number | | | | | |
|---|---|---|---|---|---|---|
| Age | Boys | Girls | Total | Boys | Girls | % |
| 12 | 149 | 125 | 274 | 16 | 16 | 11.7 |
| 14 | 168 | 128 | 296 | 16 | 16 | 10.5 |
| Tota | 1317 | 253 | 570 | 32 | 32 | 11.2 |

happened because the test was administered in the teachers' staff room. During the intervals between lessons teachers frequently entered the room, so that the student had to go out and return when the bell rang again. This happened with the first four students. Thereafter the administration took place in the secondary school principal's room without any hindrances until all the pilot testing of the sample was completed.

For the pilot test, every item of all ten subtests was administered individually to each child, excluding the Mazes and Digit Span, the two tests which were eliminated from this study. The Mazes and Digit Span are used only as alternatives or as supplementary tests, if time permits. The Mazes subtest, which requires more time, may be substituted for Coding if the examiner so prefers. Any other substitution, including the substitution of Mazes for any other subtest and the substitution of Digit Span for any of the Verbal subtests, should be made only if one of the regular subtests is to be omitted because of special handicaps or accidental disruption of testing procedures. The supplementary tests may always be administered in addition to the regular battery and their inclusion is advised because of the qualitative and diagnostic information provided. In such cases, however, their scores are not used in finding IQs.

As mentioned, the subtests were administered individually to each child. At the same time a bonus system was employed by timing the verbatim responses to each item of the Picture Arrangement, Block Design and Object Assembly tests. Wechsler (1974) stated in his Manual;

> The scoring of Picture Arrangement, Block Design Object Assembly, and Coding A includes points for speed In each case, 1 or more points for quick performance may be added to the child's score but only when he receives a perfect score on the relevant item or test.

It was decided to take two age-groups of mean ages 12 and 14 years old, and to test students of these ages for the Pilot Sample. This method of testing only two age--groups was used in view of the fact that the testing of a smaller number of subjects from within every age group might be inadequate. The results from such a small number of subjects may not be as valid or reliable as that from the chosen larger samples within only two representative age groups (64 students). This procedure followed the pattern of sampling of other researchers in their Pilot Studies (Alexopoulos, 1979 and Haritos-Fatouros, 1963).

The average ages of the two groups of students assessed in the Pilot Sample were 12 years 5 months and 14 years 6 months. The means and Standard Deviations of scaled scores using the British norms for each subtest and for Verbal, Performance and Full Scale IQs were calculated. The results are shown as follows in Table 3.3 .(The scaling procedures will be discussed later, in Chapter Five).

Table 3.3

Means and Standard Deviations of Scaled Scores IQs Using British Norms in the Subtests and the Verbal, Performance and Full Scale IQ Scores for the Pilot Sample of Boys and Girls.

| Test | Gender | Mean | S.D. |
|---|---|---|---|
| Information | Boys | 8.97 | 2.33 |
| | Girls | 8.50 | 2.60 |
| | B + G | 8.74 | 2.47 |
| Similarities | Boys | 9.41 | 3.01 |
| | Girls | 9.78 | 3.11 |
| | B + G | 9.57 | 3.09 |
| Arithmetic | Boys | 8.90 | 2.07 |
| | Girls | 8.18 | 1.85 |
| | B + G | 8.51 | 1.93 |
| Comprehen | Boys | 8.12 | 2.34 |
| sion | Girls | 8.53 | 2.46 |
| | B + G | 8.32 | 2.20 |
| Vocabulary | Boys | 7.03 | 2.09 |
| | Girls | 7.12 | 2.17 |
| | B + G | 7.09 | 2.08 |

(continued Table 3.3 )

| | | | |
|---|---|---|---|
| Picture | Boys | 10.31 | 2.81 |
| Completion | Girls | 10.25 | 2.27 |
| | B + G | 10.29 | 2.58 |
| Picture | Boys | 9.79 | 2.69 |
| Arrangement | Girls | 9.45 | 2.75 |
| | B + G | 9.27 | 2.61 |
| Block | Boys | 9.92 | 2.92 |
| Design | Girls | 9.09 | 2.80 |
| | B + G | 9.46 | 2.84 |
| Object | Boys | 9.73 | 2.59 |
| Assembly | Girls | 8.92 | 2.01 |
| | B + G | 9.27 | 2.09 |
| Coding B | Boys | 7.81 | 2.16 |
| | Girls | 7.45 | 1.90 |
| | B + G | 7.61 | 2.03 |
| Verbal | Boys | 91.95 | 8.83 |
| IQ | Girls | 91.01 | 8.31 |
| | B + G | 91.48 | 8.59 |
| Perform- | Boys | 96.17 | 9.81 |
| ance IQ | Girls | 93.05 | 9.11 |
| | B + G | 94.12 | 9.54 |
| Full Scale | Boys | 93.02 | 9.06 |
| IQ | Girls | 91.14 | 8.75 |
| | B + G | 92.59 | 9.11 |

Table 3.4

Means, Standard Deviations, Gender Differences between Means, Standard Errors of the Differences, and the t-Values for the Raw Scores of Sub-tests with the Pilot Sample (N=64; Boys = 32, Girls = 32).

| Tests | Gender | No. | Mean | Diff.Betw. Means | S.Ds | Standard Err.Diff. | t-value |
|---|---|---|---|---|---|---|---|
| Infor- | Boys | 32 | 18.53 | 0.18 | 3.21 | 1.01 | 0.17 |
| mation | Girls | 32 | 18.35 | | 3.11 | | |
| Simil- | Boys | 32 | 16.86 | -0.15 | 3.59 | 1.20 | -0.13 |
| arity | Girls | 32 | 17.01 | | 2.99 | | |
| Arith- | Boys | 32 | 12.91 | 0.66 | 2.04 | 0.95 | 0.76 |
| metic | Girls | 32 | 12.25 | | 2.83 | | |
| Compre- | Boys | 32 | 20.15 | 2.15 | 3.00 | 0.97 | 1.76 |
| hension | Girls | 32 | 18.00 | | 3.95 | | |
| Vocabu- | Boys | 32 | 34.83 | 1.72 | 4.96 | 1.79 | 0.96 |
| lary | Girls | 32 | 33.11 | | 5.20 | | |
| Picture | Boys | 32 | 21.19 | 2.02 | 3.31 | 1.17 | 1.87 |
| Compl. | Girls | 32 | 19.17 | | 2.89 | | |
| Picture | Boys | 32 | 27.45 | 3.32 | 7.33 | 2.40 | 1.41 |
| Arrange. | Girls | 32 | 24.13 | | 6.04 | | |
| Block | Boys | 32 | 34.91 | 3.81 | 9.18 | 2.00 | 1.21 |
| Design | Girls | 32 | 31.10 | | 8.64 | | |
| Object | Boys | 32 | 23.79 | 1.50 | 4.09 | 2.48 | 0.96 |
| Assem. | Girls | 32 | 22.29 | | 4.75 | | |
| Coding | Boys | 32 | 42.00 | 3.76 | 9.29 | 2.48 | 1.18 |
| B | Girls | 32 | 45.76 | | 3.73 | | |

* Significant at 5% level.

Table 3.6

Percentages of Students Passing the Information Subtest Items at Ages12 and 14 Years in the Pilot Sample.

Pilot Test Item Data - Information Subtest

Item Difficulty.

| Items | 12 Yrs. N=32 | 14 Yrs. N=32 | Total N=64 |
|---|---|---|---|
| 1- | 100 | 100 | 100 |
| 2- | 100 | 100 | 100 |
| 3- | 100 | 100 | 100 |
| 4- | 100 | 100 | 100 |
| 5- | 100 | 100 | 100 |
| 6- | 90 | 96 | 93 |
| 7- | 98 | 100 | 99 |
| 8- | 94 | 98 | 96 |
| 9- | 98 | 100 | 99 |
| 10- | 98 | 99 | 99 |
| 11- | 100 | 100 | 100 |
| 12- | 22 | 32 | 26 |
| 13- | 96 | 98 | 97 |
| 14- | 88 | 91 | 90 |
| 15- | 68 | 74 | 71 |
| 16- | 15 | 25 | 20 |
| 17- | 79 | 90 | 85 |
| 18- | 82 | 90 | 86 |
| 19- | 86 | 92 | 89 |
| 20- | 84 | 93 | 89 |
| 21- | 68 | 80 | 74 |

| | | | |
|---|---|---|---|
| 22- | 65 | 75 | 70 |
| 23- | 21 | 37 | 28 |
| 24- | 44 | 51 | 48 |
| 25- | 42 | 59 | 51 |
| 26- | 47 | 53 | 50 |
| 27- | 28 | 32 | 30 |
| 28- | 30 | 38 | 34 |
| 29- | 8 | 10 | 9 |
| 30- | 1 | 3 | 2 |

Note: The items are numbered as in the British WISC-R in each table, except for the Vocabulary test.

Table 3.5

Kuder-Richardson Reliability Coefficients using Formula 20 for Sub-tests of Pilot Sample.

| Verbal Subtests | | Performance Subtests | |
|---|---|---|---|
| 1- Information | 0.65 | 6- Picture Complete. | 0.67 |
| 2- Similarity | 0.70 | 7- Picture Arrange. | 0.71 |
| 3- Arithmetic | 0.69 | 8- Block Design | 0.68 |
| 4- Comprehension | 0.79 | 9- Object Assembly | 0.76 |
| 5- Vocabulary | 0.96 | 10- Coding B | 0.60 |

Table 3.7

Percentages of Students Passing the Items in the Picture Completion Subtest at Ages 12 and 14 Years in the Pilot Sample.

Pilot Test- Item Data                Picture Completion Subtest

Item Difficulty

| Items | 12 Yrs. N=32 | 14 Yrs. N=32 | Total N=64 |
|---|---|---|---|
| 1- | 100 | 100 | 100 |
| 2- | 100 | 100 | 100 |
| 3- | 99 | 100 | 100 |
| 4- | 100 | 100 | 100 |
| 5- | 94 | 98 | 96 |
| 6- | 96 | 98 | 97 |
| 7- | 98 | 98 | 98 |
| 8- | 86 | 88 | 86 |
| 9- | 84 | 88 | 86 |
| 10- | 82 | 88 | 85 |
| 11- | 88 | 90 | 89 |
| 12- | 83 | 88 | 86 |
| 13- | 75 | 83 | 79 |
| 14- | 81 | 89 | 85 |
| 15- | 76 | 83 | 80 |
| 16- | 85 | 87 | 86 |
| 17- | 85 | 88 | 87 |
| 18- | 84 | 88 | 86 |
| 19- | 70 | 74 | 72 |
| 20- | 66 | 70 | 68 |
| 21- | 41 | 49 | 45 |

| | | | |
|---|---|---|---|
| 22- | 25 | 31 | 28 |
| 23- | 28 | 36 | 32 |
| 24- | 30 | 41 | 36 |
| 25- | 22 | 28 | 25 |
| 26- | 3 | 5 | 4 |

Table 3.9

Percentages of Students Passing the Items of the Picture Arrangement Subtest at Ages 12 and 14 Years in the Pilot Sample.

| Pilot Test | Item Data | Picture Arrangement Subtest | |
|---|---|---|---|
| | Item Difficulty | | |
| Items | 12 Yrs. N=32 | 14 Yrs. N=32 | Total N=64 |
| 1- | 100 | 100 | 100 |
| 2- | 100 | 100 | 100 |
| 3- | 95 | 99 | 98 |
| 4- | 85 | 95 | 90 |
| 5- | 84 | 91 | 88 |
| 6- | 79 | 88 | 84 |
| 7- | 75 | 83 | 79 |
| 8- | 71 | 77 | 74 |
| 9- | 65 | 72 | 69 |
| 10- | 44 | 59 | 52 |
| 11- | 40 | 51 | 46 |
| 12- | 20 | 26 | 23 |

Table 3.8

Percentages of Students Passing Items in the Similarities Subtest at Ages12 and 14 Years in the Pilot Sample.

| Pilot Test | Item Data | | Similarities Subtest |
|---|---|---|---|
| | Item Difficulty | | |
| Items | 12 Yrs. N=32 | 14Yrs. N=32 | Total N=64 |
| 1- | 97 | 98 | 98 |
| 2- | 98 | 100 | 99 |
| 3- | 94 | 98 | 96 |
| 4- | 98 | 100 | 99 |
| 5- | 96 | 99 | 98 |
| 6- | 92 | 96 | 94 |
| 7- | 90 | 96 | 93 |
| 8- | 72 | 80 | 76 |
| 9- | 55 | 65 | 60 |
| 10- | 79 | 86 | 83 |
| 11- | 54 | 65 | 60 |
| 12- | 58 | 63 | 61 |
| 13 | 45 | 59 | 52 |
| 14- | 38 | 42 | 40 |
| 15- | 17 | 38 | 28 |
| 16- | 3 | 7 | 5 |
| 17 | 6 | 12 | 9 |

Table 3.10

Percentages of Students Passing the Items of the Arithmetic Subtest Ages 12 and 14 Years in the Pilot Sample.

| Pilot Test | Item Data | | Arithmetic Subtest |
|---|---|---|---|
| | Item Difficulty | | |
| Items | 12 Yrs. N=32 | 14 Yrs. N=32 | Total N=64. |
| 1- | 100 | 100 | 100 |
| 2- | 100 | 100 | 100 |
| 3- | 100 | 100 | 100 |
| 4- | 100 | 100 | 100 |
| 5- | 100 | 100 | 100 |
| 6- | 100 | 100 | 100 |
| 7- | 100 | 100 | 100 |
| 8- | 100 | 100 | 100 |
| 9- | 98 | 100 | 99 |
| 10- | 94 | 98 | 96 |
| 11- | 86 | 90 | 88 |
| 12- | 87 | 90 | 89 |
| 13- | 53 | 60 | 57 |
| 14- | 50 | 68 | 59 |
| 15- | 34 | 41 | 38 |
| 16- | 10 | 13 | 12 |
| 17- | 3 | 5 | 4 |
| 18- | 2 | 3 | 3 |

Table 3.12

Percentages of Students Passing the Items of the Vocabulary Subtest at Ages 12 and 14 Years in the Pilot Sample.

| Pilot Test | Item Data | | Vocabulary Subtest |
|---|---|---|---|
| | Item Difficulty | | |
| Items | 12 Yrs. N=32 | 14 Yrs. N=32 | Total N=64 |
| 1- | 90 | 94 | 92 |
| 2- | 79 | 84 | 82 |
| 3- | 98 | 100 | 99 |
| 4- | 89 | 94 | 92 |
| 5- | 98 | 100 | 99 |
| 6- | 68 | 73 | 71 |
| 7- | 29 | 32 | 31 |
| 8- | 92 | 98 | 95 |
| 9- | 87 | 91 | 89 |
| 10- | 75 | 79 | 77 |
| 11- | 59 | 68 | 64 |
| 12- | 53 | 55 | 54 |
| 13- | 48 | 58 | 53 |
| 14- | 16 | 19 | 18 |
| 15- | 35 | 45 | 40 |
| 16- | 80 | 84 | 82 |
| 17- | 49 | 71 | 60 |
| 18- | 19 | 28 | 24 |
| 19- | 62 | 78 | 70 |
| 20- | 77 | 78 | 78 |

| | | | |
|---|---|---|---|
| 21- | 8 | 14 | 11 |
| 22- | 65 | 75 | 70 |
| 23- | 23 | 29 | 26 |
| 24- | 35 | 46 | 41 |
| 25- | 33 | 37 | 35 |
| 26- | 25 | 33 | 29 |
| 27- | 13 | 19 | 16 |
| 28- | 26 | 34 | 30 |
| 29- | 92 | 96 | 95 |
| 30- | 40 | 46 | 43 |
| 31- | 25 | 31 | 28 |
| 32- | 15 | 23 | 19 |

Table 3.13

Percentages of Students Passing the Items of Object Assembly Subtest at Ages 12 and 14 Years in the Pilot Sample.

| Pilot Test | Item Data | Object Assembly Subtest | |
|---|---|---|---|
| | Item Difficulty | | |
| Items | 12 Yrs. N=32 | 13 Yrs. N=32 | Total N=64 |
| 1- | 98 | 100 | 99 |
| 2- | 91 | 93 | 92 |
| 3- | 76 | 81 | 79 |
| 4- | 62 | 73 | 68 |

Table 3.14

Percentages of Students Passing the Items of the Comprehension Subtest at Ages 12 and 14 Years in the Pilot Sample.

| Pilot Test | Item Data | | Comprehension Subtest |
|---|---|---|---|
| | Item Difficulty | | |
| Items | 12 Yrs. N=32 | 14 Yrs. N=32 | Total N=64 |
| 1- | 100 | 100 | 100 |
| 2- | 96 | 98 | 97 |
| 3- | 92 | 96 | 94 |
| 4- | 91 | 94 | 93 |
| 5- | 74 | 86 | 80 |
| 6- | 56 | 62 | 59 |
| 7- | 80 | 91 | 86 |
| 8- | 69 | 71 | 70 |
| 9- | 82 | 84 | 83 |
| 10- | 12 | 18 | 15 |
| 11- | 72 | 78 | 75 |
| 12- | 77 | 87 | 82 |
| 13- | 47 | 52 | 50 |
| 14- | 48 | 50 | 49 |
| 15- | 29 | 35 | 32 |
| 16- | 20 | 28 | 24 |
| 17- | 7 | 14 | 11 |

Table 3.11

Percentages of Students Passing the Items of the Block Design Subtest Ages 12 and 14 Years in the Pilot Sample.

Design Subtest Pilot Test Item Data Block

Item Difficulty.

| Items | 12 Yrs. N=32 | 14 Yrs. N=32 | Total N=64 |
|---|---|---|---|
| 1- | 100 | 100 | 100 |
| 2- | 100 | 100 | 100 |
| 3- | 92 | 100 | 96 |
| 4- | 89 | 94 | 92 |
| 5- | 86 | 92 | 89 |
| 6- | 78 | 85 | 82 |
| 7- | 74 | 82 | 78 |
| 8- | 62 | 71 | 67 |
| 9- | 46 | 57 | 52 |
| 10- | 20 | 29 | 25 |
| 11- | 4 | 12 | 8 |

## 3.2 Results of Item Analysis.

The Pilot Sample item data are given in Tables 3.6 to 3.14. The following is a summary of these analyses.

Sub-test (1)–Information.

It can be seen from the results in Table 3.6 that the item difficulty for students of 12 and 14 years of age indicated that the subtest items should be reordered to assure a continuous progression in the direction of increased difficulty. Most items discriminated equally well as with the British WISC-R, but three items, numbers 12, 16 and 23, are more difficult for the Iraqi students. Therefore, they need to be reorganized. The rest of the items' difficulties appear to be in the right order, with the more difficult items in the higher numbers (See Table 3.6). Also, most of the items were more difficult for the younger age-group than for the older group.

Sub-test (2)–Picture Completion.

It seemed from the difficulty value in Table 3.7 that this subtest either possessed very easy items (e.g. Items 1 to 11) or difficult ones (e.g. Items 22 to 26), with most items in the right order. For the items in this subtest the younger age-group found them more difficult than the older group.

Sub-test (3)–Similarities.

The measures of item difficulty for this subtest increased quite smoothly from 98 for Item 1, to 9 for Item 17. Item 10 appeared to be irregular in that its difficulty level of .85 was poorly placed in the order of difficulty of the items, as we can see in Table 3.8 . For all items in this subtest the younger age group found them more difficult. .

Subtest (4)–Picture Arrangement.

This subtest was found to be well-ordered in terms of difficulty, from 1.00 for the first item to 0.23 for the last item. It was felt that this subtest was very suitable to use with Iraqi children. (See Table 3.9). Also, the younger age group had more difficulty with the majority of the test items.

Sub-test (5)–Arithmetic

With this subtest the items appeared to possess the right order of increasing difficulty. This subtest, like the Picture Arrangement, was suitable for use with the present sample (See Table 3.10), as the test's difficulty progressed in the correct order. The younger students found more difficulty with the items.

Subtest (6)–Block Design

As we can see in Table 3.11, this subtest required no alteration in the order of difficulty of its items. The item difficulties increased continuously from Item 1, with a difficulty index of 1.00, to Item 11, with a difficulty index of .08. Also, the younger age–group found most of the items more difficult than did the older group. (See Table 3.11).

Subtest (7)–Vocabulary

As the Vocabulary test was specifically designed for an Iraqi student population, there was no prior information as to the difficulty level of each item; in this case the item-analysis results yielded new information on the ordering of items in terms of their continuing difficulty.

Thus, the results in Table 3.12 contain the items which need to be re-ordered in a similar way to the other subtests.

Sub-test (8)–Object Assembly.

The results shown in Table 3.13 indicated that the items in this subtest showed progressive difficulty in the required sequence. Also, the younger age–group found all the items more difficult than did the older age group.

Sub-test (9)–Comprehension.

Most items appeared to discriminate in the right order, except for Items 6 and 10, where there was evidence of a need to re–order (See Table 3.14).

Since item difficulty measures were completed for each item at the two age levels (12 and 14, years old) for the total sample of 64 subjects, the final item order of every subtest was determined by comparing these two age levels. The results showed the same progression of difficulty with each age group for each subtest.

As can be seen from the Tables (3.6 through 3.14) Subtests (1) -Information, (2) –Picture Completion, and (9)– Comprehension were re--ordered. There were no great differences in the order of items in Subtests (3) -Similarities and (4)- Picture Arrangement from those of the British version. It was decided that these differences were not great enough to warrant a change from the original. The item order of Subtest (5)- Arithmetic and (8)- Object Assembly was also the same as that of the British version.

As can be seen from Table 3.3, the British WISC-R norms may not be appropriate for Iraqi school children, because the means for most of the Verbal and Performance subtests are below 10, which is the mean for the British WISC-R standardization sample. Using the raw scores from the Iraqi pilot sample the equivalent scaled scores based on the British norms (1974) were calculated. These scores produced mean values for most of the subtests below the required mean values of 10 based on the British sample. This meant that British norms could not be directly applied to an Iraqi sample. and that new norms would have to be established for Iraqi children. The Picture Completion subtest however, has equivalent means to the British WISC-R standardization sample. (See Table 3.3).

The Standard Deviations also, of the present Pilot Sample have different values from the corresponding British Amendments of WISC-R. The results in Tables 3.3 and 3.4 show that from statistical analyses of these results it might be dangerous to apply the British WISC-R norms to Iraqi school children, and therefore Iraqi norms must be established by using a large representative sample. (See Table 3.16, which presents the differences between the British WISC-R and the present Pilot Sample in means and Standard Deviations.

Table 3.15

Differences Between the British WISC-R and the Present Pilot Sample in Means and Standard Deviations.

| Test | British WISC-R | | Pilot Sample | |
|---|---|---|---|---|
| | Mean | S.D. | Mean | S.D. |
| Information | 10.10 | 3.10 | 8.74 | 2.47 |
| Similarities | 10.00 | 2.90 | 9.57 | 2.09 |
| Arithmetic | 10.30 | 3.00 | 8.51 | 1.93 |
| Comprehension | 10.00 | 2.90 | 8.32 | 2.26 |
| Vocabulary | 10.20 | 3.20 | -- | -- |
| Picture Comp. | 10.20 | 3.20 | 10.32 | 2.58 |
| Picture Arr. | 10.30 | 3.00 | 9.27 | 2.62 |
| Block Design | 10.30 | 3.00 | 9.46 | 2.95 |
| Object Assem. | 10.20 | 3.40 | 9.42 | 2.08 |
| Coding B | 10.10 | 3.30 | 7.63 | 2.10 |

From the results, which can be seen in Tables 3.3 to 3.16, it was concluded that a re–standardization of the whole test was necessary. This was because most subtest results from the Pilot Sample using the British WISC-R norms gave a mean below 10 and a Standard Deviation below 3. Thus, it was decided to use a large, representative sample, comprising boys and girls, to standardize this test in order to render it more appropriate for use with Iraqi students.

The reliabilities of the Pilot Sample's sub-tests can be seen from Table 3.5, and these values may be considered adequate for the present sample of 64 subjects.

Tables 3.6 to 3.14 present the percentages of children passing each item in the individual subtests, where some differences in the order of the difficulty of items from that of the British study were revealed. This is another justification for the standardization of the WISC-R on a large sample of Iraqi school children.

Some supplementary information is given in Table 3.4, which presents the differences between the boys' and girls' mean scores on each subtest and the significance of these differences (t-tests). From this table we can see that the boys outshone the girls in the Information, Arithmetic, Picture Arrangement, Block Design, and Object Assembly subtest, but these differences are not significant.

The results in Table 3.4, using the raw scores, show that the girls outshone the boys in the Similarities 17. 01: 16.86; but the boys outshone the girls in the all the other subtests. Again, these differences in mean scores are not significant.

# CHAPTER FOUR

## Organisation and Administration of the Main Study

# Chapter Four

## Organisation and Administration of Modified WISC-R and Sample of Main Study.

### 4.0 The Educational System in Iraq

Education in Iraq is mandatory for children, aged six through fifteen years, or for six years of primary and the first three years of secondary school. Since 17 July 1968, education has been free for all individuals from kindergarten to university or other institutions of higher education (Ministry of Education, 1976, p.16).

The Ministry of Education and the Ministry of Higher Education are responsible for education at all levels. In 1979 the government allocated 13.5 percent of the national budget to education.

The present pattern of education in Iraq provides for two years of kindergarten, and six years in primary education in which 2,615,910 boys and girls are segregated in separate schools (except in rural and city schools due to a lack of school buildings). (Schmida and Keenum, 1983). A final examination must be passed at the end of the sixth primary year in order to enter the next level.

Secondary education is also a six--year cycle, divided into three years each of intermediate and preparatory level. At the end of the third intermediate year, students must pass an examination and receive a certificate in order to continue on to the preparatory (or secondary) level. Options at the secondary level include general secondary education in the scientific, literary, or commercial area; preparatory vocational schools (including those for nursing and health); police training; and primary teacher training schools. Many students enter the work force immediately after they complete their initial educational programmes, although some may continue on to higher training or university studies.

Upon completion of the third and final preparatory (secondary) year, students must pass a baccalaureate examination in order to qualify for university admission.

## 4.1 Sample of Schools and Students in Main Study.

The standardization sample was selected by a scheme of random sampling of schools and of children within the schools. It was decided to standardize the modified British WISC-R for the age range 12 - 15, because students of this age range go from primary schools to intermediate schools at age 12 and from intermediate schools to preparatory level at the age of 15 years. According to the 1987 Census, the Iraqi population was 16,478,716. Table 4.1 presents the size of the Iraqi population for the three regions North, Middle and South as well as the number of males and females within the whole population (with the percentage of the national population living within each region).

The population within Iraqi schools in 1988 was 1,058,485 children; this was the total number within all primary, intermediate and preparatory schools (636,400 boys and 422,085 girls). The total number of students in secondary schools was 533,646 consisting of 349,253 boys and 184,393 girls (Ministry of Education, 1988). Table 4.2 shows the numbers of boys and girls, with the percentages of region's population who were students .

Table 4.1

Population, Percentages, (in the Three Geographical Regions), and Male – Female of Iraqi Population.

| Regions | Population | percent | Male | Female |
|---|---|---|---|---|
| North | 4,079,177 | 25.17 | 8,565,273 | 7,913,443 |
| Middle | 8,935,898 | 55.50 | | |
| South | 3,463,641 | 19.43 | | |
| Total | 16,478,716 | 100 | 16,478,716 | |

Table 4.2

Number of Students and the Percentages Studying within the Three Iraqi Geographical Regions (Secondary Schools).

| Regions | Boys | Girls | Total | Percentages |
|---|---|---|---|---|
| North | 74,599 | 30,696 | 105,295 | 19.71 |
| Middle | 213,594 | 123,109 | 336,703 | 63.11 |
| South | 61,060 | 30,588 | 91,648 | 17.18 |
| TOTAL | 349,253 | 184,393 | 533,646 | 100 |

The total number of secondary schools (intermediate and preparatory) was 1,368 consisting of 637 boys' schools 201 girls' schools, and 330 mixed schools. The number of schools used for the present main study was 82, including mixed, boys' and girls' schools. This figure represents a 1:16.7 ratio of the total number of schools. (See Table 4.3). In other words, from within every group of 17 schools in the country we chose a school for our sample. Table 4.3 presents the number of schools in each region, the total number of secondary schools in Iraq, and the sampled numbers of secondary schools in each region.

Table 4.3

Total Numbers of Secondary Schools in Iraq and the Number of Boys' Girls' and Mixed Schools Sampled.

| | Total Schools | | | | Sampled Schools | | | |
|---|---|---|---|---|---|---|---|---|
| Region | Boys | Girls | Mixed | Total | Boys | Girls | Mixed | Total |
| North | 162 | 75 | 100 | 337 | 10 | 6 | 2 | 18 |
| Middle | 373 | 259 | 162 | 794 | 24 | 14 | 8 | 46 |
| South | 102 | 67 | 68 | 237 | 8 | 6 | 4 | 18 |
| TOTAL | 637 | 201 | 330 | 1368 | 42 | 26 | 14 | 82 |

Note: Although there are more mixed schools than girls' schools within the Iraqi system of education, the number of girls within some of these mixed schools is very small (in some cases only 20 to 30 girls). Therefore the chosen school sample for the main study had to contain more girls' schools to get enough girls for testing. As noted earlier, students from age 12 to 15 years study in an intermediate school. Students from age 16 to 18 years study in a preparatory school.

Both the British and the American WISC-R norms presented in Wechsler's manual were derived from groups which were representative of their school populations. These norms were established from the age of 6 years 0 months to 16 years 11 months (Wechsler, 1974). Wechsler indicated, with regard to the American WISC-R that

> although the ideal procedure would have been to use an unrestricted random sampling technique practical considerations made this impossible. Instead, a stratified sampling plan was adopted to ensure that the normative sample would include representative proportions of various classes of children. The stratification along selected variables was arranged in accordance with reports of the 1970 United State census.The variables used were:age, sex, race (White - Nonwhite), geographic region, occupation of head of household, and urban--rural region. (p.17).

An equal number of children, 100 boys and 100 girls, was selected for each age group. The ages ranged from 6 1/2 years to 16 1/2 years; the total sample contained 2200 cases. The United States was divided into the four major geographical areas specified in the Census reports: Northeast, North Central, South, and West. Cases were selected for the normative group in accordance with the proportions of children living in each region. Furthermore, Wechsler stated that like other variables urban-rural residence was not a strictly controlled stratification variable. Communities with 2500 or more inhabitants were defined as urban, and smaller communities as rural. The WISC-R standardization sample was supposed to be a cross-section of the whole population of the United States, and therefore the test was standardized for the whole population of that country. (Wechsler, 1974).

Wechsler's sampling procedures were followed in the present study; figures for the defined population were obtained with the assistance of the Ministry of Education. The data thus obtained gave the total number of children calculated to be at school, together with the number of schools in the three regions. However, due to special features of Iraqi schools, which tend to be homogeneous (e.g. the curriculum, books, schools, and teachers) the criteria for drawing the sample were able to be more limited for the present study in comparison to the American and British ones.

The criteria for the selection of the present sample (main study) were age, gender, socioeconomic status, geographical region and urban-rural variables. As with all such large--scale research it was important to decide the practical limitations that must be imposed on the efforts that were required.

### 4.2 1- Socioeconomic Status.

Socioeconomic status was based on the occupation of the head of a household. The following six occupational groups taken from the 1987 Census categories in Iraq, were selected for the basis of stratification within the sample:

1- Unskilled Workers, factory workers, farmers.

2- Skilled Workers, truck drivers, electrician workers.

3- Private employees, shop--owners, car--owners, house--owners.

4- Lease-holders, high officials.

5- Civil servants, self--employed people doctors, surgeons, lawyers.

6- Higher grade civil servants, merchants, industrialists, general managers.

Note: The information of socioeconomic status groups in this list for the occupations of household are from the 1977 Iraqi Census. In November 1987 (after the author finished the Pilot test) a new occupational division was published for the 1987 Census. The author used this one for the main study.

Students were assigned within the standardization sample in accordance with the incidence of these six occupational categories as specified for the Iraqi population. Each student was asked to try to determine his father's occupation or his last known occupation if his father was in the army. Most men in Iraq were serving in the army during the war between Iraq and Iran. If an attempt to determine the father's occupational group was unsuccessful, (dead, missing or in captivity) the mother's occupation or the father's last job was used.

### 4.3 2- Urban-Rural Status

Each child in the standardization sample was classified for urban--rural status on the basis of the location of his home. Classification by location of school was often misleading, since many children, particularly secondary school pupils, travelled daily from rural areas to schools in urban areas. Information about the Urban--Rural status of students was obtained from the Ministry of Education. Such information allowed certain areas to be classified as urban and other areas as rural.

The delimitation of urban and rural areas was carried out by objective estimations from the Census report (1987). Communities with 2500 or more inhabitants were defined as 'urban' and smaller communities as 'rural'. 'urban' areas thus included all centres of dense population and heavy industry, and 'rural' areas on the other hand, contained thinly populated areas (under 2500) which included agricultural villages and small towns.

### 4.4 3- Gender

It was decided that stratified random sampling of schools should result in a fairly equal representation of both genders, taking into account that some schools are mixed and the rest are single--sex. Although the number of boys in the secondary schools are greater than the number of girls, it was decided to take an equal numbers of boys and girls, as Wechsler (1974) did in his standardization procedures.

### 4.5 4- Geographical Region.

After careful consideration it was decided to divide Iraq into three regions, (according to the government Census 1987) North, Middle and South. Subjects were randomly selected for the normative group in accordance with the proportion of children living in each region. The correct representation was maintained within each geographical region, following Wechsler's (1974) procedure when he divided the United States into four geographical regions and selected at random a percentage of schools from within each region.

It must be pointed out that it is not difficult to decide on the best sample design when one has an adequate knowledge of the availability of information.

## 4.6 Sampling

**Emphasising Precision.**

This involves using those stratifying variables that are most appropriate, namely, gender, age, socioeconomic status, geographical region and urban--rural residence, then sampling at random within the resulting lists.

Another practical question that needed answering was, How many schools do we need to visit? The smaller the number of schools we visited the less easy it would probably be to stratify fully. Again, a decision like this depends partly on a knowledge of variables determining stratification. Because of the homogeneity of the Iraqi school system, which has a national curriculum followed by all establishments, we considered that 82 schools were adequate for the present study (a proportion of 1 out of every 17 secondary schools in the country).

In considering the above mentioned factors it was decided to select the main study sample using the following procedures.

1- The whole country was divided into three geographical regions according to the 1987 Census report (Ministry of Planning, 1987). The majority of the Iraqi population are Arabs, who occupy the south and middle region and part of the northern region. In the far north east of the country (in Kurdistan) is where the Kurds live.

The pattern of life within the Kurdish society has slowly but steadily changed from one of nomadism to one of settled communities. The kurds now pursue a basically agricultural way of life supplemented by the raising of livestock, a limited amount of trade, and the establishment of a small number of local handicrafts and industries. Like the rest of the Middle East, Kurdistan is also divided into rural and urban communities where the people live in villages, towns and urban centres. A Kurd may be best characterized as a mountaineer well adapted to the terrain and climate of his homeland. Kurdish women are normally considered more free than those among the neighbouring peoples.

On March 11th 1974 the Iraqi government decided to established more autonomy for the Kurdistan region. Article number 33 of the Kurdish Autonomy Act 1974 stated that:

(a) The Kurdistan region shall enjoy autonomy.

(b) The records of the 1957 Census shall be deemed as the basis for defining the national nature of the absolute majority of the population.

Article (2) of the 1974 declaration stated that:

(a) The Kurdish language shall be the official language in the region.

(b) Kurdish shall be the language of education for Kurds in the region.

In selecting the sample it was decided not to assess Kurdish students because they have a different language and independent schools. Appendix (B) shows a map of Iraq and the three geographical regions, as well as the autonomous region of Kurdistan.

2- The selected secondary schools were listed and numbered in each geographical region with separate lists for urban and rural schools (See Appendix E).. By using random numbers, 82 of these secondary schools were selected from all over the country roughly in proportion to the number of schools within each region. Thus, a ratio of 1:17 of all Iraqi secondary schools was sampled. Table 4.3 presents the number of secondary schools in each region and the number of sampling schools for the main study. Appendix F shows that letter of approval letter authorizing the study was obtained from the Ministry of Education in Iraq.

3- The number of students sampled for the main study was calculated according to the percentage of students in each region out of the total number of students in the secondary schools for the year 1987. (See Table 4.2). Table 4.4 shows the numbers in the sample taken randomly from each region, boys and girls, with the percentage of the sample for each of the three regions.

4- Secondary schools in Iraq are of two kinds, large (urban) and small (rural) schools. Large schools contain from 12 to 36 classes, each class having between 30 to 40 students. Small secondary schools contain 6 to 12 classes, each class having 25 to 35 students. Around 14 students from each large secondary school or 8 students from each small secondary school were selected randomly, according to the number of students we had to assess from each region and depending on whether the schools were for males only, females only or mixed, to adjust the number of boys or girls. Table 4.5 shows the number of large and small secondary schools chosen randomly from the three geographical regions and the number of the sampling schools in each region of the country, Appendix E presents a list of all large (urban) and small (rural) schools in the three geographical regions.

Table 4.4

Number in Student Sample Tested on WISC-R in the Three Geographical Regions and the Percentages from Each Region.

| Region | Boys | Girls | Total | Percentage |
|---|---|---|---|---|
| North | 70 | 70 | 140 | 17.5 |
| Middle | 250 | 250 | 500 | 62.5 |
| South | 80 | 80 | 160 | 20.00 |
| TOTAL | 400 | 400 | 800 | 100.00 |

5- Students were selected at random from each school. In Iraq, secondary schools normally have a number of different classes in each whole year--group. In order, therefore, to take a random sample of all the students in a whole year group the following procedure was adopted: The total number students in all the classes taken together was divided by the number of students required and rounded off to the nearest whole number. This integral quotient was then used in combination with the alphabetical register of the whole year--group in order to select specific students by the following means: The third student on the alphabetical list was taken and the integral quotient above was then added to the number 3, generating a new number. This student was also taken numerically from the alphabetical list. Further addition of the same number and selection of the appropriate subject from the alphabetical list was performed until sufficient subjects had been selected. For example, if the total number in a whole year group at a large secondary school was 140 students and we wished to select 14 students we divided the total number of students by 14, which is 10, we added 3 ( Most studies added 3 or 7, see Alexopoulos 1979; Al-Aubaidy, 1987) to the 10, generating a new number 13. Subject number 13 in the alphabetical register of the whole year group was taken from the list; 10 was continually added until the whole age group was sampled. It was necessary to modify this procedure in some cases due to the fact that up to 15 percent of the children fell outside the required age--range, being either one or two years above or below the age of the majority of the whole year group. In such cases–where the above selection procedure produced a subject outside the required age–group–the following procedure was adopted to overcome the difficulty. One was added to the number generated by the original procedure and this subject was selected instead of the original one. If this subject also fell outside the scope of the sample, subtraction took place instead from the number originally generated and this subject was taken. Thus, taking the previous example, this procedure would have been followed if subject 13 had been outside the required age group.

Table 4.5

Number of Large (urban) and Small (rural) Secondary Schools Chosen Randomly from Each Region for the Main Study.

| | Urban | Rural | |
|---|---|---|---|
| Region | Large School | Small School | Total |
| North | 8 | 10 | 18 |
| Middle | 27 | 19 | 46 |
| South | 8 | 10 | 18 |
| TOTAL | 43 | 39 | 82 |

Table 4.6

The Numbers of Students (Boys and Girls) Selected Randomly from Secondary Schools in Urban and Rural Areas within Each Geographical Region.

| Region | Boys | Girls | Boys | Girls | Total |
|---|---|---|---|---|---|
| | Large Schools | | Small Schools | | |
| North | 52 | 52 | 18 | 18 | 140 |
| Middle | 186 | 186 | 64 | 64 | 500 |
| South | 56 | 56 | 24 | 24 | 180 |
| Total | 294 | 294 | 106 | 106 | 800 |

The figures in Table 4.6 were based on the fact that about 27 percent of the student population are rural residents and about 73 percent from an urban residence population.

6- For each age--group, 200 subjects comprising 100 boys and 100 girls were selected. It was decided that 200 subjects for each age group were an adequate number for the revision and standardization of the British WISC-R.(See Table 4.4).

## 4.7 Scoring And Administration

Some previous experience in administering and scoring the WISC-R is necessary to ensure that procedures are carried out effectively and reliably. The experimenter gained this experience at the State University of Utah through administering the WISC-R to primary-school students as a study for his Master's degree. As psychological testing has become more common, and administrative procedures more standardized, errors attributable to test administrative conditions have probably decreased. Some misunderstanding of testing conditions, mismarking, mistakes in timing and other distractions however, do occur. One of the most important procedures the researcher undertook before starting the test was to prepare well in advance so that no emergencies affected the testing process.

The tester took care to memorize the exact British WISC-R verbal instructions so that previous familiarity with the test would help in avoiding misreading, and hesitation and permit a more natural, informal manner. Also, the researcher placed the WISC-R testing materials on the testing table within easy reach, so as not to distract the child. The testing room was selected to be free from noise and distraction and to provide adequate lighting, ventilation and seating facilities.

During the testing sessions the room was locked and a sign posited on the door with an assistant outside it to prevent the entrance of any unauthorised person. Approximately 60 to 75 minutes was spent administering the British WISC-R battery of 10 tests. The entire battery was given in a single sitting for all the children involved in this study.

A friendly relationship was established with the children taking the test by putting him at ease, engaging the child in some informal conversation, keeping him interested in the tasks at hand and encouraging the child to do his best. The test was not given to the child until he seemed relaxed enough. The children were encouraged throughout the test, and the tester was sensitive to the specific needs of the children.

The scoring, as far as possible, was uniform because all the testing was carried out by the author; so were the procedures, for all subjects, in administering the test. The importance of this is emphasised by Anastasi (1976) who considered it as one of the main conditions for the standardization of any test.(p. 25).

The test was introduced individually to each child in the following manner: " This is an intelligence test for children aged 6-16 years which was made for Americans, and Western culture I would like to adapt it to the Iraqi culture for the age--range 12-15 years only. I have chosen certain schools and certain students and you have been chosen by chance from among a group of friends to help in this adaptation. It will be used for Iraqi children of the same age as you. Try to answer each question as well and as fast as you can."

The following directions for starting each test have been suggested by Wechsler (1974):

Information Ages 11-13 years: start from Item 7; ages 14-16 years: start from Item 11. If a child aged 8-16 years obtains perfect scores on the first two items he is administered, the child obtains full credit for all earlier items. If he does not earn perfect scores on his first two items, earlier items are administered in the reverse sequence until he obtains two consecutive perfect scores.

Discontinue: after 5 consecutive failures.

Picture Completion Ages 8-16 years: start from Item 5. If a child aged 8-16 years obtains perfect scores on Items 5 and 6, he is given full credit for Items 1-4. If he fails either Item 5 or 6 we return to Item 1 and administer Items 1-4 before proceeding.

Discontinue: after 4 consecutive failures.

Similarities Begin with Item 1 for all children. Items 1-4 are scored 1 or 0, while the remaining items in the test are scored 2, 1, or 0, after counsulting the general scoring principles provided in the WISC-R Manual. Thus, 2 points are given for any general classification which is pertinent for both members of a pair (e.g. "An apple and a banana are fruits"), 1 point for any specific properties which are common to both and constitute similarity: "An apple and a banana are foods".

Discontinue: after 3 consecutive failures.

Picture Arrangement Ages 8-16 years: beginning with Item 3. If a child aged 8-16 years passes Item 3, full credit is given for Items 1 and 2. If he fails Item 3, then go back to Item 1.

Discontinue: after 3 consecutive failures. ( An item is considered failed only if both trials are failed); e.g. the initial administer of Item 3 as its subsequent administration after Items 1 and 2.

Scoring: Items 1-4: 2 points for passing on the first trial; 1 point on the second trial. There are time limits for items on this test and additional bonus points for earlier completion (See Table 4.7).

Table 4.7

Time Limit for Trials and Bonus Points in the Picture Arrangement Test.

| Items | Time Limit | Bonus Point for Each Time Limit 5 | 4 | 3 |
|---|---|---|---|---|
| 1, 2, 3, 4. | 45" | | | |
| 5, 6, 7, 8. | 45" | 1-10" | 11-15" | 16-45" |
| 9. | 60" | 1-10" | 11-20" | 21-60" |
| 10, 11, 12. | 60" | 1-15" | 16-25" | 26-60" |

Arithmetic Ages 11-13 years: begin with Item 8; ages 14-16 years: Item 10. If a child does not earn perfect scores on his first two items, the earlier items are administered in the reverse sequence until he obtains two consecutive perfect scores (not counting the starting point item). Items 1-4 are administered in the order 1, 2, 3, 4. Thus, if a child of 8 years or above does not answer both Items 5 and 6 correctly, we return to Item 1 and administer Items 1-4 before proceeding.

Discontinue: after 3 consecutive failures.

Scoring: 1 point is given for each correct response. Problems 2 and 3 are sometimes given 1/2 point each (that is if the child corrects his error, when he is recounting).

There are time limits for Problems 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 of 30"; for Problems 14, 15, the limit is 45"; and for Problems 16, 17, 18 a limit of 75" is imposed.

Bonuses included in this Arithmetic test are given in Table 4.8 .

Table 4.8

Time Limits and Bonus Points with Time in Arithmetic Test.

| Design | Time Limit | Points with Time Bonus | | | |
|---|---|---|---|---|---|
| | | 7 | 6 | 5 | 4 |
| 1, 2, 3. | 45" | | | | |
| 4 | 45" | 1-10" | 11-15" | 16-20" | 21-45" |
| 5, 6, 7. | 75" | 1-10" | 11-15" | 16-20" | 21-75" |
| 8 | 75" | 1-15" | 16-20" | 21-25" | 26-75" |
| 9 | 120" | 1-25" | 26-35" | 36-55" | 56-120" |
| 10 | 120" | 1-40" | 41-55" | 56-75" | 76-120" |
| 11 | 120" | 1-40" | 41-55" | 56-80" | 81-120" |

Block Design Ages 8-16 years: begin from Item 3. If a child aged 8-16 years passes the trial of Design 3, we give him full credit for Items 1 an 2. If he fails the first trial of Design 3, we administer the second trial. Whether he passes the second trial or not, we go back and administer Designs 1 and 2 before proceeding with the test.

Discontinue: after 2 consecutive failures. (An item is considered failed only if both trials are failed).

Vocabulary Ages 11-13 years: start from item 6; ages 14-16 years: begin with Item 8. If a child aged 8-16 years obtains perfect (2 points) scores on the first two items we give him full credit for all previous items. If not, we administer the earlier items in reverse sequence until he obtains two consecutive perfect scores (not counting the starting point item).

Discontinue: after 5 consecutive failures.

Scoring Each item is scored 2, 1, or 0 as below.

2 points if a child gives 1) A good synonym. 2) A major use. 3) One or more definitive features or primary features of objects. 4) A general classification to which the word belongs. 5) A correct figurative use of the word. 6) Several less-definitive but correct descriptive features which cumulatively indicate understanding of the word. 7) Verbs: a definitive example of action or a casual relation.

1 Point for 1) A vague or less pertinent synonym. 2) A minor use, not elaborated. 3) An attribute which is correct but not definitive or not distinguishing. 4) An example using the word itself, not elaborated. 5) A concrete instance of the word, not elaborated. 6) A correct definition of a related form of the word.

Although the translation of the Vocabulary test of each word was kept as close as possible to the English language, some linguistic and cultural differences were noted, due to the fact that every language has its own modes of expression which cannot be expressed through another language by translation. An example of that is the word 'bow', which in the English language has several meanings : 1) a bow is a piece of wood curved by a tight string, used for shooting arrows; 2) bend the head or body (as a greeting) and 3) front or forward end of a boat or ship from where it begins to curve. But in the Arabic language the translation of meaning 1) has this meaning only, i.e. a piece of wood curved by a tight string, used for shooting arrows.

Directions

Say, "I am going to say some words. Listen carefully and tell me what each word means".

Proceed with the words in the order listed (starting at the appropriate place for older, normal children), saying for each one, "What is a ------ ?" or "What does ----- mean ?" With more intelligent and older children, the formal question may be omitted after the third word (Wechsler, 1974); just pronounce the word. Make certain that you are using the local pronunciation of each word, or the pronunciation you believe to be familiar to the child.

If a child is young, or if an older child is suspected of being mentally deficient, by giving less than a 2--point response to Item 1, the word "sight", say, "Well, a "sight" of something occurs when you first see it". Do not give any further help, except as indicated below.

A child is not given credit for merely pointing to an object (e.g., your ear ). If this occurs, say " Tell me in words what a ------ is". Responses are sometimes based on homonyms and are not given credit, and the examiner must ask, "What else does ------- mean ? Do not spell the word for the child.

Similarly, a child may hear a word incorrectly and respond to a different word (e.g., he may define 'red' instead of 'head,'). If this occurs, the examiner must say, "Listen carefully. What does -------- mean ? "Occasionally, it is difficult for the examiner to decide whether the child does or does not know the meaning of a word. In such instances it is permissible to say, "Explain what you mean. Tell me more about it", or some make similar neutral inquiry. However, no other form of questioning may be used. This same inquiry is permitted when the child's response could thus be evoked. (In the case of a clear-cut 0-point response or a clear-cut 1-point response, no such inquiry should be made).

All word meanings recognised by standard dictionaries are acceptable (and scored according to the quality of the definition). Regionalisms and slang not found in dictionaries are to be scored 0. If such a response is given, or if the examiner is in doubt about the acceptability of a response, he should ask the child for another meaning.

Object Assembly We begin with the sample item and then proceed to Item 1 for all children.

Discontinue: We give the entire test to all children.

There are bonuses and time limits for this subtest, as shown in Table 4.9.

Comprehension We begin with Item 1 for all children.

Discontinue: after 4 consecutive failures.

Scoring. Each item is scored 2, 1, or 0, depending on the degree of understanding expressed and the quality of the response. The examiner should match the child's response against the general criteria and the sample answers given in the WISC-R Manual for each question.

Table 4.9

Time Limits and Bonus Points for the Object Assembly Test.

| Item | Time Limit | | | Points with Time Bonus | |
|---|---|---|---|---|---|
| | 9 | 8 | 7 | 6 | 5 |
| 1 | 120" | 1-20" | 21-30" | 31-120" | |
| 2 | 150" | 1-15" | 16-20" | 21-35" | 36-150" |
| 3 | 150" | 1-25" | 26-35" | 36-50" | 51-150" |
| 4 | 180" | 1-35" | 36-50" | 51-75" | 76-180" |

Coding B is for children 8 years and over.

Scoring: 1 point for each item filled in correctly. ( the five sample items are not included in the child's score).

# CHAPTER FIVE

## Results of the Study

Chapter Five

# CHAPTER FIVE

## 5.0 Results and Analyses of The Data of Main Study

Ten of the twelve subtests of the WISC-R were standardized, the subtests Digit Span and Mazes being both excluded from this study. For each of the ten subtests the distribution of raw scores for each age--level of a mixed group of boys and girls, was converted to a scale score with a mean of ten and Standard Deviation of three. For the students ranging in age from 12 years 0 months to 15 years 11 months 30 days, Appendix C presents the scaled score equivalents to raw scores, at fourmonth intervals, for each of the ten tests. The scaled scores range from 1 to 19, which provides a range of three Standard Deviations on either side of the mean as Wechsler (1974) did when he standardized his test. Appendix D shows the diagram Wechsler included for the scaled scores ranging from 1 to 19, noting that

> In keeping with the author's theory of measurement, which stresses the comparison of a child with his chronological age peers, the WISC-R IQs are based on scaled scores derived separately for each age group.

The same scaling procedure was followed in the present study. This was accomplished by preparing a cumulative frequency distribution of raw scores for each age group, normalizing the distribution, and computing the appropriate scaled score for each raw score. For each test, the progression of scaled scores from each age to the next was examined, and some minor irregularities that were found were eliminated by smoothing. These calculations were done on the University of Baghdad computer, using the Wechsler program which gave the scaled scores corresponding to raw scores suitably normalized for each subscale, at four-month intervals. These scaled equivalents of raw scores for each age group are given in Appendix C.

A similar method to that used in the WISC-R Manual (1974) was followed to calculate the Verbal, Performance and Full Scale IQ values for the WISC-R applied to Iraqi school children. The sums of scaled scores of each age--group (using equivalent values for raw scores as given in Appendix C) for the Verbal, Performance and Full Scale results were obtained for each child in the standardization sample. These were based on ten subtest scores (five Verbal and five Performance tests). For each age group, the means and Standard Deviations of the three separate sums of scaled scores were completed as shown in Table 5.1 .

Table 5.2 presents the three sums of scaled scores of the British Amendments, the American standardization of the WISC-R and the corresponding results found in the present study . From Table 5.2 it can be seen that there was some degree of similarity in the results within each age group in the present study, American standardization sample and British Amendments, as is to be expected since all the results were scaled in the same way.

For the ten subtests the means and Standard Deviations of the scaled scores were calculated, and these results are presented in Table 5.3, scaled to a mean of ten and Standard Deviation of three, as was done by Wechsler (1974) with his data.

Table 5.4 presents the IQ equivalents of scaled scores for the Verbal Scales. Table 5.5 shows the IQ equivalents of sums of scaled scores for the Performance Scales. Table 5.6 presents the Full Scale IQ for the Iraqi school children from age 12 years 0 months 0 days to 16 years 0 months 0 days. Here again we followed the same procedure as Wechsler (1974), so that the IQs ranged from +3 2/3 to -3 2/3 Standard Deviations. The Verbal IQ scores were found to range from 46 to 152 if M = 100 and SD. = 15. The conversion formula for translating each sum of scaled scores into an IQ score is:

$$IQ = \left(\frac{X-\bar{X}}{S.D.}\right) \times 15 + 100$$

where X = sum of scaled scores for a student's achievement on either Verbal, Performance or Full Scale subtests.

$\bar{X}$ = mean of sums of scaled scores within an age group.

S.D. = Standard Deviation of sums of scaled scores within this age group.

For example: A sum of scaled scores of 10 is converted as follows:

$$IQ = \left(\frac{10 - 49.96}{12.03}\right) X\ 15 + 100 = (-3.3216\ X\ 15) + 100$$

as -3.3216 X 15 = -49.82543, then

IQ= -49.82543 + 100 = 50.17457. This was rounded = 50

(Data taken from Table 5.1)

Table 5.1

Means and Standard Deviations of Sums of Scaled Scores on Verbal, Performance and Full Scale IQs by Age, for Standardization Sample (N= 200 for Each Age Group).

| Group | Verbal | | Performance | | Full Scale | |
|---|---|---|---|---|---|---|
| Age | Mean | S.D. | Mean | S.D. | Mean | S.D. |
| 12 1/2 | 49.50 | 12.04 | 49.68 | 11.39 | 99.18 | 20.57 |
| 13 1/2 | 49.87 | 11.97 | 50.93 | 10.22 | 100.80 | 21.13 |
| 14 1/2 | 50.13 | 12.20 | 50.10 | 10.86 | 100.23 | 21.76 |
| 15 1/2 | 50.35 | 11.90 | 49.75 | 10.70 | 100.10 | 21.19 |
| Total | 49.96 | 12.03 | 50.11 | 10.79 | 100.28 | 21.16 |

Table 5.2

Means and Standard Deviations of Sums of Scaled Scores on the Verbal, Performance,- and Full Scales by Age for American WISC-R, British Standardization Sample and Present Study.

| Study | Age | Verbal | | Performance | | Full Scale | |
|---|---|---|---|---|---|---|---|
| | | Mean | S.D. | Mean | S.D. | Mean | S.D. |
| A.W-R. | 12 1/2 | 50.00 | 12.20 | 49.60 | 11.20 | 99.60 | 21.80 |
| B.W-R. | 12 1/2 | 49.98 | 12.25 | 49.59 | 11.22 | 99.57 | 21.82 |
| I.W-R. | 12 1/2 | 49.50 | 12.03 | 49.68 | 11.39 | 99.18 | 20.57 |
| A.W-R. | 13 1/2 | 50.60 | 12.70 | 51.00 | 11.20 | 101.60 | 21.70 |
| B.W-R. | 13 1/2 | 50.56 | 12.71 | 51.05 | 11.18 | 101.61 | 21.67 |
| I.W-R. | 13 1/2 | 49.87 | 11.97 | 50.93 | 10.22 | 100.80 | 21.13 |
| A.W-R. | 14 1/2 | 50.20 | 12.60 | 49.60 | 10.70 | 99.80 | 21.40 |
| B.W-R. | 14 1/2 | 50.22 | 12.64 | 49.63 | 10.67 | 99.85 | 21.36 |
| I.W-R. | 14 1/2 | 50.13 | 12.20 | 50.10 | 10.86 | 100.23 | 21.76 |
| A.W-R. | 15 1/2 | 50.23 | 11.60 | 50.47 | 10.82 | 101.26 | 20.30 |
| B.W-R. | 15 1/2 | 50.63 | 11.68 | 50.56 | 10.90 | 101.19 | 20.19 |
| I.W-R. | 15 1/2 | 50.35 | 11.90 | 49.75 | 10.70 | 100.10 | 21.21 |

A.W-R.= American WISC-R. B.W-R.= British WISC-R.

I.W-R.= Iraqi WISC-R.(present study).

Table 5.3

Means and Standard Deviations of Scaled Scores by Age-Group, for Each Subtest in the Present Study.

| Test | 12 yrs. | | 13 yrs. | | 14 yrs. | | 15 yrs. | |
|---|---|---|---|---|---|---|---|---|
| | Mean | S.D. | Mean | S.D. | Mean | S.D. | Mean | S.D. |
| Info. | 9.9 | 2.9 | 10.0 | 3.0 | 10.1 | 3.0 | 9.9 | 3.0 |
| Simi. | 9.8 | 3.0 | 9.9 | 3.0 | 10.1 | 2.9 | 10.0 | 2.9 |
| Ari. | 9.9 | 3.0 | 10.0 | 2.9 | 9.9 | 2.9 | 9.9 | 3.0 |
| Com. | 10.1 | 3.0 | 9.8 | 2.9 | 10.0 | 3.0 | 9.9 | 2.8 |
| Voca. | 9.9 | 3.0 | 9.9 | 3.0 | 10.2 | 3.0 | 10.1 | 3.0 |
| P.C. | 9.9 | 3.0 | 10.0 | 3.1 | 10.3 | 3.1 | 9.7 | 2.8 |
| P.A. | 10.0 | 3.1 | 10.0 | 3.0 | 10.1 | 3.0 | 9.8 | 2.9 |
| B.D. | 9.8 | 3.0 | 10.0 | 3.0 | 9.9 | 2.8 | 10.0 | 3.0 |
| O.A. | 9.9 | 3.1 | 10.0 | 3.0 | 10.2 | 3.0 | 10.2 | 3.0 |
| Cod. | 9.9 | 3.1 | 9.9 | 3.0 | 10.0 | 2.8 | 9.8 | 2.9 |

In.= Information Simi.= Similarity Ari.= Arithmetic

Com.= Comprehension Voca.= Vocabulary

P.C.=Picture Completion P.A.= Picture Arrangement

B.D.= Block Design O.A.= Object Assembly

Table 5.4

IQ Equivalents of Sums of Scaled Scores for Verbal Scale, for the Total Present Sample (N=800) (Based on a Mean of Sum of Scaled Scores = 49.96 and a Standard Deviation = 12.03)

| Sum of Scaled Scores | IQ | Sum of Scaled Scores | IQ | Sum of Scaled Scores | IQ |
|---|---|---|---|---|---|
| 6 | 46 | 35 | 81 | 64 | 118 |
| 7 | 46 | 36 | 83 | 65 | 119 |
| 8 | 48 | 37 | 84 | 66 | 120 |
| 9 | 49 | 38 | 86 | 67 | 121 |
| 10 | 50 | 39 | 86 | 68 | 122 |
| 11 | 51 | 40 | 88 | 69 | 124 |
| 12 | 53 | 41 | 89 | 70 | 125 |
| 13 | 54 | 42 | 90 | 71 | 126 |
| 14 | 55 | 43 | 91 | 72 | 127 |
| 15 | 56 | 44 | 93 | 73 | 129 |
| 16 | 58 | 45 | 93 | 74 | 130 |
| 17 | 60 | 46 | 95 | 75 | 131 |
| 18 | 60 | 47 | 96 | 76 | 132 |
| 19 | 61 | 48 | 98 | 77 | 134 |
| 20 | 63 | 49 | 99 | 78 | 135 |
| 21 | 64 | 50 | 100 | 79 | 136 |
| 22 | 65 | 51 | 101 | 80 | 137 |
| 23 | 66 | 52 | 103 | 81 | 139 |
| 24 | 68 | 53 | 104 | 82 | 140 |
| 25 | 69 | 54 | 105 | 83 | 141 |

(continued Table 5.4)

| | | | | | |
|---|---|---|---|---|---|
| 26 | 70 | 55 | 106 | 84 | 142 |
| 27 | 71 | 56 | 108 | 85 | 144 |
| 28 | 73 | 57 | 109 | 86 | 145 |
| 29 | 74 | 58 | 110 | 87 | 146 |
| 30 | 75 | 59 | 112 | 88 | 147 |
| 31 | 76 | 60 | 113 | 89 | 149 |
| 32 | 78 | 61 | 114 | 90 | 150 |
| 33 | 79 | 62 | 115 | 91 | 152 |
| 34 | 80 | 63 | 116 | 92 | 152 |

Verbal score is the sum of scaled scores on 5 tests.

Table 5.5

IQ Equivalents of Sum of Scaled Scores for Performance Scale for Present Sample (Based on a Mean of Sum of Scaled Scores = 50.11 and a Standard Deviation = 10.79).

| Sum of Scale Scores | IQ | Sum of Scale Scales | IQ | Sum of Scale Scales | IQ |
|---|---|---|---|---|---|
| 9 | 43 | 38 | 83 | 66 | 122 |
| 10 | 44 | 39 | 85 | 67 | 123 |
| 11 | 46 | 40 | 86 | 68 | 125 |
| 12 | 47 | 41 | 87 | 69 | 126 |
| 13 | 48 | 42 | 89 | 70 | 128 |

(continued Table 5.5)

| | | | | | |
|---|---|---|---|---|---|
| 14 | 50 | 43 | 90 | 71 | 129 |
| 15 | 52 | 44 | 92 | 72 | 130 |
| 16 | 53 | 45 | 93 | 73 | 132 |
| 17 | 54 | 46 | 94 | 74 | 133 |
| 18 | 55 | 47 | 96 | 75 | 135 |
| 19 | 57 | 48 | 97 | 76 | 136 |
| 20 | 58 | 49 | 98 | 77 | 137 |
| 21 | 60 | 50 | 100 | 78 | 139 |
| 22 | 61 | 51 | 101 | 79 | 140 |
| 23 | 63 | 52 | 103 | 80 | 142 |
| 24 | 64 | 53 | 104 | 81 | 143 |
| 25 | 65 | 54 | 105 | 82 | 144 |
| 26 | 66 | 55 | 107 | 83 | 146 |
| 27 | 68 | 56 | 108 | 84 | 147 |
| 28 | 69 | 57 | 110 | 85 | 149 |
| 29 | 71 | 58 | 111 | 86 | 150 |
| 30 | 72 | 59 | 112 | 87 | 151 |
| 31 | 73 | 60 | 114 | 88 | 153 |
| 32 | 75 | 61 | 115 | 89 | 154 |
| 33 | 76 | 62 | 117 | 90 | 155 |
| 34 | 78 | 63 | 119 | | |
| 35 | 80 | 64 | 119 | | |
| 36 | 80 | 65 | 121 | | |
| 37 | 84 | | | | |

Performance score is the sum of scaled scores on 5 tests.

Table 5.6

IQ Equivalents of Sums of Scaled Scores for Full Scale, for Present Sample (Based on a Mean of Scaled Scores = 100.28 and a Standard Deviation =21.16).

| Sum of Scale Scores | IQ | Sum of Scale Scores | IQ | Sum of Scale Scores | IQ |
|---|---|---|---|---|---|
| 15 | 40 | 72 | 80 | 128 | 120 |
| 16 | 40 | 73 | 81 | 130 | 121 |
| 17 | 41 | 74 | 81 | 131 | 122 |
| 18 | 42 | 75 | 82 | 132 | 122 |
| 19 | 42 | 76 | 82 | 133 | 123 |
| 20 | 43 | 77 | 83 | 134 | 124 |
| 21 | 44 | 78 | 84 | 135 | 125 |
| 22 | 45 | 79 | 85 | 136 | 125 |
| 23 | 45 | 80 | 86 | 137 | 126 |
| 24 | 46 | 81 | 86 | 138 | 127 |
| 25 | 47 | 82 | 87 | 139 | 127 |
| 26 | 47 | 83 | 87 | 140 | 128 |
| 27 | 48 | 84 | 88 | 141 | 129 |
| 28 | 49 | 85 | 89 | 142 | 130 |
| 29 | 49 | 86 | 90 | 143 | 130 |
| 30 | 50 | 87 | 91 | 144 | 13 |
| 31 | 51 | 88 | 91 | 145 | 132 |
| 32 | 52 | 89 | 92 | 146 | 132 |
| 33 | 52 | 90 | 93 | 147 | 133 |
| 34 | 53 | 91 | 93 | 148 | 134 |
| 35 | 54 | 92 | 94 | 149 | 135 |

(continued Table 5.6)

| | | | | | |
|---|---|---|---|---|---|
| 36 | 54 | 93 | 95 | 150 | 135 |
| 37 | 55 | 94 | 96 | 151 | 136 |
| 38 | 56 | 95 | 96 | 152 | 137 |
| 39 | 57 | 96 | 97 | 153 | 137 |
| 40 | 57 | 97 | 98 | 154 | 138 |
| 41 | 58 | 98 | 98 | 155 | 139 |
| 42 | 59 | 99 | 99 | 156 | 139 |
| 43 | 59 | 100 | 100 | 157 | 140 |
| 44 | 60 | 101 | 101 | 158 | 141 |
| 45 | 61 | 102 | 101 | 159 | 142 |
| 46 | 62 | 103 | 102 | 160 | 142 |
| 47 | 62 | 104 | 103 | 161 | 143 |
| 48 | 63 | 105 | 104 | 162 | 144 |
| 49 | 64 | 106 | 104 | 163 | 144 |
| 50 | 64 | 107 | 105 | 164 | 145 |
| 51 | 65 | 108 | 105 | 165 | 146 |
| 52 | 66 | 109 | 106 | 166 | 147 |
| 53 | 66 | 110 | 107 | 167 | 147 |
| 54 | 67 | 111 | 108 | 168 | 148 |
| 55 | 68 | 112 | 108 | 169 | 149 |
| 56 | 69 | 113 | 109 | 170 | 149 |
| 57 | 69 | 114 | 110 | 171 | 150 |
| 58 | 70 | 115 | 110 | 172 | 151 |
| 59 | 70 | 116 | 111 | 173 | 152 |
| 60 | 71 | 117 | 112 | 174 | 152 |
| 61 | 72 | 118 | 113 | 175 | 153 |

| | | | | | |
|---|---|---|---|---|---|
| 62 | 73 | 119 | 113 | 176 | 154 |
| 63 | 74 | 120 | 114 | 177 | 154 |
| 64 | 74 | 121 | 115 | 178 | 155 |
| 65 | 75 | 122 | 115 | 179 | 156 |
| 66 | 76 | 123 | 116 | 180 | 157 |
| 67 | 76 | 124 | 117 | 181 | 158 |
| 68 | 77 | 125 | 118 | 182 | 160 |
| 69 | 78 | 126 | 118 | 183 | 161 |
| 70 | 79 | 127 | 119 | 184 | 162 |
| 71 | 79 | | | | |

These Full Scale equivalent IQ scores are based on the sum of scaled scores on ten tests.

There is no child in this study or in the British WISC-R who obtained a sum score of less than 5 in the Verbal or Performance tests, since as for all ages, raw scores of even 0 have been assigned scaled scores of at least 1 for each of the tests scaled on a mean of 100 and a Standard Deviation of 15. ( Wechsler used the same procedure).

Table 5.7 shows how the British and Iraqi WISC-R Full Scale IQs seem to compare in terms of the proportions of the sample populations which fall within the different intelligence levels. These results are based on a British standardization sample of N = 2200 students, and the present study standardization sample of N =800 students.

Table 5.7

Percentages of the Student Samples within Each IQ Range Based on Full Scale.

| IQ | Classification | Theoretical Normal Curve | Percent Br.WISC-R | Percent Ir.WISC-R |
|---|---|---|---|---|
| 130& over | Very Superior | 2.2 | 2.3 | 2.0 |
| 120-129 | Superior | 6.7 | 7.4 | 7.0 |
| 110-119 | Bright | 16.1 | 16.5 | 16.8 |
| 90-109 | Average | 50.0 | 49.4 | 50.9 |
| 80-89 | Low Average | 16.1 | 16.2 | 15.2 |
| 70-79 | Borderline | 6.7 | 6.0 | 5.7 |
| 69- | Mentally Defic. | 2.2 | 2.2 | 2.4 |

## 5.1 Reliability

The reliability coefficients for each subtest and for Verbal, Performance and Full Scale for each age group were calculated by two methods, namely the Split--half and Kuder-Richardson Formula 20. The Split-half was accomplished by using the odd-numbered items as one set of scores and the even-numbered items as the other set. In other words, for each test we obtained separate scores for the odd numbered and even-numbered items. The correlation between these two scores gave an estimate of the reliability. The reliability coefficient is a measure of the amount of inconsistency; it does not indicate, however, the causes of any lack of consistency. It tells how much the scores may be expected to vary, not why they vary. Thus, if we have reliability coefficients for any two tests, obtained under similar conditions, we can tell which of the two tests measures more consistently. But without further analyses we cannot know why one test measures more accurately than the other. The standard error of measurement was used because the reliability of the test may be expressed in terms of this error (See Brown 1976). To estimate the child's true score the magnitude of the error component is required. The purpose of these statistical analyses of the sample results was to establish positively whether or not the present WISC-R was a reliable test for the assessment of Iraqi school children.

Speed is a factor that can influence reliability. In fact, the Split half and Kuder Richardson formula 20 reliability procedures are inappropriate when speed is a major factor in test performance. (Brown, 1976 p.88).

Table 5.8 presents the reliability coefficients of the scaled scores of each of the subtests with the corresponding Verbal, Performance and Full Scale IQs. The reliabilities are of two kinds: Split half and Kuder-Richardson Formula 20. The Kuder-Richardson correlation, which provides a measure of internal consistency, was not appropriate for Coding B, because the test is a speeded test. It was difficult to find the reliability by Kuder-Richardson Formula 20 for the Object Assembly as no meaningful result was found. The same was true of the British WISC-R.

The standard error of measurement was used to estimate the degree of possible variation in students' individual scores for each age--roup. Table 5.9 presents the standard error of measurement for each of the four age groups for the scaled scores. The standard error of measurement was computed by the formula SEm = S.D. $\sqrt{(1-r)}$ where S.D.= Standard Deviation of the scaled score units and r = reliability coefficient. Thus, for the Information test, S.D. = 2.9 (See Table 5.3) for 12 year olds and r= 0.81 (See Table 5.8), the value of SEm = 2.9 $\sqrt{1-0.81}$ = 2.9 X 0.43 = 1.26. (See Table 5.9). (See Wechsler 1974)

As we can see from Table 5.10, the standard error of measurement of scaled scores differs within each of the tests. Wechsler (1974, p. 29) noted that:

> Awareness of these facts is of particular importance when the clinician evaluates children's test profiles.

Table 5.11 shows the reliabilities of the present study for the four age groups of each set of scaled scores and the respective reliabilities for the British WISC-R standardization sample, (Split-half only). From this table we can conclude that the reliability of the present test is similar to the British one. Table 5.10 presents the standard error of measurement of each set of scaled scores for each age-group (based on Split half reliabilities) and for the British WISC-R standardization sample. These results also indicate a reliability of the present results to similar those of the British WISC-R. Table 5.12 presents the intercorrelations of the scaled scores of the subtests for each age group 12, 13, 14, and 15 year-old boys and girls taken together. These results show that the intercorrelations between subtests were found to correlate moderately with each other, and highly between Verbal, Performance and Full Scale. Similar results were found with the British WISC-R.

Table 5.8

Reliability Coefficients of the Tests and IQ Scales by Age (Split-half and Kuder-Richardson formula 20) N= 200 for Each Age Group

| Test | 12 years | | 13 years | | 14 years | | 15 years | |
|---|---|---|---|---|---|---|---|---|
| | S-H | K-R20 | S-H | K-R20 | S-H | K-R20 | S-H | K-R20 |
| Infor. | 0.81 | 0.77 | 0.79 | 0.73 | 0.79 | 0.74 | 0.86 | 0.72 |
| Simil. | 0.78 | 0.71 | 0.70 | 0.68 | 0.80 | 0.67 | 0.78 | 0.75 |
| Arith. | 0.77 | 0.74 | 0.72 | 0.65 | 0.73 | 0.68 | 0.75 | 0.82 |
| Comp. | 0.80 | 0.83 | 0.81 | 0.89 | 0.76 | 0.90 | 0.69 | 0.85 |
| Voca. | 0.90 | 0.86 | 0.82 | 0.91 | 0.87 | 0.94 | 0.88 | 0.90 |
| P.C. | 0.76 | 0.81 | 0.70 | 0.76 | 0.77 | 0.65 | 0.69 | 0.72 |
| P.A. | 0.73 | 0.89 | 0.71 | 0.86 | 0.69 | 0.81 | 0.65 | 0.83 |
| B.D. | 0.82 | 0.91 | 0.80 | 0.94 | 0.79 | 0.88 | 0.82 | 0.95 |
| O.A. | 0.52 | - | 0.49 | - | 0.64 | - | 0.73 | - |
| Coding | - | - | - | - | - | - | - | |
| Verbal | 0.91 | 0.92 | 0.90 | 0.91 | 0.90 | 0.94 | 0.92 | 0.94 |
| Perfr. | 0.88 | 0.93 | 0.85 | 0.86 | 0.82 | 0.88 | 0.86 | 0.95 |
| F. S. | 0.94 | 0.96 | 0.93 | 0.95 | 0.89 | 0.92 | 0.91 | 0.96 |

S-H = Split-half

K-R20 = Kuder-Richardson formula 20

Note: The Reliability Coefficients of the Kuder-Richardson Formula 20 were computed from scaled scores. The verbal score is the sum of scaled scores on 5 tests. Performance Scale is the sum of the scaled scores on 5 tests. Full Scale is the sum of scaled scores on 10 tests.

Table 5.9

Standard Error of Measurement (SEm) of the Scaled Scores and IQs, by Age. N= 200 for Each Age Group. (Based on Split-half Reliabilities).

| Test | 12 Yrs. | | 13 Yrs. | | 14 Yrs. | | 15 Yrs. | |
|---|---|---|---|---|---|---|---|---|
| | S-H | K-R20 | S-H | K-R20 | S-H | K-R20 | S-H | K-R20 |
| Inf. | 1.26 | 1.60 | 1.37 | 1.52 | 1.37 | 1.62 | 1.12 | 1.56 |
| Sim. | 1.40 | 1.87 | 1.64 | 1.77 | 1.29 | 1.65 | 1.36 | 1.69 |
| Ari. | 1.50 | 1.44 | 1.23 | 1.86 | 1.19 | 1.57 | 1.50 | 1.61 |
| Com. | 1.43 | 1.10 | 1.29 | 0.94 | 1.61 | 0.93 | 1.81 | 1.10 |
| Voc. | 0.94 | 1.08 | 1.27 | 0.90 | 1.08 | 0.91 | 1.03 | 0.98 |
| P.C. | 1.45 | 1.24 | 1.69 | 1.47 | 1.72 | 1.79 | 1.68 | 1.71 |
| P.A. | 1.61 | 1.02 | 1.61 | 1.06 | 1.61 | 1.28 | 1.89 | 1.19 |
| B.D. | 1.27 | 0.97 | 1.34 | 0.89 | 1.28 | 1.04 | 1.27 | 0.76 |
| O.A. | 2.23 | - | 2.26 | - | 1.80 | - | 1.55 | - |
| Cod. | - | - | - | - | - | - | - | - |
| V.IQ | 3.60 | 4.19 | 3.78 | 3.84 | 3.85 | 3.96 | 3.36 | 3.62 |
| P.IQ | 4.87 | 3.42 | 3.95 | 4.89 | 3.76 | 4.31 | 4.00 | 3.52 |
| F.IQ | 5.03 | 4.07 | 5.59 | 4.63 | 5.79 | 4.15 | 5.36 | 4.17 |

Note: The standard errors of measurement were computed from scaled scores.

Table 5.10

Standard Error of Measurement (SEm) of the Scaled Scores and IQs for the Present Study (Based on Split-half Reliabilities) and for British WISC-R Standardization Sample.

| | SEm Present Study Based | | | | SEm British WISC-R Based | | | |
|---|---|---|---|---|---|---|---|---|
| Test | 12Ys. | 13Ys. | 14Ys. | 15Ys. | 12Ys. | 13Ys. | 14Ys. | 15 Ys. |
| Inf. | 1.26 | 1.37 | 1.37 | 1.12 | 1.06 | 1.14 | 1.08 | 0.93 |
| Sim. | 1.40 | 1.64 | 1.29 | 1.36 | 1.14 | 1.34 | 1.34 | 1.50 |
| Ari. | 1.50 | 1.23 | 1.19 | 1.50 | 1.38 | 1.30 | 1.45 | 1.25 |
| Com. | 1.43 | 1.29 | 1.61 | 1.80 | 1.08 | 1.27 | 1.22 | 1.52 |
| Voc. | 0.94 | 1.27 | 1.08 | 1.03 | 1.08 | 1.36 | 1.19 | 1.27 |
| P.C. | 1.46 | 1.69 | 1.72 | 1.68 | 1.50 | 1.61 | 1.54 | 1.89 |
| P.A. | 1.61 | 1.61 | 1.61 | 1.89 | 1.51 | 1.59 | 1.54 | 1.59 |
| B.D. | 1.27 | 1.34 | 1.28 | 1.27 | 1.13 | 1.14 | 1.14 | 1.09 |
| O.A. | 2.23 | 2.26 | 1.18 | 1.55 | 1.84 | 1.71 | 1.68 | 1.82 |
| Cod. | - | - | - | - | - | - | - | - |
| V. | 3.60 | 3.78 | 3.85 | 3.36 | 3.13 | 3.42 | 3.40 | 3.42 |
| P. | 3.87 | 3.95 | 3.76 | 4.00 | 4.58 | 4.96 | 4.74 | 4.84 |
| F.S. | 5.03 | 5.59 | 5.79 | 5.36 | 3.96 | 3.23 | 3.15 | 3.19 |

Note: The standard errors of measurement (SEm) of the scaled scores were computed from scaled scores.

Table 5.11

Reliability Coefficients for the Present Study (Split half Only) and for the British WISC-R.

| Test | Split half Present Study 12yrs. | 13yrs. | 14yrs. | 15yrs. | Split half British WISC-R 12yrs. | 13yrs. | 14yrs. | 15yrs. |
|---|---|---|---|---|---|---|---|---|
| Inf. | 0.81 | 0.79 | 0.79 | 0.86 | 0.87 | 0.87 | 0.88 | 0.90 |
| Sim. | 0.78 | 0.70 | 0.80 | 0.78 | 0.84 | 0.79 | 0.81 | 0.74 |
| Ari. | 0.75 | 0.72 | 0.73 | 0.75 | 0.80 | 0.81 | 0.73 | 0.80 |
| Com. | 0.77 | 0.80 | 0.71 | 0.82 | 0.88 | 0.89 | 0.91 | 0.90 |
| Voc. | 0.90 | 0.82 | 0.87 | 0.88 | 0.87 | 0.81 | 0.82 | 0.72 |
| P.C. | 0.76 | 0.70 | 0.69 | 0.64 | 0.75 | 0.75 | 0.72 | 0.68 |
| P.A. | 0.73 | 0.71 | 0.60 | 0.58 | 0.78 | 0.72 | 0.74 | 0.73 |
| B.D. | 0.82 | 0.80 | 0.79 | 0.82 | 0.86 | 0.86 | 0.84 | 0.85 |
| O.A. | 0.48 | 0.43 | 0.64 | 0.73 | 0.63 | 0.72 | 0.72 | 0.68 |
| Cod. | - | - | - | - | - | - | - | - |
| Ver. | 0.91 | 0.90 | 0.90 | 0.92 | 0.96 | 0.96 | 0.96 | 0.95 |
| Per. | 0.88 | 0.85 | 0.88 | 0.86 | 0.91 | 0.90 | 0.89 | 0.90 |
| F.S. | 0.94 | 0.93 | 0.89 | 0.91 | 0.96 | 0.96 | 0.96 | 0.95 |

Note: The Reliability Coefficients of the Split--half were computed from scaled scores. The Verbal score is the sum of scaled scores on 5 tests. Performance score is the sum of scaled scores on 5 tests. Full Scale is the sum of scaled scores on 10 tests.

Table 5.12

Intercorrelation of the Subtests, by Age-Group (Scaled Scores)

Age 12 1/2 100 boys and 100 girls

| Test | Information | Similarities | Arithmetic | Vocabulary | Comprehension | Picture Completion | Picture Arrangement | Block Design | Object Assembly | Coding B | V. | P. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Similarities | .52 | | | | | | | | | | | |
| Arithmetic | .49 | .34 | | | | | | | | | | |
| Vocabulary | .65 | .58 | .51 | | | | | | | | | |
| Comprehension | .43 | .60 | .41 | .55 | | | | | | | | |
| Picture Compilation | .50 | .47 | .23 | .38 | .40 | | | | | | | |
| Picture Arrangement | .51 | .30 | .28 | .44 | .34 | .57 | | | | | | |
| Block Design | .56 | .53 | .55 | .50 | .44 | .61 | .51 | | | | | |
| Object Assembly | .45 | .46 | .40 | .39 | .35 | .48 | .53 | .62 | | | | |
| Coding B | .30 | .20 | .28 | .31 | .29 | .11 | .17 | .39 | .24 | | | |
| Verbal | .77 | .68 | .70 | .81 | .35 | .41 | .49 | .59 | .50 | .33 | | |
| Performance | .58 | .52 | .48 | .45 | .43 | .63 | .66 | .76 | .74 | .47 | .64 | |
| Full Scale | .72 | .68 | .63 | .67 | .54 | .60 | .63 | .83 | .75 | .32 | .88 | .90 |

Note: The Coefficients of Correlation were computed from Scaled Scores.
Verbal Score is the Sum of scaled scores on 5 tests : Performance score is the Sum of scaled scores on 5 tests : Full Scale is the Sum of scaled scores on ten tests. Digit Span and Mazes are omitted.

Table 5.12 (continued)
Age 13 1/2

| Test | Inform. | Similar. | Arith. | Vocab. | Comprehen. | P.C. | P.A. | B.D. | O.A. | Coding B | V. | P. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Similar. | .54 | | | | | | | | | | | |
| Arith. | .51 | .41 | | | | | | | | | | |
| Vocab. | .67 | .62 | .53 | | | | | | | | | |
| Comp. | .45 | .57 | .44 | .53 | | | | | | | | |
| P.C. | .48 | .50 | .31 | .39 | .43 | | | | | | | |
| P.A. | .49 | .37 | .30 | .42 | .36 | .51 | | | | | | |
| B.D. | .54 | .55 | .48 | .52 | .45 | .60 | .50 | | | | | |
| O.A. | .39 | .35 | .31 | .37 | .33 | .40 | .43 | .55 | | | | |
| Coding B | .36 | .31 | .20 | .25 | .32 | .23 | .21 | .39 | .43 | | | |
| V. | .79 | .73 | .68 | .86 | .49 | .44 | .47 | .62 | .41 | .35 | | |
| P. | .59 | .56 | .47 | .51 | .48 | .59 | .62 | .75 | .63 | .38 | .91 | |
| Full S. | .73 | .69 | .65 | .74 | .58 | .63 | .52 | .81 | .67 | .36 | .90 | .88 |

Table 5.12 (continued)
Age 14 1/2

| | Inform. | Similar. | Arith. | Vocab. | Comprehen. | P.C. | P.A. | B.D. | O.A. | Coding B | V. | P. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Similar. | .60 | | | | | | | | | | | |
| Arith. | .59 | .48 | | | | | | | | | | |
| Vocab. | .73 | .68 | .57 | | | | | | | | | |
| Compil. | .51 | .61 | .40 | .62 | | | | | | | | |
| P.C. | .49 | .48 | .32 | .43 | .46 | | | | | | | |
| P.A. | .45 | .41 | .36 | .43 | .39 | .50 | | | | | | |
| B.D. | .52 | .55 | .51 | .56 | .48 | .64 | .53 | | | | | |
| O.A. | .43 | .47 | .36 | .35 | .40 | .46 | .48 | .59 | | | | |
| Coding B | .26 | .32 | .28 | .22 | .30 | .19 | .23 | .31. | 33 | | | |
| V. | .81 | .75 | .69 | .84 | .58 | .47 | .50 | .65 | .53 | .36 | | |
| P. | .54 | .58 | .49 | .53 | .48 | .62 | .64 | .77 | .66 | .39 | .92 | |
| Full S. | .75 | .70 | .61 | .67 | .56 | .63 | .58 | .80 | .65 | .34 | .92 | .91 |

Table 5.12 (continued)

Age 15 1/2 years

| | Inform. | Similar. | Vocab. | Comprehen. | P.C. | P.A. | B.D. | O.A. | Coding B | V. | P. | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Similar. | .58 | | | | | | | | | | | |
| Arith. | .52 | .43 | | | | | | | | | | |
| Vocab. | .71 | .66 | .52 | | | | | | | | | |
| Comprehen. | .50 | .59 | .38 | .60 | | | | | | | | |
| P.C. | .47 | .44 | .31 | .40 | .48 | | | | | | | |
| P.A. | .33 | .36 | .34 | .44 | .37 | .45 | | | | | | |
| B.D. | .51 | .50 | .52 | .54 | .48 | .61 | .49 | | | | | |
| O.A. | .41 | .43 | .29 | .33 | .37 | .46 | .45 | .58 | | | | |
| Coding B | .27 | .30 | .31 | .23 | .28 | .20 | .24 | .29 | .28 | | | |
| V. | .74 | .69 | .71 | .83 | .40 | .39 | .46 | .57 | .48 | .29 | | |
| P. | .56 | .55 | .48 | .54 | .46 | .60 | .68 | .73 | .65 | .35 | .96 | |
| Full S. | .71 | .68 | .58 | .66 | .52 | .62 | .54 | .81 | .66 | .33 | .88 | .90 |

## Gender Differences

The means and Standard Deviations of scaled scores were also calculated in the present study for each subtest within each age--roup. Boys' and girls' scaled scores were calculated separately. (These results are discussed in Chapter 6). Table 5.14 shows the means and Standard Deviations of scaled scores obtained in the present study for the four age--roups 12, 13, 14 and 15 years for boys and girls separately in the ten subtests of the WISC-R. These figures are in agreement with the British ones (See Table 5.14). Table 5.15 gives the means and Standard Deviations of Verbal, Performance and Full Scale IQs for boys and girls separately; differences between scaled score means (boys minus girls); standard errors of the mean differences; and t values which assess the significance of these mean differences between boys and girls for the three scale tests of Verbal, Performance and Full Scale IQs. Significant differences were found between the boys' and girls' scores in favour of the boys in all subtests except Coding B.

Tables 5.16, 5.17, 5.18, and 5.19 give us the means, standard errors of the means, differences between means for boys and girls (boys minus girls), standard errors of difference, and t values for the raw scores of the subtests of the four age groups of the present study, namely, 12, 13, 14 and 15 years. These results show that the boys obtained higher scores than the girls in most WISC-R subtests with significant differences of 1 % and 5 % levels within each age--roup.

## Age Differences

Table 5.20 shows the means and Standard Deviations of the Verbal, Performance and Full Scale IQs for each of these four age groups, the differences between means for the age groups 15-14, 15-13, and 15-12 years, and t values for these differences. It was decided to use this method to find out if there were any significant differences between these ages. The non-significant means differences indicate that the scale procedures based on selected means (50 for Verbal and Performance, 100 for Full Scale IQs) for each year group were applied correctly.

**Order of Difficulty within each Subtest Based on the Main Sample.**

In order to find the difficulty level for each item, the percentage of passes for each age group was obtained for every item in the test. It will be remembered that the criterion for passing an item, for this study, was any score greater than zero.

Table 5.21 presents the proportion of subjects (boys and girls together) passing each item for the four age groups, and the order of difficulty for each item of each subtest for the four age groups combined. Table 5.13 presents the Spearman rank-order correlation coefficients of items within each subtest between the new order and the WISC-R British Amendments. Most items are in agreement as to the order of difficulty with the British WISC-R, as we can see from Table 5.13.

Table 5.13
Spearman Rank Order Correlation Coefficient

Between New Order Subtests and the British WISC-R.

| | | | |
|---|---|---|---|
| Information | 0.91 | Picture Completion | 0.89 |
| Similarities | 0.94 | Picture Arrangement | 0.94 |
| Arithmetic | 0.98 | Block Design | 0.99 |
| Vocabulary | — | Object Assembly | 0.73 |
| Comprehen. | 0.92 | Coding B | — |

Table 5.14
Means and Standard Deviations of Scaled Scores of the Ten Subtests for Girls and Boys within Each of the Four Age Groups in the Present Study

| Test | Gender | 12 years | | 13 years | | 14 years | | 15 years | |
|---|---|---|---|---|---|---|---|---|---|
| | | M. | S.D. | M. | S.D. | M. | S.D. | M. | S.D. |
| Information | B | 10.26 | 3.19 | 10.93 | 3.30 | 11.71 | 2.89 | 10.97 | 2.35 |
| | G | 9.48 | 2.67 | 9.02 | 3.15 | 8.63 | 2.92 | 9.23 | 3.41 |
| Similarities | B | 10.07 | 3.13 | 10.12 | 3.26 | 10.40 | 2.84 | 10.44 | 2.81 |
| | G | 8.90 | 2.84 | 9.31 | 2.97 | 9.83 | 3.02 | 9.74 | 3.13 |
| Arithmetic | B | 10.85 | 2.65 | 10.73 | 3.16 | 11.04 | 2.77 | 10.80 | 2.72 |
| | G | 9.16 | 2.87 | 8.89 | 2.78 | 10.11 | 3.15 | 9.25 | 2.89 |
| Comprehension | B | 10.20 | 3.10 | 10.85 | 2.97 | 10.62 | 2.72 | 11.14 | 2.70 |
| | G | 8.73 | 2.88 | 9.08 | 2.84 | 9.29 | 2.91 | 9.36 | 3.13 |
| Vocabulary | B | 10.85 | 3.42 | 10.21 | 3.19 | 10.39 | 2.93 | 10.61 | 2.82 |
| | G | 9.78 | 3.09 | 9.30 | 3.22 | 9.81 | 2.87 | 9.68 | 3.38 |
| Picture Compilation | B | 9.78 | 3.03 | 10.21 | 2.96 | 11.01 | 3.06 | 10.19 | 2.81 |
| | G | 9.65 | 2.82 | 9.20 | 2.90 | 9.66 | 2.85 | 9.48 | 3.28 |
| Picture Arrangement | B | 9.94 | 3.31 | 11.05 | 2.47 | 10.36 | 2.78 | 10.31 | 2.60 |
| | G | 8.98 | 2.85 | 9.80 | 3.15 | 9.61 | 3.23 | 9.67 | 3.12 |
| Block Design | B | 10.41 | 3.25 | 10.03 | 3.06 | 10.49 | 2.98 | 10.44 | 2.87 |
| | G | 9.67 | 2.85 | 9.79 | 2.73 | 9.76 | 3.07 | 9.30 | 2.90 |
| Object Assembly | B | 10.17 | 3.03 | 10.38 | 2.81 | 10.86 | 2.74 | 10.71 | 2.30 |
| | G | 9.73 | 2.89 | 9.69 | 3.09 | 9.18 | 3.20 | 9.29 | 3.42 |
| Coding B | B | 8.87 | 2.92 | 9.15 | 3.08 | 8.97 | 2.90 | 9.36 | 2.67 |
| | G | 10.91 | 2.49 | 10.36 | 3.31 | 10.39 | 2.89 | 10.19 | 3.19 |

Table 5.15

Mean IQs, Standard Deviations, Standard Errors of the Mean, Differences between Means, Standard Errors of Difference and t Values for Four Age-Groups (12, 13, 14 and 15 years old) for Boys and Girls

| Test | Gender | Age | Number | Mean | Standard Deviation | Standard Error of the Mean | Difference Between Means | Standard Error of the Difference | t-Value |
|---|---|---|---|---|---|---|---|---|---|
| Verbal | B | 12 | 100 | 101.96 | 14.89 | 2.15 | 4.92 | 2.63 | 1.76* |
| | G | | 100 | 97.04 | 13.10 | 1.91 | | | |
| Performance | B | 12 | 100 | 99.47 | 15.83 | 2.42 | 0.60 | 2.89 | 0.20 |
| | G | | 100 | 98.87 | 13.11 | 1.98 | | | |
| Full Scale | B | 12 | 100 | 101.84 | 15.78 | 2.23 | 4.83 | 2.95 | 1.68 |
| | G | | 100 | 97.01 | 13.02 | 1.86 | | | |
| Verbal | B | 13 | 100 | 103.87 | 14.30 | 2.10 | 6.64 | 2.98 | 2.31* |
| | G | | 100 | 97.23 | 14.49 | 2.24 | | | |
| Performance | B | 13 | 100 | 102.91 | 14.82 | 2.13 | 5.74 | 2.79 | 1.98* |
| | G | | 100 | 97.17 | 14.25 | 2.09 | | | |
| Full Scale | B | 14 | 100 | 105.96 | 14.73 | 2.16 | 9.24 | 2.83 | 3.17** |
| | G | | 100 | 96.72 | 14.48 | 2.14 | | | |
| Verbal | B | 14 | 100 | 105.36 | 12.63 | 1.82 | 8.48 | 2.76 | 3.06** |
| | G | | 100 | 96.88 | 15.14 | 2.34 | | | |
| Performance | B | 14 | 100 | 103.05 | 13.17 | 1.98 | 5.74 | 2.98 | 1.97** |
| | G | | 100 | 97.31 | 16.06 | 2.53 | | | |

Table 5.15 (continued)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Full Scale | B | 14 | 100 | 105.92 | 12.11 | 1.75 | 9.11 | 2.84 | 3.21** |
| | G | | 100 | 96.81 | 16.24 | 2.41 | | | |
| Verbal | B | 15 | 100 | 106.22 | 12.85 | 2.13 | 9.25 | 2.68 | 3.29** |
| | G | | 100 | 96.97 | 15.34 | 2.39 | | | |
| Performance | B | 15 | 100 | 104.10 | 13.04 | 1.98 | 7.04 | 3.15 | 2.49* |
| | G | | | 97.06 | 15.33 | 2.41 | | | |
| Full Scale | B | 15 | 100 | 105.72 | 12.36 | 1.80 | | | |
| | G | | 100 | 96.89 | 16.35 | 2.38 | 7.83 | 2.67 | 2.73* |

* Significant at the 0.05 level

** Significant at the 0.01 level

Table 5.16
Differences in Mean Raw Scores (Boys minus girls), Standard Errors of Difference and t-Values. Age 12 years, on Subtests

| Test | Maximum Mark | Number | Gender | Mean Raw Score | Standard Error of the Mean | Difference Between Means | Standard Error of Difference | t-Value |
|---|---|---|---|---|---|---|---|---|
| Information | 30 | 100 | B | 20.21 | 0.538 | 0.41 | 0.875 | 0.87 |
| | | 100 | G | 19.80 | 0.476 | | | |
| Similarities | 30 | 100 | B | 15.17 | 0.670 | 0.64 | 0.981 | 1.20 |
| | | 100 | G | 14.53 | 0.531 | | | |
| Arithmetic | 18 | 100 | B | 13.67 | 0.362 | 1.47 | 0.540 | 3.26** |
| | | 100 | G | 12.20 | 0.279 | | | |
| Comprehension | 34 | 100 | B | 18.20 | 0.532 | 0.92 | 1.210 | 1.39 |
| | | 100 | G | 17.38 | 0.487 | | | |
| Vocabulary | 64 | 100 | B | 33.70 | 1.103 | 0.60 | 1.643 | 0.37 |
| | | 100 | G | 33.10 | 1.065 | | | |
| Picture Compiletion | 26 | 100 | B | 16.34 | 0.537 | 0.54 | 0.974 | 0.80 |
| | | 100 | G | 15.80 | 0.423 | | | |
| Picture Arrangement | 48 | 100 | B | 19.14 | 1.362 | 0.74 | 1.965 | 0.45 |
| | | 100 | G | 18.40 | 1.114 | | | |

Table 5.16 (continued)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Block Design | 62 | 100 | B | 29.50 | 0.997 | 2.38 | 1.807 | 1.10 |
| | | 100 | G | 27.12 | 0.981 | | | |
| Object Assembly | 33 | 100 | B | 17.93 | 0.844 | 1.92 | 1.741 | 1.61 |
| | | 100 | G | 16.01 | 0.871 | | | |
| Coding | 93 | 100 | B | 39.60 | 1.067 | -0.15 | 1.530 | -0.08 |
| | | 100 | G | 39.75 | 1.136 | | | |

* Significant at the 0.05 level

** Significant at the 0.01 level

Table 5.17

Differences in Mean Raw Scores (Boys and Girls), Standard Errors of Difference and t-Values. Age 13 years, on Subtests

| Test | Maximum Mark | Number | Gender | Mean Raw Score | Standard Error of the Mean | Difference Between Means | Standard Error of Difference | t-Value |
|---|---|---|---|---|---|---|---|---|
| Information | 30 | 100 | B | 21.85 | 0.534 | 2.22 | 0.743 | 3.36** |
| | | 100 | G | 19.63 | 0.480 | | | |
| Similarities | 30 | 100 | B | 16.71 | 0.731 | 1.73 | 0.569 | 2.27* |
| | | 100 | G | 14.98 | 0.642 | | | |
| Arithmetic | 18 | 100 | B | 14.19 | 0.237 | 0.46 | 0.472 | 1.02 |
| | | 100 | G | 13.73 | 0.362 | | | |
| Comprehension | 34 | 100 | B | 18.85 | 0.345 | 1.66 | 1.361 | 2.55* |
| | | 100 | G | 17.19 | 0.321 | | | |
| Vocabulary | 64 | 100 | B | 35.15 | 1.512 | 0.48 | 1.750 | 0.32 |
| | | 100 | G | 34.67 | 1.210 | | | |
| Picture Compiletion | 26 | 100 | B | 17.31 | 0.432 | 1.33 | 0.807 | 1.12 |
| | | 100 | G | 15.98 | 0.376 | | | |
| Picture Arrangement | 48 | 100 | B | 20.01 | 1.291 | 1.72 | 1.938 | 0.91 |
| | | 100 | G | 18.29 | 1.113 | | | |
| Block Design | 62 | 100 | B | 29.97 | 1.032 | 1.46 | 1.237 | 0.68 |
| | | 100 | G | 28.51 | 1.161 | | | |

Table 5.17 (continued) Age 13

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Object Assembly | 33 | 100 | B | 18.73 | 0.765 | 2.12 | 1.442 | 1.64 |
| | | 100 | G | 16.61 | 0.563 | | | |
| Coding B | 93 | 100 | B | 40.07 | 1.073 | -3.04 | 1.841 | -1.87 |
| | | 100 | G | 43.11 | 1.241 | | | |

* Significent at the 0.05 level
** Significent at the 0.01 level

Table 5. 18

Differences in Mean Raw Score (Boys and Girls), Standard Errors of Difference and t-Values (Aged 14 years) on Subtests

| Test | Maximum Score | Number | Gender | Mean Raw Score | Standard Error of the Mean | Difference Between Means | Standard Error of Difference | t value |
|---|---|---|---|---|---|---|---|---|
| Information | 30 | 100 | B | 22.19 | 0.580 | 1.77 | 0.671 | 2.85** |
| | | 100 | G | 20.42 | 0.475 | | | |
| Similarities | 30 | 100 | B | 16.33 | 0.654 | 0.38 | 0.830 | 0.54 |
| | | 100 | G | 15.95 | 0.551 | | | |
| Arithmetic | 18 | 100 | B | 16.12 | 0.352 | 1.81 | 0.437 | 5.17** |
| | | 100 | G | 14.31 | 0.298 | | | |
| Comprehension | 34 | 100 | B | 20.06 | 0.423 | 1.70 | 0.896 | 2.65* |
| | | 100 | G | 18.36 | 0.405 | | | |
| Vocabulary | 64 | 100 | B | 37.80 | 1.149 | 2.31 | 1.726 | 1.47 |
| | | 100 | G | 35.49 | 0.972 | | | |
| Picture Completion | 26 | 100 | B | 18.41 | 0.463 | 1.68 | 0.738 | 2.40* |
| | | 100 | G | 16.73 | 0.376 | | | |
| Picture Arrangement | 48 | 100 | B | 23.30 | 1.117 | 4.33 | 1.810 | 2.32* |
| | | 100 | G | 18.97 | 1.022 | | | |
| Block Design | 62 | 100 | B | 32.78 | 1.274 | 3.05 | 2.019 | 1.50 |
| | | 100 | G | 29.63 | 1.431 | | | |

Table 5.18 Age 14

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Object Assembly | 33 | 100 | B | 21.24 | 0.653 | 2.72 | 0.969 | 2.26* |
| | | 100 | G | 18.42 | 0.540 | | | |
| Coding | 93 | 100 | B | 47.29 | 1.521 | -4.60 | 2.165 | -2.34* |
| | | 100 | G | 51.90 | 1.307 | | | |

* Significant at the 0.05 level
** Significicent at the 0.01 level

Table 5.19

Differences in Mean Raw Scores (Boys and Girls), Standard Errors of Difference and t-Values (Age 15 years) on Subtests

| Test | Maximum Mark | Number | Gender | Mean Raw Score | Standard Error of the Mean | Difference Between Between Means | Standard Error of Difference | t-Value |
|---|---|---|---|---|---|---|---|---|
| Information | 30 | 100 | B | 23.28 | 0.532 | 2.16 | 0.631 | 3.66** |
| | | 100 | G | 21.12 | 0.471 | | | |
| Similarities | 30 | 100 | B | 17.85 | 0.562 | 0.92 | 0.836 | 1.50 |
| | | 100 | G | 16.93 | 0.376 | | | |
| Arithmetic | 18 | 100 | B | 17.35 | 0.301 | 1.62 | 0.465 | 3.85** |
| | | 100 | G | 15.73 | 0.263 | | | |
| Comprehension | 34 | 100 | B | 20.78 | 0.542 | 1.66 | 0.943 | 2.76* |
| | | 100 | G | 19.12 | 0.511 | | | |
| Vocabulary | 64 | 100 | B | 41.92 | 0.985 | 2.04 | 1.734 | 1.45 |
| | | 100 | G | 39.88 | 0.831 | | | |
| Picture Completion | 26 | 100 | B | 19.12 | 0.397 | 1.11 | 0.827 | 1.81 |
| | | 100 | G | 18.01 | 0.385 | | | |
| Picture Arrangement | 48 | 100 | B | 26.51 | 0.947 | 3.87 | 1.206 | 2.44* |
| | | | G | 22.64 | 0.873 | | | |
| Block Design | 62 | 100 | B | 36.76 | 1.228 | 2.96 | 1.978 | 1.45 |
| | | 100 | G | 33.80 | 1.180 | | | |

Table 5.19 (continued)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Object Assembly | 33 | 100 | B | 24.81 | 0.776 | 4.70 | 0.890 | 4.12** |
| | | 100 | G | 20.11 | 0.634 | | | |
| Coding | 93 | 100 | B | 48.94 | 1.563 | -3.10 | 2.327 | -1.57 |
| | | 100 | G | 52.04 | 1.471 | | | |

* Significant at the 0.05 level
** Significant at the 0.01 level

Table 5.20

Mean IQs, Standard Deviations, Difference Between Means, and t-Values for 15-14 years, 15-13 years and 15-12 years (N=200). Based in Scaled Scores of both Boys and Girls.

| IQ Test | Age | Number | Mean | S.D. | Difference Between Means | t-Value |
|---|---|---|---|---|---|---|
| Verbal | 15 | 200 | 50.35 | 11.90 | 0.22 | 0.12 |
| | 14 | 200 | 50.13 | 12.20 | | |
| Performance | 15 | 200 | 49.75 | 10.70 | -0.35 | -0.23 |
| | 14 | 200 | 50.10 | 10.86 | | |
| Full Scale | 15 | 200 | 100.10 | 21.21 | -0.13 | -0.04 |
| | 14 | 200 | 100.23 | 21.73 | | |
| Verbal | 15 | 200 | 50.35 | 11.90 | 0.48 | 0.28 |
| | 13 | 200 | 49.87 | 11.97 | | |
| Performance | 15 | 200 | 49.75 | 10.70 | -1.18 | -0.80 |
| | 13 | 200 | 50.93 | 10.22 | | |
| Full Scale | 15 | 200 | 100.10 | 21.21 | -0.70 | -0.24 |
| | 13 | 200 | 100.80 | 21.13 | | |
| Verbal | 15 | 200 | 50.35 | 11.90 | 0.85 | 0.50 |
| | 12 | 200 | 49.50 | 12.03 | | |

Table 5.20 (continued)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Performance | 15 | 200 | 49.75 | 10.75 | 10.70 | 0.07 | 0.04 |
| | 12 | 200 | 49.68 | 11.39 | | | |
| Full Scale | 15 | 200 | 100.10 | 21.21 | | | |
| | 12 | 200 | 99.18 | 20.57 | 0.92 | 0.31 | |

* Significant at the 0.05 level
** significant at the 0.01 level

Verbal score is the sum of scaled scores on 5 tests: Performance score is the sum of scaled scores on 5 test
Full Scale score is the sum of scaled scores on 10 tests.

Table 5.21
Proportions of Boys and Girls Together Passing Each Item within Each of the Four Age-Groups and the Order of Difficulty for Each Item of Each Subtest for the Four Age Groups Combined for the Present Study

1 INFORMATION

| WISC-R Order Items | Year 12 | Year 13 | Year 14 | Year 15 | Average Passing % | Order of Difficulty |
|---|---|---|---|---|---|---|
| 1 | 1.00 | 1.00 | 1.00 | 1.00 | 100 | 1 |
| 2 | 1.00 | 1.00 | 1.00 | 1.00 | 100 | 2 |
| 3 | 1.00 | 1.00 | 1.00 | 1.00 | 100 | 3 |
| 4 | 1.00 | 1.00 | 1.00 | 1.00 | 100 | 4 |
| 5 | 1.00 | 1.00 | 1.00 | 1.00 | 100 | 5 |
| 6 | .98 | .99 | 1.00 | 1.00 | 99.25 | 7 |
| 7 | 1.00 | 1.00 | 1.00 | 1.00 | 100 | 6 |
| 8 | .96 | .97 | 1.00 | .99 | 98 | 10 |
| 9 | .97 | .98 | 1.00 | 1.00 | 98.75 | 8 |
| 10 | .86 | .95 | .96 | .95 | 93 | 12 |
| 11 | .96 | .98 | 1.00 | 1.00 | 98.50 | 9 |
| 12 | .17 | .14 | .19 | .23 | 18.25 | 27 |
| 13 | .95 | .97 | .99 | .99 | 97.50 | 11 |
| 14 | .68 | .71 | .75 | .79 | 73.25 | 16 |
| 15 | .65 | .78 | .83 | .89 | 78.75 | 15 |
| 16 | .05 | .11 | .14 | .28 | 14.50 | 28 |
| 17 | .81 | .89 | .93 | .94 | 89.25 | 14 |
| 18 | .65 | .69 | .69 | .78 | 70.25 | 19 |
| 19 | .85 | .90 | .93 | .99 | 91.75 | 13 |
| 20 | .41 | .56 | .48 | .42 | 46.75 | 21 |
| 21 | .36 | .48 | .38 | .49 | 42.75 | 23 |
| 22 | .58 | .67 | .71 | .89 | 71.25 | 18 |

Table 5.21 (continued)

| | | | | | | |
|---|---|---|---|---|---|---|
| 23 | .09 | .18 | .19 | .28 | 18.50 | 26 |
| 24 | .20 | .28 | .26 | .49 | 30.75 | 25 |
| 25 | .23 | .31 | .47 | .69 | 42.50 | 24 |
| 26 | .63 | .71 | .79 | .75 | 72 | 17 |
| 27 | .30 | .32 | .46 | .73 | 45.25 | 22 |
| 28 | .41 | .42 | .54 | .68 | 51.25 | 20 |
| 29 | .03 | .04 | .10 | .09 | 6.50 | 29 |
| 30 | .00 | .04 | .07 | .06 | 4.25 | 30 |

2 PICTURE COMPLETION

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | 1.00 | 1.00 | 1.00 | 1.00 | 100 | 1 |
| 2 | .96 | 1.00 | .99 | 1.00 | 98.75 | 3 |
| 3 | 1.00 | 1.00 | 1.00 | 1.00 | 100 | 2 |
| 4 | .94 | .99 | 1.00 | 1.00 | 98.25 | 4 |
| 5 | .93 | .98 | .97 | .99 | 96.75 | 5 |
| 6 | .81 | .83 | .89 | .91 | 86 | 9 |
| 7 | .93 | .95 | .99 | 1.00 | 96.75 | 6 |
| 8 | .65 | .70 | .63 | .83 | 70.25 | 16 |
| 9 | .64 | .69 | .71 | .84 | 72 | 14 |
| 10 | .58 | .64 | .69 | .72 | 65.75 | 17 |
| 11 | .89 | .92 | .97 | .98 | 94 | 7 |
| 12 | .74 | .78 | .81 | .79 | 78 | 12 |
| 13 | .73 | .79 | .84 | .86 | 80.50 | 11 |
| 14 | .78 | .83 | .86 | .89 | 84 | 10 |
| 15 | .49 | .56 | .60 | .72 | 59.25 | 18 |
| 16 | .87 | .82 | .88 | .89 | 86.50 | 8 |
| 17 | .70 | .78 | .83 | .79 | 77.50 | 13 |
| 18 | .68 | .69 | .75 | .70 | 70.50 | 15 |
| 19 | .34 | .41 | .47 | .44 | 41.50 | 18 |
| 20 | .29 | .39 | .42 | .41 | 37.75 | 20 |
| 21 | .15 | .14 | .18 | .17 | 16 | 22 |

Table 5.21 (continued)

| | | | | | | |
|---|---|---|---|---|---|---|
| 23 | .09 | .18 | .19 | .28 | 18.50 | 26 |
| 24 | .20 | .28 | .26 | .49 | 30.75 | 25 |
| 25 | .23 | .31 | .47 | .69 | 42.50 | 24 |
| 26 | .63 | .71 | .79 | .75 | 72 | 17 |
| 27 | .30 | .32 | .46 | .73 | 45.25 | 22 |
| 28 | .41 | .42 | .54 | .68 | 51.25 | 20 |
| 29 | .03 | .04 | .10 | .09 | 6.50 | 29 |
| 30 | .00 | .04 | .07 | .06 | 4.25 | 30 |

2 PICTURE COMPLETION

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | 1.00 | 1.00 | 1.00 | 1.00 | 100 | 1 |
| 2 | .96 | 1.00 | .99 | 1.00 | 98.75 | 3 |
| 3 | 1.00 | 1.00 | 1.00 | 1.00 | 100 | 2 |
| 4 | .94 | .99 | 1.00 | 1.00 | 98.25 | 4 |
| 5 | .93 | .98 | .97 | .99 | 96.75 | 5 |
| 6 | .81 | .83 | .89 | .91 | 86 | 9 |
| 7 | .93 | .95 | .99 | 1.00 | 96.75 | 6 |
| 8 | .65 | .70 | .63 | .83 | 70.25 | 16 |
| 9 | .64 | .69 | .71 | .84 | 72 | 14 |
| 10 | .58 | .64 | .69 | .72 | 65.75 | 17 |
| 11 | .89 | .92 | .97 | .98 | 94 | 7 |
| 12 | .74 | .78 | .81 | .79 | 78 | 12 |
| 13 | .73 | .79 | .84 | .86 | 80.50 | 11 |
| 14 | .78 | .83 | .86 | .89 | 84 | 10 |
| 15 | .49 | .56 | .60 | .72 | 59.25 | 18 |
| 16 | .87 | .82 | .88 | .89 | 86.50 | 8 |
| 17 | .70 | .78 | .83 | .79 | 77.50 | 13 |
| 18 | .68 | .69 | .75 | .70 | 70.50 | 15 |
| 19 | .34 | .41 | .47 | .44 | 41.50 | 28 |
| 20 | .29 | .39 | .42 | .41 | 37.75 | 20 |
| 21 | .15 | .14 | .18 | .17 | 16 | 22 |

Table 5.21 (continued)

| | | | | | | |
|---|---|---|---|---|---|---|
| 22 | .09 | .12 | .14 | .14 | 12 | 24 |
| 23 | .07 | .13 | .17 | .15 | 13 | 23 |
| 24 | .20 | .23 | .29 | .28 | 25 | 21 |
| 25 | .11 | .09 | .13 | .14 | 11.75 | 25 |
| 26 | .02 | .04 | .01 | .05 | 3 | 26 |
| | | | 3 SIMILARITIES | | | |
| 1 | 1.00 | 1.00 | 1.00 | 1.00 | 100 | 1 |
| 2 | .99 | 1.00 | 1.00 | 1.00 | 99.75 | 2 |
| 3 | .98 | .97 | 1.00 | .99 | 98.50 | 3 |
| 4 | .95 | .97 | .99 | .99 | 97.50 | 5 |
| 5 | .97 | .96 | 1.00 | .98 | 97.75 | 4 |
| 6 | .80 | .84 | .79 | .91 | 83.50 | 7 |
| 7 | .73 | .78 | .80 | .77 | 77 | 8 |
| 8 | .35 | .38 | .36 | .39 | 37 | 11 |
| 9 | .38 | .42 | .45 | .48 | 43.25 | 9 |
| 10 | .79 | .84 | .90 | .87 | 85 | 6 |
| 11 | .20 | .26 | .30 | .36 | 28 | 13 |
| 12 | .39 | .41 | .42 | .43 | 41.25 | 10 |
| 13 | .27 | .30 | .34 | .32 | 30.75 | 12 |
| 14 | .06 | .09 | .07 | .06 | 7 | 16 |
| 15 | .11 | .09 | .13 | .15 | 12 | 15 |
| 16 | .02 | .08 | .06 | .09 | 6.25 | 17 |
| 17 | .10 | .12 | .15 | .16 | 13.25 | 14 |
| | | | 4 PICTURE ARRANGEMENT | | | |
| 1 | .95 | .93 | .96 | .99 | 95.75 | 1 |
| 2 | .83 | .88 | .93 | .94 | 89.50 | 4 |
| 3 | .86 | .89 | .94 | .95 | 91 | 3 |
| 4 | .65 | .72 | .78 | .83 | 74.50 | 5 |

Table 5.21 (continued)

| | | | | | | |
|---|---|---|---|---|---|---|
| 5 | .87 | .91 | .96 | .98 | 93 | 2 |
| 6 | .72 | .70 | .74 | .73 | 72.25 | 6 |
| 7 | .68 | .71 | .73 | .75 | 72.25 | 7 |
| 8 | .61 | .67 | .64 | .68 | 65 | 8 |
| 9 | .50 | .52 | .61 | .58 | 52.75 | 9 |
| 10 | .28 | .26 | .35 | .33 | 30.50 | 11 |
| 11 | .30 | .32 | .37 | .37 | 34 | 10 |
| 12 | .07 | .09 | .11 | .13 | 10 | 12 |

5 ARITHMETIC

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | 1.00 | 1.00 | 1.00 | 1.00 | 100 | 1 |
| 2 | 1.00 | 1.00 | 1.00 | 1.00 | 100 | 2 |
| 3 | 1.00 | 1.00 | 1.00 | 1.00 | 100 | 3 |
| 4 | 1.00 | 1.00 | 1.00 | 1.00 | 100 | 4 |
| 5 | 1.00 | 1.00 | 1.00 | 1.00 | 100 | 5 |
| 6 | .98 | 1.00 | 1.00 | 1.00 | 99.50 | 6 |
| 7 | .99 | .99 | 1.00 | 1.00 | 99.50 | 7 |
| 8 | .96 | .98 | .99 | 1.00 | 98.25 | 8 |
| 9 | .94 | .96 | .98 | .99 | 96.75 | 9 |
| 10 | .86 | .83 | .91 | .94 | 88.50 | 10 |
| 11 | .71 | .73 | .75 | .78 | 74.25 | 12 |
| 12 | .70 | .74 | .78 | .81 | 75.75 | 11 |
| 13 | .38 | .41 | .43 | .48 | 42.50 | 14 |
| 14 | .48 | .56 | .61 | .72 | 59.25 | 13 |
| 15 | .25 | .31 | .32 | .39 | 31.75 | 15 |
| 16 | .18 | .20 | .22 | .25 | 21.25 | 16 |
| 17 | .08 | .12 | .16 | .19 | 13.75 | 17 |
| 18 | .01 | .05 | .08 | .11 | 6.25 | 18 |

Table 5.21 (continued)

6 BLOCK DESIGN

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | 1.00 | 1.00 | 1.00 | 1.00 | 100 | 1 |
| 2 | 1.00 | 1.00 | 1.00 | 1.00 | 100 | 2 |
| 3 | .96 | .95 | .97 | .99 | 96.75 | 3 |
| 4 | .85 | .89 | .93 | .95 | 90.50 | 4 |
| 5 | .78 | .81 | .83 | .89 | 82.75 | 5 |
| 6 | .71 | .76 | .80 | .86 | 78.25 | 6 |
| 7 | .67 | .73 | .74 | .79 | 72.25 | 7 |
| 8 | .63 | .65 | .69 | .72 | 67.25 | 8 |
| 9 | .17 | .20 | .24 | .25 | 21.50 | 9 |
| 10 | .03 | .06 | .09 | .10 | 6.75 | 11 |
| 11 | .05 | .08 | .11 | .13 | 9.25 | 10 |

[1] Pilot Test
[2] Order of Difficulty

7 VOCABULARY

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | .95 | .93 | .98 | .99 | 96 | 3 |
| 2 | .97 | .98 | 1.00 | 1.00 | 98.75 | 1 |
| 3 | .96 | .98 | 1.00 | 1.00 | 98.50 | 2 |
| 4 | .91 | .93 | .97 | .97 | 94.50 | 4 |
| 5 | .87 | .91 | .95 | .98 | 92.75 | 6 |
| 6 | .84 | .89 | .92 | .94 | 90.25 | 7 |
| 7 | .83 | .88 | .91 | .95 | 89.25 | 8 |
| 8 | .96 | .93 | .96 | .98 | 94.25 | 5 |
| 9 | .88 | .84 | .86 | .90 | 87 | 9 |
| 10 | .76 | .79 | .83 | .88 | 81.50 | 13 |
| 11 | .81 | .85 | .80 | .90 | 8[illegible] | [illegible] |
| 12 | .80 | .83 | .87 | .90 | 85.50 | 10 |
| 13 | .79 | .82 | .86 | .89 | 84.50 | 11 |
| 14 | .76 | .79 | .82 | .86 | 80.25 | 14 |
| 15 | .71 | .75 | .78 | .81 | 77.50 | 15 |

Table 5.21 (continued)

| | | | | | | |
|---|---|---|---|---|---|---|
| 16 | .70 | .74 | .78 | .80 | 75.50 | 16 |
| 17 | .67 | .72 | .79 | .83 | 75.25 | 18 |
| 18 | .66 | .69 | .73 | .78 | 71.50 | 19 |
| 19 | .70 | .73 | .78 | .81 | 75.50 | 17 |
| 20 | .61 | .165 | .69 | .75 | 67.75 | 20 |
| 21 | .58 | .63 | .68 | .72 | 62.75 | 22 |
| 22 | .59 | .62 | .66 | .70 | 64.25 | 21 |
| 23 | .51 | .57 | .60 | .61 | 57.25 | 23 |
| 24 | .50 | .54 | .56 | .59 | 54.75 | 24 |
| 25 | .42 | .49 | .53 | .57 | 50.25 | 25 |
| 26 | .39 | .42 | .47 | .51 | 44.75 | 26 |
| 27 | .35 | .40 | .43 | .47 | 41.25 | 27 |
| 28 | .36 | .37 | .41 | .46 | 40 | 29 |
| 29 | .37 | .40 | .42 | .45 | 41 | 28 |
| 30 | .32 | .38 | .42 | .48 | .40 | 30 |
| 31 | .25 | .29 | .31 | .35 | 40 | 31 |
| 32 | .20 | .24 | .29 | .32 | 27 | 32 |

8 OBJECT ASSEMBLY

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | .86 | .92 | .99 | 1.00 | 94.25 | 1 |
| 2 | .69 | .76 | .84 | .91 | 80 | 2 |
| 3 | .30 | .35 | .44 | .52 | 40.25 | 3 |
| 4 | .23 | .29 | .34 | .39 | 31.25 | 4 |

PAGE
NUMBERING
AS
ORIGINAL

Table 5.21 (continued)

9 COMPREHENSION

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | .96 | 1.00 | 1.00 | 1.00 | .99 | 1 |
| 2 | .78 | .84 | .89 | .92 | 95.25 | 2 |
| 3 | .80 | .85 | .89 | .94 | 87 | 3 |
| 4 | .71 | .79 | .84 | .88 | 80.50 | 4 |
| 5 | .63 | .67 | .72 | .76 | 69.50 | 7 |
| 6 | .48 | .52 | .57 | .60 | 54.25 | 11 |
| 7 | .68 | .71 | .76 | .80 | 73.75 | 5 |
| 8 | .37 | .39 | .44 | .49 | 42 | 14 |
| 9 | .43 | .49 | .54 | .59 | 51.25 | 12 |
| 10 | .02 | .05 | .09 | .11 | 6.75 | 17 |
| 11 | .25 | .29 | .37 | .42 | 33.25 | 15 |
| 12 | .52 | .58 | .61 | .66 | 59.25 | 10 |
| 13 | .39 | .41 | .44 | .47 | 42.75 | 13 |
| 14 | .58 | .64 | .70 | .76 | 67 | 8 |
| 15 | .65 | .69 | .72 | .77 | 70.75 | 6 |
| 16 | .60 | .64 | .69 | .71 | 66 | 9 |
| 17 | .09 | .11 | .14 | .19 | 13.25 | 16 |

## Socioeconomic Status and IQ.

In the past the effects of social class have often been underestimated, according to Jensen (1979). Today on the other hand, it is sometimes implied that differences in measured intelligence reflect only differences in social class or in early environment. Swift (1967), for instance, in an excellent attempt to study the social and environmental factors affecting success in the 'eleven-plus' selection procedure, and to refine the rather crude indices of 'social class' that have generally been used in experimental studies, writes of "ability" in inverted commas, as if to imply that individual differences ascribable to factors other than social class do not exist. Thus, it was important to discover the relationship between IQ and social class in this study in Iraq.

It should be noted that one of the aims of this research was to examine the relationship between fathers' occupations and the measured intelligence of their children. (See Chapter One, 1.1). A more sophisticated division of occupations was employed in the final main study than in the pilot sample. This division was based mainly on the income of fathers (National Census, 1987). The classification of fathers' occupations for the final study, in accordance with the National Census, 1987, was as follows:

1- Unskilled workers, farmers, cattle- raisers.

2- Skilled workers.

3- Professional workers, builders who have offices, electricians, shop-owners, car-owners.

4- Lease-holders, high officials (i.e. graduates).

5- Self-employed professionals (e.g. doctors, dentists, university teachers).

6- Rentiers, merchants, industrialists.

The data classification of the occupational status of households for the present study were done by the Educational and Psychological Research Centre, University of Baghdad. The categorizations of the occupations of the Pilot Sample were based on the 1977 Census. When the researcher finished the pilot testing stage a new division of occupations appeared which was based on the 1987 Census published in November 1987. The main differences between the old and the new divisions are that in the latter list each category is made clearer, with more information helping classification.

Table 5.22 shows the fathers' occupations in each of the above categories and the corresponding means and Standard Deviations for Full Scale IQs of the children within each category for all four age groups, 12, 13, 14 and 15 years combined.

Table 5.22

Numbers, Means and Standard Deviations for the Whole Sample of Full Scale IQs in each of the Occupation Groups of the Fathers. N=800.

| Group | Number | Mean | Standard Deviation |
|---|---|---|---|
| 1 | 263 | 94.57 | 23.20 |
| 2 | 192 | 98.13 | 21.64 |
| 3 | 137 | 103.98 | 18.42 |
| 4 | 88 | 108.69 | 22.36 |
| 5 | 76 | 110.47 | 19.72 |
| 6 | 44 | 113.95 | 19.28 |
| Total | 800 | 100.28 | 21.16 |

Table 5.23 gives the results obtained from an analysis of variance, indicating that there was a significant difference between the groups' IQ means ($p < .001$). Table 5.22 shows the means and Standard Deviations of the six groups' Full

Scale mean IQs. These means indicate that the higher the social class, the higher the mean IQ score.

Table 5.23

Analysis of Variance of Full Scale IQ within Father Occupations Groups for the Whole Sample (All Age Groups).

| Source | D.F. | Sum of Squares | Mean Prob. | F Ratio | F Prob. |
|---|---|---|---|---|---|
| Between Groups | 5 | 8738.60 | 1747.72 | 25.80 | 0.001 |
| Within Groups | 791 | 53577.52 | 62.78 | | |
| Total | 796 | 62316.12 | | | |

## 5.2 **Family Size and IQ.**

By the term 'family size' the author means the number of children in the family. The data on family size were obtained directly from the children during the testing period. Each subject was asked to give the number of his/her brothers and sisters. It was assumed that the students tested in this sample had no reason for giving inaccurate data and therefore there was no real need for verification of the accuracy of their statements from teachers or parents.

Table 5.24 presents the number of students within each category of family size for each age group tested in the present study. Table 5.25 shows the means and differences between means for the Full Scale IQs.

It was thought that these two groupings of 'family size' were meaningful in order to find and test the differences in mean IQs. between categories.

Results (See Table 5.25) of family size and IQ show that this factor may affect the intelligence of the children, in that the fewer the children the family have, the higher the IQ of the children.

Table 5.24

Numbers of Children Tested within Each Category of Family Size for Each Age Group, Main Study.

| Family | Size | N=12yrs | N=13yrs. | N=14yrs | N=15yrs. |
|---|---|---|---|---|---|
| 1- | 1 to 3 | 40 | 38 | 22 | 36 |
| 2- | 4 to 6 | 90 | 102 | 96 | 92 |
| 3- | 7 to 9 | 50 | 40 | 56 | 48 |
| 4- | 10 and up | 20 | 20 | 26 | 24 |
| Total | | 200 | 200 | 200 | 200 |

## 5.3 Urban-Rural Residence and IQ.

Communities with 2,500 or more inhabitants were defined as urban, and smaller communities as rural. (See Chapter Four). Table 5.26 presents the means, differences between means, Standard Deviations and the t values for the samples of students from urban areas and those from rural areas for each age-level. It was intended to discover differences between performance of children in urban residence and these in rural residence, bearing in mind the need for such data in Iraq. It will be of considerable value to provide some knowledge of discrepancies in such student performance for the use of specialists and educators.

From Table 5.26 it is seen that for all age-groups, urban students are on average significantly more intelligent than those from rural areas.

Table 5.25

Mean IQs and Standard Deviations for the Two Types of Categories of Full Scale Family Size for Each Age group of 12, 13, 14 and 15 year-olds and t-Values

| Age | Family Size | Subjects Number | Full Scale IQ | | | t-Value |
|---|---|---|---|---|---|---|
| | | | M. | Difference between Means | S.D. | |
| 12 | 1-3 children | 40 | 105.37 | 15.45 | 15.04 | 15.04** |
| | 4 or more | 160 | 89.92 | | 11.18 | |
| 13 | 1-3 children | 38 | 102.75 | 12.09 | 15.76 | 3.65** |
| | 4 or more | 162 | 90.66 | | 10.83 | |
| 14 | 1-3 children | 22 | 105.83 | 13.59 | 15.12 | 3.73** |
| | 4 or more | 178 | 92.24 | | 13.31 | |
| 15 | 1-3 children | 36 | 100.91 | 2.15 | 15.16 | 0.56 |
| | 4 or more | 164 | 98.76 | | 16.39 | |

* Significant at 5%

Table 5.26

Mean IQs and S.Ds for Urban–Rural Residence for Students Aged 12, 13, 14 and 15 Years and t–Values.

| Region | Age | Number of Subjects | Mean | Mean Diff. | S.D. | t value |
|---|---|---|---|---|---|---|
| Urban | 12 | 141 | 103.97 | 11.61 | 15.26 | 4.00** |
| Rural | | 59 | 92.36 | 12.41 | | |
| Urban | 13 | 148 | 102.29 | 13.86 | 14.62 | 4.80** |
| Rural | | 52 | 88.43 | 11.94 | | |
| Urban | 14 | 146 | 105.32 | 11.20 | 15.83 | 3.69** |
| Rural | | 54 | 94.12 | 12.60 | | |
| Urban | 15 | 153 | 105.19 | 10.74 | 15.64 | 3.35** |
| Rural | | 47 | 94.45 | 12.87 | | |

** Significant at 1 %.

* Significant at 5 %.

## 5.4 Number of Years of Schooling of Fathers and Mothers and IQ.

The variable 'number of years of schooling' of father and mother was categorised for each student as follows:

1= No formal schooling

2= Some years in primary school.

3= Finished primary school and completed some years in intermediate school.

4= Finished intermediate school and completed some years in secondary school.

5= Finished secondary school.

6= Finished university or equivalent institute or took a higher degree.

This information was obtained from the students by asking them, and from the information provided to the school headmaster by parents every year. The data obtained about number of years of schooling of father and mother in the present study were classified by the Research Centre in the Ministry of Education.

As can be seen, these category divisions are the same as those used in the pilot sample; the means of the Full Scale IQs are given in Table 5.27 for each category. Table 5.28 presents the analyses of variance of Full Scale IQ between the six groups, specifying the number of years the parents attended school, and the F ratio and F probability.

The one--way analysis of variance shown in Table 5.28 indicates that there were significant differences in mean IQs at $p < 0.01$ level between groups, associated with the educational level of parents.

Table 5.27

Means and Standard Deviations of the Full Scale IQs in Each of Years of Schooling Groups, in the Total Sample N=800.

| Group | Mean | S.D. |
|---|---|---|
| 1 | 94.18 | 24.31 |
| 2 | 97.96 | 22.07 |
| 3 | 104.73 | 18.15 |
| 4 | 105.34 | 19.79 |
| 5 | 106.45 | 20.76 |
| 6 | 111.67 | 20.42 |
| Total | 100.28 | 21.16 |

Table 5.28

Analysis of Variance of Full Scale IQ between Parents' Years of Schooling Groups

| Source | D.F | Sum of Squares | Mean | F Ratio | F Prob. |
|---|---|---|---|---|---|
| Between Groups | 5 | 3760.34 | 752.06 | 5.59 | 0.01 |
| Within Groups | 791 | 50773.05 | | 67.25 | |
| Total | 796 | 54533.39 | | | |

The students' IQ mean scores were found to be dependent on the education of their parents in that students coming from homes with poorly educated parents tended to be less intelligent than those from homes with more educated parents. From an analysis of variance these group differences in IQ were shown to be highly significant.

## 5.5 Geographical Region and IQ.

For this study Iraq was divided into the three geographical regions specified under the 1987 Census Report: North, South and Middle. Table 5.29 presents the mean Full Scale IQs and Standard Deviations for the children living in the three geographical regions for the four age groups. This table also contains the results of the t--analyses on the differences between the IQ means from within the different regions. These results show that students living in the Middle region of Iraq obtained significantly higher IQ scores than those from the Northern or Southern regions.

As far as the author knows, there is no study that has been carried out in Iraq to discover differences between students' performances on IQ tests administered within the three Iraqi regions. The superiority of students living in the middle region is clearly manifested.

Table 5.29

Number of Children, Boys and Girls, Tested with WISC-R in the Three Geographical Regions: North, Middle and South, Mean, Standard Deviation, Differences Between Means and t-Values for the Average of all Age-Groups for the Full Scale IQ.

| Region | Number | Full Scale Mean | Full Scale S.D. | Difference between Means | t-Value |
|---|---|---|---|---|---|
| North | 140 | 96.81 | 13.11 | -6.97 | -4.00** |
| Middle | 500 | 103.78 | 14.62 | | |
| North | 140 | 96.81 | 13.11 | -2.17 | -1.02 |
| South | 160 | 98.98 | 12.94 | | |
| Middle | 500 | 103.78 | 14.62 | 4.80 | 2.92** |
| South | 160 | 98.98 | 12.94 | | |

* Significant at 0.05 level
** Significant at 0.01 level

# CHAPTER SIX

## Discussion, Summary and Conclusion

## Chapter Six

CHAPTER SIX

# Discussion, Summary and Conclusion

## 6.0 Standardization of the Scale

There has been a growing awareness in Iraq of the need to assess children's intelligence and abilities, to aid evaluation of the various obstacles to education like learning difficulties and behavioural and emotional problems, as well as to identify and assess gifted children and those with other special needs.

Some educational and psychological institutions in the past and even at the present time have attempted to meet the increasing demand for an assessment of children's intelligence by applying standardized tests such as the WISC, WISC-R and Scales of the Progressive Matrices, to other cultures and in other countries. To place a child in a special education programme the school depends on the teacher's judgement or the children's subject marks (Maths, Reading ...etc.). IQ scores are often derived from foreign norms or by educators and psychologists intuitively estimating the level of children's intellectual functioning. This practice can have and does have major consequences for the children's lives and those of their parents. Psychologists and educators have therefore begun to recognise the need for Iraqi norms.

Clarke and Clarke (1975) stated the advantages of assessment as functioning: "1) To describe the individual as he is at a particular point in time, upon intellectual, social, emotional or other variables with reference to a normative or contrast population; 2) To predict the individual's probable status at later points in time; 3) To provide a behavioural profile of assets and deficits as a starting point for remedial programmes; 4) To provide an objective means of checking progress of an individual or a group"

Wechsler (1974) prepared a cumulative frequency distribution of raw scores for each age group, normalized the distribution, and computed the appropriate scaled score for each raw score (Wechsler, 1974).

Anastasi (1976) preferred to avoid normalized standard scores, which is not the stance adopted by the present author. Anastasi states:

> Whenever feasible, it is generally more desirable to obtain a normal distribution of raw scores by proper adjustment of the difficulty level of test items rather than by subsequently normalizing a markedly abnormal distribution.

From Table 5.7 we can see that this approach is supported when the scores are adjusted to follow an approximate normal distribution. Moreover, as we can see from Tables 5.2 and 5.3, the means and Standard Deviations for each subtest and for the sums of scaled scores of Verbal, Performance and Full Scale of the present study are similar to those of the British and American WISC-R studies, with small discrepancies.

Tables 5.2 and 5.3 give scores for each age- group (boys and girls together), of means and Standard Deviations for the present study and for both the British and American WISC-R. From the results in Table 5.2 we can conclude that the present sample has a similar mean and Standard Deviation of scaled scores resulting from the same scaling procedures. Schwarz and Krug (1972) made four recommendations which can be applied to the present work:

> In summary, four major points have been made about specific test adaptations. The first was that changes of content do not cheapen or degrade a standard test,and do not imply that less able applicants will be selected. Often, adapting a test is the only way of identifying people as able as those selected in

> other countries with the traditional versions. The second was that although adaptation is normally required in a developing country, these changes can be verified before the project proceeds. The third is that elaborate background research is seldom necessary to produce effective tests and should not be programmed when there is an immediate need for operational testing procedures. And fourth was that training failures can as easily result from deficiencies in the course as deficiencies in testing, and that a project which would cure only the ills on the tests may therefore not yield the improvements desired.

It must be noted that the present study is in some respects more elaborate than some other researches mentioned in the review chapters, because of the great care and time taken in the selection of controlling variables and in the statistical treatment put into this research. Therefore, its strength lies in the strict adherence to the most appropriate procedures governing major test construction developed for the WISC-R intelligence test. Where appropriate the present author has modified test items and constructed his own subtest, e.g. Vocabulary.

## 6.1 Reliability

'Reliability' refers to the consistency of scores obtained by the same person when retested with the same test on different occasions or with different sets of equivalent items. Or in different terms, we can say that reliability underlies the computation of the error of measurement of a single score, whereby we can predict the range of fluctuation likely to occur in a single individual's score as a result of irrelevant chance factors.

The concept of 'test reliability' has been used to cover several aspects of score consistency. In its broadest sense, 'test reliability' indicates the extent to which individual differences in test scores are attributable to 'true' differences in the characteristic under consideration and the extent to which they are attributable to chance errors.

Table 5.8 shows the reliability coefficients for the Split-half and the Kuder-Richardson Formula 20 for each of the ten subtests at the four age- levels and for Verbal, Performance and Full Scale IQs. As we can see from that table, the subtests of Information, Vocabulary and Block Design have the highest reliabilities, ranging from 0.72 to 0.86 for the Information subtest, with average reliability for the four age-levels of 0.81 for the Split-half and 0.74 for Kuder-Richardson Formula 20; and from 0.82 to 0.94 for the Vocabulary subtest with average reliability of 0.87 for the Split-half, and 0.90 for the Kuder-Richardson Formula 20; for the Block Design subtest a reliability from 0.79 to 0.95 with average reliability of 0.81 for the Split-half and 0.92 for the Kuder-Richardson Formula 20 for the four age-levels. Object Assembly was the least reliable subtest, ranging from 0.49 to 0.73 with average reliability, for the four age- levels, and of 0.60 for the Split-half. This low reliability was due to the fact that this subtest contained four items only. These results were in agreement with Wechsler's British findings (1974), as we can see from Table 5.11.

Table 5.8 shows the reliability coefficients of Verbal, Performance and Full Scale IQ in the present sample. From these results we can see that the highest reliability coefficients are for the Full Scale test (average value of 0.92), ranging from 0.89 to 0.96; and the second highest reliability coefficients are with the Verbal Scale test, ranging from 0.90 to 0.94, with average value of 0.92. The Performance test, on average, was the least reliable (average value of 0.88), ranging from 0.82 to 0.95. Table 5.11 also shows the reliability coefficients both for the British WISC-R and for the present study. From Table 5.11 it can be judged that the present reliability results are satisfactory when compared with the corresponding reliabilities obtained from the British standardization of WISC-R.

In general, the Verbal tests show greater reliability than the Performance tests. The small number of items in some of the Performance tests may account for this, as lower reliability is a common feature of most Performance tests, especially for the Object Assembly. (See Table 5.11). Generally speaking, the reliabilities of the various subscales within the present sample are comparable with those of the British WISC-R, since the Verbal part (Information, Comprehension, Vocabulary, Arithmetic and Similarities subtests) shows that the changes made on the various test items were successful and suitable for the Iraqi children.

Full Scale tests show a higher value for reliability than the Verbal and Performance test scales. These results are natural and expected, since reliability normally increases with the length of the test. Comparing the reliabilities in Table 5.8 with the studies on WISC and WISC-R reviewed in Chapter Two, it can be seen that some of the present results are higher and more reliable than those in other studies (Dastoor and Emovon, 1972). This is due perhaps to the fact that our sample was larger and more widespread than many other studies. Moreover, the WISC-R test is more reliable than the WISC on which the previous studies were based. (Sattler, 1982).

Table 5.9 presents the standard error of measurement based on the two techniques of finding reliabilities (Split-half and Kuder-Richardson Formula 20) of scaled scores on each of the WISC-R subtests (except Coding B) and the three Scale IQs. The standard error of measurement (SEm) is the function of the reliability coefficient and an indication of the confidence one can have in making judgements about a child's true performance on a particular test. Table 5.10 presents the standard errors of measurement of the present study and the British WISC-R findings and shows that they are comparable. This result of the present study is also in agreement with the American WISC-R standardization sample.

## 6.2 Validity.

Wechsler established the WISC-R test's external validity by calculating the degree of correlation it had with other recognised intelligence tests such as the Wechsler Preschool and Primary Scale of Intelligence (WPPSI), the Wechsler Adult

Intelligence Scale (WAIS), and the Stanford-Binet Intelligence Scale (Form L-M, 1972 norms). (Wechsler, 1974).

It was difficult in this study to use any other test as an external criterion to test the validity of the present WISC-R, because the new test (Vocabulary) needed adaptation and standardization which took a great deal of time and effort. The internal validity of a test requires that all the subtests should correlate highly with the total score obtained on the whole scale, while they should correlate moderately with each other. The reason for this is that when the subtests correlate highly with the whole test they are measuring the same thing. The moderate validity correlations show that they are measuring various aspects of the same thing.

Table 5.12 shows the inter--subtest correlations for ages 12, 13, 14 and 15 years obtained in the present research, with the correlations of each subtest with the Verbal, Performance and Full Scale scores, and the correlations of the Verbal and Performance scores with the Full Scale Scores. From Table 5.12 it can be concluded that inter--subtest correlations were of the same order of magnitude as those given in the British WISC-R Manual for the standardization samples at ages 12 1/2, 13 1/2, 14 1/2 and 15 1/2 (ranging from 0.11 to 0.90 at the average age of 12 1/2, from 0.20 to 0.90 at the average age of 13 1/2, from 0.21 to 0.90 at the average age of 14 1/2 and from 0.19 to 0.92 at the average age of 15 1/2).

The lowest intercorrelations were between Coding B and the other Verbal subtests. The highest intercorrelations in both British WISC-R and the present study were found between the Verbal subtests, with the exception of Arithmetic. But the correlations of the Arithmetic test with the rest of the Verbal tests increased with age level from age 12 1/2 to 15 1/2 years old. This is probably due to the fact that the Arithmetic subtest had items at the more difficult level in which verbal comprehension plays a considerable part.

Performance subtests in general have lower intercorrelations than those for the Verbal subtests, indicating that they are measuring different aspects of intelligence. The Coding B test had negligible correlations with the other Performance subtests; the lowest of these correlations was at age 12 years; the highest correlations were with the Object Assembly test. There was also an increase in the inter-correlation at age 13 between Coding B and the other Performance subtests except with Picture Arrangement. These findings are in agreement with the British and American WISC-R.

Most of the Verbal subtests correlated more highly with one another than with the Performance subtests. The overall (IQ) Performance scale was less homogeneous than the overall (IQ) Verbal scale; and subtests on the Verbal and Performance scales correlated more highly within the scales than between the two sets of tests overall, indicating that they were measuring different aspects of the intelligence, as is to be expected.

The Verbal IQ scores correlate more highly with each Verbal subtest than the Performance IQ score with each Performance subtest scaled score. It can be concluded that the Performance subtests were somewhat less homogeneous than the Verbal scale.

The correlations between total Verbal and Performance IQ scores ranged from 0.64 to 0.96 within the different age--groups, with an average value of 0.86. Thus, the two parts of the Full Scale IQ scale had much in common, but the correlations between them were not consistently high. Retention of the two separate IQ scores measuring both Verbal IQ and Performance IQ is therefore justified.

## 6.3 Gender Differences.

Seashore (1950) in his standardization study made reference to gender differences based on means and Standard Deviations of the three WISC scales. The boys generally did better than the girls, the boys' superiority in that study increasing with the increase in age (10 to 15 years).

The results of the present study are given in Tables 5.14 and 5.15 . The figures in these tables show that the boys gained slightly better results than the girls in all the subtests and in all age--groups--12, 13, 14 and 15 years, except for the subtest of Coding B, where we can see that the girls excelled the boys in all four age--groups of our study. The same thing happened with the three IQ scales for the four age- levels 12, 13, 14 and 15 years (See Table 5.15); the boys' scores were found to be higher than the girls'. When the 't' tests were applied in order to assess the significance of the differences between the boys and girls in the IQ scales, it was found that significant differences occurred with the Verbal and Full Scale for the 12, 13, 14 and 15 year--old samples, and for both the Verbal scale for the 13 year--old group and the Performance scale for the 14 year--old group. All other differences between girls and boys were not significant.

As mentioned above, Table 5.14 shows that the boys excelled the girls, particularly within the older age- groups of this sample, that is, for the 14 and 15 year-old groups. The girls also showed greater variability than the boys in most of the subtests. This may be due to the cultural differences in Iraq in the ways boys and girls are raised and educated. No other obvious reason can be found, and in the future researchers will have to take this into consideration.

As we can see from Table 5.16, the differences in means between the boys and the girls were tested for each subtest at each age level in raw score units. From these 't' analyses it was found that the 12 year-old group of boys excelled the girls in Arithmetic ($p < 0.01$). All other differences were negligible. (See Table 5.16).

Table 5.17 presents the 13 year-old group results. Here the boys did better in all Verbal and Performance subtests except for Coding B, in which the girls excelled the boys, but not significantly. The boys did better especially in Information, which was significant at the 0.01 level, and also in Similarities at the $p < 0.05$ significance level. Other differences in the tests were not significant.

Table 5.18 presents the results of the 14 year-old group. In this age-group the boys showed superior scores to the girls. Significant differences ($p < 0.01$ level) occurred with the Information and Arithmetic subtests, and with the Comprehension, Picture Completion, Picture Arrangement and Object Assembly, where the boys' scores were significantly higher than the girls' at the $p < 0.05$ level. Again, the girls excelled the boys, at the $p < 0.05$ level, in the Coding B test.

The results in Table 5.19 show the mean differences between the boys and the girls of the 15 year- old group in which the boys did better, at the $p < 0.01$ level in Information, Arithmetic and Object Assembly; and significantly better than the girls, at the $p < 0.05$ level in Comprehension and Picture Arrangement. The only subtest in which the girls did better was Coding B, but this difference was not significant for this age group.

In this regard some scientists have suggested that some of the differences in ability between the two sexes may be due to heredity (Vandenberg, 1972). It is generally known that girls are usually superior in verbal development and boys do better on tests of mechanical and special reasoning. The results from the present study, however, point to the overall superiority of the boys in Iraq.

Money (1968) reported that XO females (Turner's syndrome) are generally average in verbal ability but poor in special ability. This difference could indicate that one chromosome X is sufficient for verbal development but that a second X or a Y is needed for the full development of special orientation and form perception. Moreover, the fact that males (XY) are generally better in these abilities than females (XX) suggests that it is the Y which contributes to special abilities.

Miller (1985) reported the first evidence of gender differences in cell numbers for the human brain content. The brains of 13 men and 18 women were obtained and examined by D.F. Swaab and E. Fliers (1985) of the Netherlands

Institute for Brain Research in Amsterdam. One area, called the sexually dimorphic nucleus of the preotic area (SDN-POA) was found to be on average 2.5 times larger in men than in women Swaab and Fliers, (1985) reported that "In both sexes, the volume of this area and the number of cells within it decreased with age."

A review of literature by Maccoby and Jacklin (1974) revealed that, among the gender differences which the authors could find, girls have greater verbal ability than boys and that boys excel in mathematical ability, from about age 12. Kagan and Welsh (1966) found that boys tend to be more reflective, on average, than girls. The Scottish standardization of the WISC showed the same results and concluded in this study that the boys did better than the girls at the age of 9 years. This was only slightly significant, but there are significant differences, at the $p < 0.05$ level, in Information and Comprehension. For the 13 year-olds boys' and girls' differences were more apparent. Boys were superior in Information and Picture Completion at the $p < 0.01$ level; and in Vocabulary and Block Design tests at the $p < 0.01$ level. Girls were superior in Coding B at the 0.05 level. The same results were found in the present study. Also, this study agrees with the study conducted by Haritos-Fatouros (1963), who concluded that the boys excelled girls in all ten subtests but in Coding B the girls did significantly better than the boys, at the $p < 0.05$ level. Similar results were found in the study by Alexopoulos (1979), in which the boys' results were higher than the girls' in all WISC-R subtests except Coding B.

Leith and Basselt (1968) investigated differences in children's impulsive reflexivity style. Using a matching figures task, they showed that boys' errors correlate significantly only with their response speed whilst girls' errors do so only with their intelligence. The author concluded that there are gender differences in cognitive styles between boys and girls. Smith (1984) utilized a test of field dependency and creativity for a sample of ten-year-old boys and girls, to investigate the relationship between cognitive style dimensions. He concluded that boys performed significantly better, at the $p < 0.01$ level, than girls.

Seashore et al (1950) wrote, after describing the gender differences between boys and girls on the Verbal, Performance and Full Scale IQ:

> How shall one interpret these sex differences? Three explanations come to mind: 1. The tests are fair to both boys and girls, and boys actually do excel girls, especially at the later ages. 2. Boys and girls are the same in mental ability,but the chosen test items turned out to be slightly biased in favour of the boys. 3. Again, assuming that the general ability is not sex- differentiated, the sampling of boys was somehow chosen with a slight bias. The data at hand do not permit a resolution of these three choices. The safest assumption is that factors described in (2) and (3) are involved. Terman and Merrill found the same situation in their 1937 Revision of the Stanford Binet examination, and likewise could find no definitive answer for their data. (pp. 106-107)

## 6.4 Age Differences

Table 5.20 presents mean IQ scaled scores in Verbal, Performance and Full Scale tests, Standard Deviations, differences between means and 't' values for the mean differences between the 15 and 14 year-olds, 15 and 13 year- olds, and 15 and 12 year-olds. From Table 5.20 we can conclude that age- group differences between standardized IQs are not significant. The main reason for using this comparison is to find if there are any significant differences between age groups. As all the subtests were scaled to give a mean of 10 for each age group, it would have been surprising if real differences had occurred. These results provide a check on the standardization procedures.

## 6.5 Item Analysis

**Order of Difficulty from Main Sample.**

Table 5.21 shows the proportion of children passing each item within the four age--groups separately. The average percentages of students passing within the four age--groups and the order of difficulty of each item for the ten subtests in the present sample are presented. As we can see from Table 5.21, the order of the items obtained for the present sample was similar to that given in the British WISC-R in many cases. In the following sections we shall discuss the order of difficulty for each subtest.

### Verbal Test.

The Information Subtest.

In this subtest there was some slight disagreement, as we can see in Table 5.21, when the order of difficulty of our sample was compared with the British WISC-R standardization sample. Most items kept roughly the same order of difficulty in this study as the British WISC-R, but Items 12 and 16 appear to be very much more difficult for the Iraqi sample.

The Similarities Subtest.

The order of difficulty for the items in this subtest seemed to produce no major discrepancies between the British and the Iraqi samples. Items 1, 2, 3, 9 and 15 had the same order of difficulty, as did more than a third of the questions. Items, 4, 5, 6, 7, 8, 11, 12, 13, 14, 16 and 17 were slightly different from the British WISC-R order. Some of these items were slightly easier or slightly more difficult than the British standardization sample.

The Arithmetic Subtest.

This test had few minor discrepancies in the order of difficulty, with a similar general progression indicated from less to more difficult items as in the British sample. This agreement is perhaps due to the similarity of Mathematics syllabuses for children in Iraqi and British schools.

The Vocabulary Subtest.

This was a completely new test, constructed by the author using the pilot sample. As we can see from Table 5.21, there was almost total agreement between the pilot and the main study samples. These results indicate that considerable confidence can be placed in this new Iraqi subscale.

The Comprehension Subtest.

In the Comprehension subtest the order of difficulty was somewhat different from that of the British WISC-R test. Interestingly, some of the more difficult items with the British sample proved to be less difficult for the Iraqi sample. The opposite trend occurred with Item 10, which was found to be more difficult with the Iraqi sample.

**Performance Test.**

The Picture Completion Subtest.

In this subtest the order of difficulty of the items was generally similar to the British sample, with a few items being displaced somewhat more than others, e.g. Item 16. The greatest discrepancies were in Items 8 and 16. The rest of the items seemed to be adequate.

The Picture Arrangement Subtest.

There was considerable agreement as to the difficulty order between the British WISC-R and the present study. The only item which showed a slight discrepancy was Item 5, which was less difficult for the Iraqi sample.

The Block Design Subtest.

There was almost complete agreement among the items in this subtest in the order of their difficulty between our samples and the British ones. Only Items 10 and 11 of the British test were slightly displaced.

The Object Assembly.

In this subtest there was total agreement with the order of difficulty given by the British WISC-R.

## General Comments On Item Analyses.

An order of difficulty is indicated by the passes for each item in each subtest. As we know, the order of difficulty of test items is not something which can be established once and for all. This can be expected to vary somewhat even in samples from the same population.

Item analysis is very important for test construction because it summarises the results for each item on the test being constructed. From the data in Table 5.21 it seems that the items of the whole test make up a useful and reliable instrument in assessing intelligence with an Iraqi school population. Indeed, many of the items behave in the same simple way as they did in other WISC-R studies. No major differences between our items and those sampled in these analyses, occurred in the present study. Guilford (1954) wrote:

> Some generalizations can be made with respect to the dependability and the comparative values of the different methods of item analysis. First,it can be said that indices of difficulty are much more stable than indices of item validity. The type of test and of tested population would undoubtedly have bearings on the stability of item indices (p.215).

Bhatt (1972) pointed out that Wechsler did not mention any method of item analysis either in the WISC or WAIS manuals or in his book, The Measurement and Appraisal of Adult Intelligence.

## 6.6 Socioeconomic Status and IQ.

An analyses of variance was conducted on the results of IQ with socioeconomic status to find out if there were any significant differences between the groups. Tables 5.22 and 5.23 present the results of the group IQ means and the results of a one- way analysis of variance to discover any significant differences between socioeconomic groups. The F ratio was 25.80, indicating an F probability of $p < 0.001$, which is very significant. Thus, clear differences were found among the six socioeconomic groups in the present study in intelligence as measured by the WISC-R. The mean Full Scale IQs for school children bore a definite and progressive increase relative to an increase in the status of fathers' occupation, which is a ready index of socioeconomic status. The children of professionals, rentiers and merchants had a higher mean IQ of 112.95 than those of the other groups. The children of unskilled workers had a mean IQ of, 92.56 the lowest among the groups, while the children of semiskilled workers had a mean median value of 98.03.

The results for parents' socioeconomic status, therefore indicated a high association with children's IQ. (See Table 5.23). Most standard intelligence tests since the days of Binet have yielded substantial differences between members of different socieconomic classes. Terman (1918) reported mean IQs of 107, 100, and 93 for children classified as 'superior', 'average', and 'inferior' in social status (in agreement with the present study). In 1937 Stanford-Binet's IQ means for children of professional families were 18 to 20 points higher than those of children of labourers, and only 10 percent of the latter exceeded the average IQ of the former (McNemar, 1970).

Some psychologists have contended that standard tests systematically favour the middle class (Eells and Tyler, 1951) and have attempted to produce more 'culture fair' tests. Raymond B. Cattell's Culture--Fair Test, and Raven's Progressive Matrices Test utilize abstract problems, hopefully of equal novelty to all social classes. The tests do not involve language. However, average differences between social classes and between ethnic groups in the U.S.A. often remain conspicuous even in culture-fair measurements.

A study by Dastoor and Emovon. (1972) showed the same results as the present study. Children from low socioeconomic classes gain a low IQ in comparison to those of higher socioeconomic groups. Majoribanks (1972) in the U.S.A. studied 11 and 13 year-old boys from the middle and lower social classes and yielded a correlation between total IQ score of 0.31 ($p < 0.01$) with father's education and 0.43 ($p < 0.01$) with father's occupation. These results support the present study showing that socioeconomic status correlates somewhat higher with IQ than does the educational level of the head of the house. Michael (1979) found a correlation of .34 ($p < 0.01$) between social position and IQs Thus, there is some consistency in correlational pattern across studies, between the socioeconomic status and IQ, which fits in with the finding of the present study.

## 6.7 Family Size and IQ.

Table 5.24 presents the number of students within different family sizes for each age- group. From this table we see that the majority of families possessed 4 to 6 children, i.e. 90, 102, 96 and 92 children for the 12, 13, 14 and 15 year-old groups respectively. The minority of families had 1 to 3, i.e. 40, 38, 22 and 36 children for the 12, 13, 14 and 15 year-old groups respectively.

The results of the relationship between family size and IQ are shown in Table 5.25 . From this table we can see that family size is an important factor which may affect the intelligence of children. Table 5.25 presents the means of the Full Scale IQs for the age groups 12, 13, 14 and 15 years. Significant differences ($p < 0.01$ level) were found between the families having 1 to 3 children and families having 4 (and more) children in all age groups, except for the 15 year-old group where the difference was not significant. The highest mean IQ difference was found with the age group 12 and the lowest with the age group 15 years.

The results in Table 5.25 show that the fewer children the family has, the higher the IQ tests results appear to be. This may be due to the fact that small size families can afford to spend more time helping their children to develop cognitive skills. In support of this study Feldman (1981) concluded that increasing family size may result in a decrease in the amount of time or opportunity spouses have to do things together and this in turn may affect the child's social performance.

Greenberg (1985) blames the increase in size of families for the drop in average intelligence scores with the Scholastic Aptitude Test (SAT) in 1963 for the United States of America. In general he reported that the smaller the family, the higher the children's intellectual development and scholastic achievement. As families have recently become smaller, the current upswing in scores will most likely continue. Zajonc, (1984) and several other scientists report a number of apparently positive effects that the small family size has on children's intellectual development.

In Los Angeles at the annual meeting (1984) of the American Association for the Advancement of Science, it was reported that an only child is exposed mainly to his parents' adult environment--the way they interact and deal with their problems--and to adult language. In contrast, the report notes, a child in a family of 10, whose oldest sibling is 12, is surrounded by intellectually immature individuals with less--developed vocabularies. The scores of intelligence are rising, the report says, primarily because of the shrinking size of the family unit.

One of the reports in that meeting (AAAS) by Blake (1984) of the University of California at Los Angeles, claimed that family size also has a "gigantic" effect on other aspects of a child's education, including grades and whether he/she graduates from high school and goes on to college. Analyzing data from two national surveys of 56,000 white fathers, Blake found that next to a father's educational level, family size is the most important variable in measuring progress in school.

In a separate study of IQ, Higgins (1983) reports that the large family is associated with children's lower IQs. In his analysis of 300 families, Higgins reported that "parents of large families tend to have lower IQs" and concludes that the children, therefore, have inherited similar IQ levels.

According to Zajonc (1984), the findings on family size show that those at the top of the birth order have the highest IQ scores, which suggests that the optimum situation seems to be a two-child family with a spacing of more than two years between children.

## 6.8 Urban-Rural Residence and IQ.

Table 5.26 shows the Full Scale IQ means of the children from urban and rural areas, where the urban residence children did better than the children of rural residence. Differences analysed by 't' tests show that in all groups -- 12, 13, 14 and 15 years--there were significant differences in mean IQ scores at the 1 % level. These significant differences may have resulted from the fact that there are differences between urban and rural life styles and in the quality of both economic status and teaching as between the urban and rural areas.

The results of the present study are in agreement with those obtained by Wilson (1982) when she established local rural norms for Alaska. Low scores in Verbal, Performance and Full Scale of WISC-R were found in her study. Also, McNemar (1964) reported that the test results of urban and suburban children resembled each other, but that rural children averaged 5 to 10 points lower in intelligence tests. Alexapoulos (1979) in his study showed that urban children scored significantly higher with WISC-R than their rural counterparts. Kokkevi (1975) found the same results as this study--that rural subjects scored lower than urban subjects in WISC-R.

In supporting the present findings Wechsler (1958) reported that the urban-rural residence differences appeared to be related to level of intelligence. Tyler (1965) supported Wechsler's view, and Farber (1968) concluded, "There is a general tendency for mental retardation to be more prevalent in rural than in urban areas." (p.78).

## 6.9 Years of Schooling of Parents and IQ.

A one- way analysis of variance established the influence of educational factors associated with fathers, on their children's intelligence. Tables 5.27 and 5.28 present the Full Scale IQ mean values for "the parents' years of schooling" groups. From the F ratio group differences for IQ means were found to be significant at the 0.01 level. A similar but more significant result was found for Full Scale IQ means and parents' socieconomic status. Many researchers have found similar results in which the socioeconomic status appears to be more influential than the father's educational level.

Routsoni (1965) provided information on the influence of differential environmental conditions. He found that the better- educated parents tend to have children with higher IQs; the correlations obtained for his study of 13 and 14 year-olds were significant at the $p < 0.01$ level. Reschly (1976) reported that the urban-rural comparisons of intelligence were found to be significant in favour of urban children.

Table 6.1

Correlations between Father's Education and Child's IQ.

| Child's Age | Boys's Corr. | Girls's Corr. |
|---|---|---|
| 1 yrs. | 0.01 | 0.17 |
| 3 = | -0.08 | 0.32* |
| 4 = | 0.10 | 0.36* |
| 5 = | 0.10 | 0.34** |
| 6 = | 0.23 | 0.43** |
| 7 = | 0.17** | 0.44** |
| 8 = | 0.34** | 0.41** |
| 9 = | 0.27** | 0.41** |
| 10 = | 0.29** | 0.40** |
| 11 = | 0.27** | 0.42** |
| 12 = | 0.35** | 0.42** |
| 13 = | 0.41** | 0.33** |

From Honzik (1963).

* Significant at 0.05 level

** Significant at 0.01 level

Shaw and White, (1965), in an investigation of high school students with above-average intelligence, obtained results indicating that where the boys and girls had above average IQs, their parents' educational level was high. In Kimball's (1952) study, very intelligent adolescent boys had well- educated fathers. With fathers who had a lower degree of education their children tended to have low grades. Honzik's (1963) sample of individuals described as representative of the children born in Berkeley (USA) between January 1st 1928 and June 30th 1929, obtained the correlations shown in Table 6.1 between father's education and the child's IQ at different ages. With increasing age both boys' and girls' IQs appeared to be significantly related to educational level reached by parents.

## 6.10 Geographical Region and IQ.

Table 5.29 presents the results of IQ differences between the three major geographical regions. It shows that the highest difference is found between the Northen and the Middle areas (p 0.01 level). Also, there is a significant difference between the Middle and the Southern areas (p 0.01 level). The difference between the South and the North is not significant.

These results in favour of the Middle region may be due to the fact that: 1) There are better- qualified and experienced teachers there, because most graduate teachers prefer to work in the Middle region; 2) Better economic opportunities exist there; 3) Some minorities live and study in the Arab areas in the north e.g. the Kurds and Turks (the Turks came and settled in the Northern region when the Ottomans occupied Iraq); and 4) Security is often more stable in the Middle region.

Summary

The present study has provided secondary schools in Iraq with a very appropriate intelligence test which has been standardized on a comprehensive sample of the school population. By retaining much of the content of the original British WISC-R test, this Iraqi version of that well-known and extensively used intelligence test, the WISC-R (Iraqi version) can now be used with confidence in selecting pupils studying in intermediate schools or at a later stage in the preparatory schools. In particular, administrators, psychologists, remedial education staff and so on, will find it useful to have a standardized intelligence test available for assessing students in order to make decisions about their education.

# BIBLIOGRAPHY


Ahumada,I. Adaptation and Standardization of the WISC in Latn-American Countries. Revista Mexicana de Psicologia, 2, 10, 813-823,1966.

Aiken,L. R. Psychological Testing and Assessment (5th ed.). Boston: Allyn and Bacon, (1985).

Al-Aubaidy S. Standardization of the WIPPS Performance Scale, Ph.D. Thesis, University of Baghdad, (1987).

Alexopoulos, D. S. Revision and Standardization of the WISC-R for the age range 13-15 years in Greece, Unpublished Ph.D. Thesis (1979).

Allen, M. J. & Yen, W. M. Introduction to Measurement Theory California: Wadsworth, (1979).

Altus, W.D. The Differential Validity and Difficulty of Subtests of the Wechsler Mental ability Scale.Psychological Bull. 42, 231-249, 1953.

The American Association for the Advancement of Science Annual HandBook Report,San Francisco, (1984)

Anastasi, A. Psychological Testing. New York Macmillan, (1976)

Anastasi, A. Psychological Testing (2ed, Ed.). Mcmillan Company, (1961).

Anastasi, A. Differential Psychology. New York, The Macmillan Company, (1954).

Anastasi, A. & Cordova, F. A. Some effects of Bilingualism upon the Intelligence Tests Psychology of Puerto-Rican Children in New York City. Journal of Educational Psychology,44, 1-19, 1953.

Andrade, I.R. Considerateness Acerca de la estandari de Pruebas a Lation America, Conillustraciones de la Adaptacion del WISC a Mexico. Revista Mexicana de Psicologia, 2, 10, 1966.

Anderson, M. Use of the WISC-R with a Learning Disabled Population: Some Diagnostic Implications, Psychology in the Schools, 13, 381-386, 1976

Applebaum,I.R. & Bee, J.M. Social Class and Test Performance: Comparative Validity of the Peabody with WISC and WISC-R for two reader. Economic Groups Psychology Report, 40, 139 -145,(1978).

Arsenian, S. Psychological Testing of American Minorities: New York, (1945).

Arthur, G. A. Point Scale of Performance Tests. New York: Commonwealth Fund, 1930.

Arthur, G. New point Performance Scale. Journal of Applied Psychology, 9, 390-416, 1925.

Bailey, D. B. & Harbin, G. L. Non - discriminatory Evaluation. Exceptional Children, 8, 590 - 596, 1980

Bannatyne, A. Diagnosis: A note on Recategorization of the WISC Scale Scores. Journal of Learning Disabilities, 7, 272-273, 1974.

Bannatyne, A. Language, Reading and learning Disabilities. Springfield, Charles C. Thomas Press. (1971).

Beck, N. C., Ray, J. S., Sedemberg, M., Young, T. R. & Gamache, M.P.Development and Cross-Validation on Short Form for the WISC-R. Journal of Consulting and Clinical Psychology, 51, 1983 864-869.

Bernstein, B. Social Structure, Language and Learning Educational Research, 3, 163-176, 1961.

Bhatt, M. G. Adaptation and standardization of the Wechsler Intelligence Scale for Gujarati Children, Indian Educational Review 7,31-52, 1972

Bhusham, V. Adaptation on an Intelligence Test From English to French. Journal of Education Measurement, V.11 No.1 Spring 1974.

Binet, A. and Simon, T. The Development of Intelligence in Children. New York: Arno Press, (1973).

Binet, A. & Simon, T. The Development of Intelligence in Children. Brandow Publishing. (1916)

Binet, A. & Simon, T. Le Development de Intelligence Chez lees enfants. L' Annee Psychology, 14,1- 94, 1905.

Binet, A. & Henri, V.Le Psychologies individual. L'· Annee Psychology, 2,411-465, 1895.

Bonilha, L. C. Adaptation of the Verbal WISC for Children Aged 7-15 Years. Revista de Psicologia Normal e Palogica, 14, (3) 123-134, 1968.

Brittain, D. An Investigation into the Charactor and Experiment Scale IQ.? Psychology in the School, 17, (1), 40-46, 1976.

Brown,F. G. Principles of Educational and Psychological Testing. Second Ed., Holt, Rinehart and Wirston, USA, (1976).

Brown, David and Otts, M. Study investigating the Feasibility of Creating a short Form of WISC-R for Disabled Readers. Paper Presented at the Annual Meeting of the Mid-South Educational Research Association (1983).

Bruner, J. S., Shapiro, N. & Tagiuri, R. The Meaning of Traits in isolation and in Combination.In R. Tagiuri & L. Petrollo (Eds.) Perception and Interpersonal Behaviour. Stanford, California: Stanford University Press, 1958.

Burt, C. The Evidence for the Concept of Intelligence. British Journal of Educational Psychology, 1955, 25, 158-178.

Butcher, H.J. Human Intelligence its Nature and Assessment. London: Methuen (1968).

Cantor, N. Prototypiality and Personality Judgments. Ph.D.Dissertation, Department of Psychology, Stanford University, (1978).

Carroll,John. On the Scientific Basis of Ability Testing. American Psychologists, V. 36, No.10., p.1012-1020, October. 1982.

Cattell, J. M. Mental Tests and Measurement. Mind 1890, 15, 373-381.

Cattell,J. M. and Farrand, L. Physical and Mental Measurement of the Students of Columbia University.Psychological Review, 3, 618 -648, (1896).

Cattall,R. B. Abilities: Their Structure, Growth, and Action. Boston: Houghton Mifflin, (1971).

Chomsky, N. "IQ Lists: Building Blocks for the New Class System". Ramparts, (July 1972) 24-25.

Clarizio, H. F. In Defense of the IQ. Test. School Psychology, 8, 79-88, (1979).

Clarke, A.D.and Clarke, M. Recent advances in the Study of Subnormality MIND (National Association for Mental Health) 1975.

Cleary, T. A. Testing Bias: Prediction of Grades of Negro and White Students in Integrated Colleges. Journal of Education Measurement, 2, 115, 123, 1975.

Cole, Michael."An Ethnographic Psychology of Congestion".Cross-Cultural Perspectives on Learning Edu cation Richard Brishlin, S.Bochner and W. Lonner New York (1975).

Cole, O. Bias in Testing. American Psychologist. 10, 1066 -1077, 1981.Coles, Gerald. "The Learning Disabilities Test Batery :Empirical and Social Issues" Harvard Education Review 48, 3 August (1978).

Comptroller General of United States. Opportunity to improve Indian Education in Schools operated by the Bureau of Indian Affairs. Washington, D. C.: General Account Office, April, 27, 1972.

Connelly, James. Recategorized WISC-R Score Patterns of older and younger Referred Tlinit Indian Chil dren. Psychology in the School Vo. 20, 271 275, (1983).

D'Andrade, Roy. "Cultural Constructions of Reality" Cultural Illness and Health, Eds. Laura Nader and Thomas Maretzt: (Washington D.C. American Anthropological Association, 1973, 78, 144-149.

Dastoor, D. P. & Emovon, A.C. Performance of nine-years old Nigerian Children on Block Design Test, West African Journal of Education, (1972).

Dean, R. S. Predictive of the WISC-R with Mexican American Children. Journal of Special Psychology (17) 1 55-58, (1979).

Dean, R. S. Bilingualism and the Measurement of Intelligence: Review of a Decade of Research. The Journal of Genetic Psychology 103, 259-282, (1980).

Dean, R. S. The Validity and Reliability of abbreviated Version of WISC-R. Journal of Educational and Psychological Measurement, 37 , 1111-1115. (1976).

Dean, R. S. Analysis of the PIAT with Anglo and MexicanAmerican Children. Journal of School Psychology. 15, 1977, 329-333.

Delvaux R. Adaptation of the WISC Verbal with Congolese African Journal of Education, 11 (3), 1969.

Deutsch, Cynthia. "Social Class and Child Development". Review of Child Development Research Chicago University of Chicago Press, (1973).

Dirks, Jean. Can Short - Form WISC-R Tests Identify Children with High Full Scale IQ.? Psychology in School, 17, (1), 40-46, 1980.

Dirks, J., Wessels, F. & Quarforth, A. WISC-R and the Short Form, for Borderline Children. Handbook, Calefornia, (1980).

DuBois, P. H. A History of Psychological Testing. Boston: Allyn and Bacon, (1970).

Eells, Kenneth D. & Tyler, R.Intelligence and Cultural Differences(Chicago:The University of Chicago. Press, 1951).

Ertz, H. WISC-R Scores among Matched Pairs of non-Indian and American Indian Students who were Displaying learning Disorders. White Cloud Journal, Vol. 16, 1985.

Evans, B. & Waites, B. IQ. and Mental Testing An Unnatural Science and its Social History. Atlantic Highlands, N. J.: Humanities Press, (1981).

Eysenck, H. J. & Kamin, L. The Intelligence Controversy. John Wiley, (1981).

Eysenck, H. J. The Structure and Measurement of Intelligence Berlin: Spring-Verlag, (1979).

Eysenck, H. J. The Measurement of Intelligence. Baltimor: Williams & Wilkins. (1973).

Farber, B. Mental Retardation: Its Social Context and Social Consequences. Boston: Houghton, Mifflin (1968).

Feldman, L. Beyond Universals: Toward a Development Psychology of Education.Educational Research, Vol. 10, (9), 21-26, 1981.

Figueroa, R. A., Delgado, S. & Ruiz, M. The System of Multicultural Pluralistic Assessment . School Psychology Digest, 8, 28 - 35, 1984.

Finch, A. G., Ollexdick, T. H. & Ginn, F. W. WISC Short Form with Mentally Retarded Children. American Journal of Mental Deficiency,78, 144-149 1973

Finch, A. J. & Others. Short-form and Factor-Analytic Studies of the WISC-R with Behaviour Problem Children.Journal of Abnormal Child Psychology Vol. 7 (3) 337-344, Sep. 1979.

Galton, F. Genius Inquiry into its Laws and Consequences. New York st. Martin's Press, (1869 editions: reprinted in 1978).

Galton F. Inquiries into Human Faculty and its Development. New York st. Martin's Press, 1966.

Galton, F. Heredity Genius: An Inquiry into its Laws and Consequences. London: MacMillan, (1869).

Ghuman, P. P. Dasen and V. Bevali (Eds.), N. D. Cognitive Development in Indian and Pakistani Children, Forthcoming, (1982).

Gilbert, Harry B. "Individual and Group Differences". Crowell - Collier Corporation, (1971).

Glasser, A. J. Interpretation Clinica de la Escala de Inteligencia Wechsler Para ninos. Madrid, T. E. A. (1977).

Goh, D. W. Empirical Selection in the Design of Intelligence Short Form the WISC-R Example. Applied sychological Measurement, 3, 75-82, 1979.

Golden, M. & Birners E. "Social Class and Cognitive Development in Infancy, Merrill - Palmer Quarterly 14, 139-149, 1968.

Goldman, R. D. and Hewitt, B. N. Predicting the Success of Black, Chicano, Oriental and White Students Journal of Educational Measurement, 11, 129-135, 1974.

Goodenough, M. Goodenough Scale Evaluation of Human Figures drawn by Schizophrenic and non-Psychotic Adults.Journal of Clinical Psychology, 5, 396-398, 1949.

Goody, M. A comparative Study of Cross Cultural Patterns of Intellectual Abilities on WISC and WISC-R Among Anglo, Chinese, and Hispanic Educationally. Handicapped Boys with Reading Disabilities. University of Sun Francisco, (1981).

Gottesman, I. L. "Statement of Irving L. Gottesman", Environment, Intelligence, and Scholastic Achievement. (Washington, D. C. : Government Printing Office, 1972 37p.).

Gould, S. J. The Mismeasure of Man, New York Norton, (1981).

Greenberg, J. Family Size Tend to SAT, IQ Scores. Science News of the Week. Vol. 127, 340, 1985

Guilford, J. P. The Nature of Human Intelligence New York McGraw - Hill, (1967).

Guilford, J. P., & Hoepfner, R. Structure of Intellect Factors and their Tests,Reports from Psychology Labrary, University of South California, No. 36, (1966)

Guilford, J. P. Psychometric Methods, (2ed. Ed.). New York: McGrae Hill, (1954).

Gutkin, T. & Reynolds, C. R. Factor Similarity of the WISC-R for Anglo and Chicanos referred for Psychological Services. Journal of School Psychology, 18, 34-39, 1980.

Guttman, L. A Faceted Definition of Intelligence. In R. Eifermann (Eds.) Scripta Hierosolymitana, Magnes Press, (1965).

Hackett, E. F. The Reliability of Short Form of the WISC-R with Educably mentally Handicapped Subject Dissertation Abstracts International, 1979, 3, 75-82.

Hamadi, S. K. Al-Thawra Newspaper No. 2835, 25th October (1977)

Haritos - Fatouros, M. A study of the WISC Applied to Greek School - Children. Unpublished Master Thesis: University of London, (1963)

Harrington, G. Intelligence Tests may Favour the Majority Groups in Population. Nature, 258, 708-709, (1975).

Hays, J. R. & Smith, A. L. Comparison of WISC-R and Culture Fair Intelligence Test Scores for three Ethnic Groups of Juvenile Delinquents. Psychological Reports, 31, 1980, 465-468.

Healy, W. & Fernald, G. M. Tests for Practical Mental Classification. Psychology Monog. 13, No. 54, 1911

Helmstadter, G. C. Principles of Psychological Measurement.New York; Appleton-Centre Crofts, (1964).

Higgins,J. V. Family Size and IQ. Michigan. State University in East Lansing Report, (1983).

Hirshoren, A. & Karale, K. Profile Analysis of the WISC-R a Continuing Malpractice, The Excep tional Child, 23, 83-87, (1975).

Hobb, K.L. WISC-R Split - Half Short Form. Los Angeles, Western Psychological Services, (1980).

Honzik, M. P. A Sex Difference in the age of onset of the Parent child Resemblance in Intelligence. Journal of Educational Psychology, 54, 231-237, 1963.

Horn, J. L. & Cattell, R. B. #Refinement and Test of the Theory of Fluid and Crystallized Intelligence, Journal of Educational Psychol-ogy. 57, 253-270, (1966).

Hynd, G. W. & Garicia, W. Intellectual Assessment of the Native American Student. School of Special Digest, 8, (4) 446-454, 1979.

Jensen, J.E. Bias in Mental Testing. New York, The Free Press (1980).

Jensen, A. R. Bias in Mental Testing, New York; Free Press (1979).

Jensen, A. R. Genetics and Education. New York; Harper and Row, (1972).

Jensen, A. R. Hierarchial Theories of Mental Ability. In Dockwell Education on Intelligence, London Methaen (1969).

Jensen, A. R. Jensen's Theory of Intelligence: A reply. Journal of Education Psychology, 60, 427-431. (1970).

Jensen, A. R. Varieties of Individual Differences in Learning. In R. Gagne (Ed.) Learning and Individual Differences. Columbus Merrill, (1967).

Jensen, A. R. Bias in Mental Testing. New York: Free Press (1980).

Jensen, A. R. How Much can we Boost IQ and Scholastic Achievement. Harverd Education Review, 34, 1- 12, (1972).

Justman, J. Assessment of Intelligence of Disadvantaged Children. Education, 87, 354-362, 1967.

Kagan, J. P. & Welsh, L. Conceptual Impulsivity and Inductive reasoning. Child Development, No.37, 583-594, 1966.

Karnes, F. A., & Brown, K. A Short-Form of the WISC-R for Students. Psychology in the School, Vol.18 (2), 169-173, 1981.

Kaufman, Alan. "The WISC-R and Learning Disabilities Assessment: State of the Art", Journal of Learning Disabilities (November 1981).

Kaufman, Alan, Schiff, M. & Kaufman, N."Scatter Analysis of WISC-R Profiles for Learning Disband Children with Superior Intelligence" Journal of Learning Disabilities,14, 7 Aug. (1981) 401

Kaufman, Alan S. Intelligence Testing with the WISC-R. New York: John Wiley & Sons (1979).

Kaufman, Alan S. A Four Test Short Form on the WISC-R. Contemporary Education Psychology, 1, 180-196 (1976).

Kaufman, J. M. & Hallahan, D.P. (Eds.). Teaching Children with Learning Disabilities: Personal Pers pectives Columbus Ohio: Charles E. Merrill Publication, Co., (1976).

Kaufman, Alan S. Factor analysis of the WISC-R at 11 age Levels Between 6 1/2 and 16 1/2 Years Journal of Consulting and Clinical Psychology, 43, (2) 135-147, (1975).

Kennedy, L.P. & Elder, S.T. WISC-R: An abbreviated Verson. Journal of Clinical Psychology, 38, 1982 174-178.

Khayyer, M. & Mojdehi, H. Intelligence: Iranian Male Delinquents Compared with Delinquents on Selected WISC Scales, Psychological Reports, V.44, P. 282, (1979).

Kimball,B. The Sentence Completion Technique in a Study of Scholastic underachievement. Journal of Consulting Psychology. 16, 353-358, 1952.

Knox, H. A. A Scale Based on the work at Ellis and for estimating Mental Defect. Journal of American Mental Assoc., 62, 741-747, 1914.

Komm, Richard. "A comparison of the WISC-R with the BITCH as a Measured by a Conventional Achiev ement Test within a Black Population at Different Social Class Level (Educational Dissertation University of San Francisco, (1978).

Kokkevi, A. "The Wechsler Test of Adults: Mental Retardation in Evaluation of Cognitive Psychological Testing, No. 52, 1975.

Korchin, S. J. Clinical Psychology and Minority Problems American Psychologist, 35, 262-269, 1980.

Lamb, Michael E. The Role of the Father in Child Development. A Wiley Interscience. Publication. New York, (1966)

Leith, G. and Bassett, R. Mode of Learning and Personality. Birmingham University Research Report on Programmed Learning, No. 14, 120, 1968.

Lemgruber, V. & Paine P. Adaptation of the WISC Verbal Scale. Brasileiros de Psicologia V.33 pp. 32- 56, 1981.

Lunemann, A. The Correlational Validity of IQ. as a Function of Ethnicity and Desegregation. Journal of School Psychology, 12 (3) 263-268, 1974.

Maccoby, E. E., & Jacklin, C. N. The Psychology of Sex Differences. Stanford California: Stanford University Press, (1974).

MaCllough, C. & Jacqueline L. The use of WISC-R Scale in the assessment of Native Americans of the Columbia River Basin, Psychology in School, V. 22, 23-28, 1985.

MacMillan, D. & Meyers, E. The Non-discrimin-atory Testing Provision. Viewpoints,53, 39-55, 1977.

MacNemar, B. A Critical Examination of the University of Iowa Studies on Environmental influences upon the IQ. Psychological Bulletin, 37, 1970 63-69.

Malin, A. J. An Indian Adaptation of the WISC. Journal of Vocational and Educational Psychology, V.10 (4), 128-131, 1954.

Maltzman, C. "Cross - Cultural Patterns of Responses to the WISC & WISC-R by Chinese - American, Hispanic - American and Occioental - American Educationally Hendicapped. San Francisco,(1983)

Manson, G. E. Personality Difference in Intelligence Test Performance.Journal of Apple. Psychology, 9, 1925, 230-255.

Maretzki, Thomas. "Preface" Cultural Illness and Health. Eds.Laura Nader and Thomas Maretzki, Washington, C. American Abthropology Associa-tion, (1973).

Marjoribanks, K. Environment, Social Class, and Mental Abilities. Journal of Educational Psychology 63, 103-107, 1972.

Marmoral, E. I. & Brown, F. Comparison of Bender -Gestalt and WISC correlations for Puerto Rican, White and Negro Children. Journal of Clinical Psychology, 31, 465-468, (1975).

Marsh, Herbert. Ethical and Economic Issues: Trustee Interest and Involvement in Academic Policies for Faculty Consulting. New York, (1977).

Martinez, J. L. Chicano Psychology. New York: Academic Press, (1977).

Matarazo, Joseph. Wechsler's Measurement and Appraisal Adult Intelligence. Baltimore: The Williams Company, (1972).

McFie, J. Preliminary Results With an Individual Intelligence Test.West African Journal of Education. VXI No.1 P. 5-8 (1970)

McNemar, Quinn. Lost: Our Intelligence. Why ? American Psychologist, 19, 871-882, 1964.

McShane, D.A. Recategorized Wechsler Scores of American Indian Children. Ph.D. Thesis,George Peabody College, University Microfilms, (1980).

McSane, D. A. & Plas, J. M. Wechsler Scale Performance Patterns of American Indian Children. Psychology in the School 19 (1), 8-17 1982.

McShane, D. A. A Review of Scores of American Indian Children on the WISC-R. White Cloud Journal, 1 (4), 3-10, 1982.

McShane, D. A. Bias in Assessment of American Indian Children. Listening Post, 1, 8-19, 1979.

Mercer, J. & Lewis, J. Using the System of Multicultural Pluralistic Assessment to Identify the Gifted Minority Child.National State Leadership Training Institute on the Gifted and Talented, Los Angeles, California, (1971).

Mercer, J. Validity "Bias" and Fairness: An Analysis. A Journal of Educational Studies, Vol. 9, (1), 1-16, 1976.

Mercer, J. Sociocultural Factors in Labelling Mental Retardates. Peabody Journal of Education, 48, (3), 188-203, 1971.

Mercer, S. System of Multicultural Paralistic Assessment (SOMPA) Technical. New York, Psychological Co. (1979).

Mercer, S. Sociocultural Factors in Labelling Mental Retardates. Peabody Journal of Education, 48 (3) 1980, 188-203.

Michael, R. Family Configuration and Achievement.Social Psychology Quarterly, V.42, (2), 135-148,1979.

Miele, F. Cultural Bias in the WISC. Intelligence, 3, 149-164, (1979).

Miller, J. A. Sex Differences found in Human Brains. Science News of the Week, Vol. 127, 1985.

Ministry of Education, Iraq. A Plan for Pre-School Education. 1975/1980.

Ministry of Education. The Seventh annual Educational Conference. Baghdad, Iraq, (1976).

Ministry of Planing, Iraqi General Census for the Year 1987,

Mishara, S. P. Evidence of Item Bias in the Verbal Subtests of the WISC-R for Mezican - American Children. Journal of Psychoeducational Assessment, 1, 321-328, 1983.

Money, John. Cognitive Deficits in Turner's Syndrome. Progress in Human Behaviour Genetics. Baltimore Johns Hopkins Press, (1968).

Moran, R. B. Observations and Recommendation on the Purto - Rican Version of the WISC. Pedagogia, 10, 89 - 98, (1962).

Naglieri, J. A. Does the WISC-R Measure Verbal Intelligence for non - English Speaking Children? Psychology in School, 19, 1982, 478-479.

National Institute of Health. Annual book, New York, (1979).

Neisser, V. The Concept of Intelligence. In R.J. Sternbery & D.K. Detterman (Eds.) Human Intelligence Prespectives on its Theory and Measurement. Norwood, N.J. : Ablex, (1979).

Newman, H. H., & Others Twins: A Study of Heredity and Environment. Chicago University of Chicago Press, (1955).

Oakland, T. D. Research on the ABIC and ELP: A revisit to an old topic. The School Psychology Digest, 8, 211-213, 1979.

Oakland, T. Non Biased Assessment in Counselling Issues and Guide-lines Measurement and Evaluation in Guidance, 15, 107-116, 1982.

Oakland, T. Predictive Validity of the WISC-R IQs. and Estimated Learning Potent. A Paper Presented at the Annual Convention of the American Psychological, Montreal (1977).

Oakland, T. & Matuszek, P. Using Tests in nondiscriminatory Assessment of Minority Children (pp.52-69). New York: Brunner/Mazel (1977).

Oakland, T. Psychological and Educational Assessment of Minority Children. New York Brunner, Mazel Inc. (1977).

Ogha, John. Minority Education and Caste (New York Academic Press) 1978..

Ollendick, M., Genshafft, J. & Geushaft, J. Comparison of Bilingual Children on the WISC-R and the Escala De Intelligencia Wechsler Para Ninos. Psychology in Schools, 18, (2) 157-163, 1977.

Olmedo, E. A Culturation: A Psychometric Perspective. American Psychologist; V.34, No.11, 1061-1070, Nov. 1980.

Olmedo, O. Testing Linguistic Minorities. American Psychologist, 10, 1981, 1078-1085.

Oplesch, M., Genshaft, J. Comparison of Bilingual Children on the WISC-R and the Escala De Intelligencia Wechsler Para Ninos. Psychology in the Schools, 18 (2) 159-163, 1981.

Padilla, E. R. & Roll, S. Standardization WISC-R for Mexican Students.Interamerican Journal of Psychology, 162, 122-128, (1982).

Paine, P. A. & Garcia, V. L. A Revision and Standardization of the WISC Verbal Scale for Children for use in Brasil.Interamerican Journal of Psychology, 8, 3-4, 225-231, 1974.

Palmer, M. & Gaffny, A. D. Effects of Administra-tion of the WISC in Spanish and English and Relationship of Social Class to Performance. Psychology in Schools 9, 61-64, 1972.

Peel, M. & Lambert, F. The WISC-R and Internal Evidence of Test Bias with Minority Groups. Journal of Consulting and Clinical Psychology, 49, 919-927, 1979.

Piaget, J. Need and Significant of Cross-Culture Studies in Genetic Psychology. In B. Inhelder and H. H. Chipman (Eds.) Piaget and his school. New York: Springer-Veriag (1976).

Piaget, J. & Inhelder, B. The Psychology of the Child. New York: Basic Book (1969).

Pintner, R. Intelligence Testing Methods and Results. New York: Holt, (1923).

Prewitt - Diaz, Joseph, O. & Munoz, G. A Reliabil-ity Study of the WISC with Purto-Rican Children Grade K-9. in Purto Rico. (1980).

Ramirez, M. & Cataneda, A. Cultural Democracy, Development and Education. New York Academic Press,(1974).

Ramos Lopez, E. C. Adaptacion de los Subtests de Vocabu-lario de las Escalas de Inteligencia Stanford Binet Forma L-M y Wechsler Para Nino. Revista de Psicologia Generaly Aplicada, 25, 102, 53-68, 1970.

Rasbury, W. C.,Falgut, J. C. & Perry, N. W. A Yudin - Type Short Form of the WISC-R; Two aspects of Validation. Journal of Clinical Psychology, 34, 120-128, 1978.

Rees, M. A. Gwenda, M. A. Thesis, Wales, An Investigation into the Vocabulary Development in Welsh and English in pupils 9 to 10 years in a Mixed Language area in Relation to a Measure of Bilingual Background.(1952).

Reschly, D. J. Evaluation of the effects of SOMPA measures on Classification of Students as Midly Ment aly Retarded. American Journal of Mental Deficiency, 85 (1), 16-20, 1981.

Reschly, D. J. "Psychological Testing in Educational Classification and Placement". Americam Psychologist. 36, 1094-1102, 1981.

Reschly, D. & Sabers, D. Analysis of test Bias in four Groups with Regression Definition. Journal of Educational Measurement, 16, (1), 1-9, 1979.

Reschly, D. J. Validity of the WISC-R Factor Scores in Predicting Achievement and for four Sociocultural Groups. Journal of School Psychology, 17, 355-359, (1979).

Reschly, D. J. & Sabers, D. L. Analysis of Test Bias in Three Groups with the regression Definition. Journal of Educational Measurement, V. 16, No.1 Spring (1979).

Reschly, D. J. & Reschly, J. WISC-R Factor Structure among Anglos, Black Chicanos, and Native American · Papagos. Journal of Consulting and Clinical Psychology, 46, 417-422, 1978.

Reschly, D. J. & Jipson, F. J. Ethnically, Geographic local, Age, Sex, and Urban-Rural Residence as Variables in the Prevalence of Mild Retardation. American Journal of Mental Deficiency. 81, 154-161, (1976).

Reschly, D. J. & Jipson, F. J. Ethnicity, Geographic Local, Age, Sex, and Urban-Rural Residence as Variables in the Provalence of Mild Retardetion. American Journal of Deficiency, Vol. 2, 159-161, 1976.

Resnick, L. The Native of Intelligence. Hillsdale, N. J.: Lawrence Eribaum Associate. (1976).

Resnick, R.J. An abbreviated form of the WISC-R: Is it Valid? Psychology in Schools 14,426-429, 1977.

Resnick, R. J. & Entin, A. D. Is an abbreviated Form of the WISC Valid for Afro-American Children ? Journal of Consulting and Clinical Psychology 36, 97-99, 1971.

Reynolds, C. & Jensen, A. WISC-R Subscale Patterns of Abilities of Black and Whites Matched on Full Scale IQ. Journal of Educational Psychology, 75 (2), 207-214, 1983.

Reynolds, C. R. The Problem of Bias in Psychological Assessment in C. R. Reynolds & T. B. Gutkins (Eds) The Handbook of School Psychology, 1982, 178 - 208.

Reynolds, C. R. & Gutkin T. B. A Regression analysis of Test Bias on the WISC-R for Anglos and Chicanos Referred for Psychological Services. Journal of Abnormal Children Psychology, 8, (2) 237- 243, 1980.

Routsoni, A. Rorschach Responses of Eight, Nine, Ten, and Eleven year-old Greek Children:A preliminary Study; The Rorschach Newsletter, 10 (2), 1965.

Rutter, M. & Madge, N. Cycles of Disadvantage. London: Henemann Educational, (1976).

Samuda, R. J. Psychological Testing of American Minorities Issues and Consequences. New York Dodd, Mead & Company, (1973).

Sanchez, G. I. Scores of Spanish Speaking Children on Reported Tests. Journal of Genetic Psychology, 40, 223-231, 1932.

Sandoval J. & Mile, W. Accuracy of Judgments of WISC-R Item Difficulty for Minority Groups, Journal of Consulting Clinical Psychology. V. 48, No. 2, 249-253, 1980.

Sandoval, J. The WISC-R and Internal Evidence of Test Bias with Minority Groups. Journal of Consulting and Clinical Psychology, 47, (5), 919 -926, 1979.

Sans, M. C. The Wechsler Intelligence Scale and its Adaptation in Argentinian Children, Acta Psiqui at Psico Amer Lat, 30, 292-296, 1984.

Sataz, P. & Mogel, S. An abbreviation of the WAIS for Clinical use. Journal of Clinical Psychology 18, 77-79, 1962.

Sataz, P. and others, An Abbreviation of the WISC for Clinical use. Journal of Consulting Psychology 31, 108, 1967.

Sattler, A. S. Assessment of Children's Intelligence and Special Abilities (2ed.) Boston: Allyn and Bacon, (1982).

Sattler, Jerome M. Relationship Between PPVT and WISC-R in Children with Reading Disabilities. Psychology in Schools, 17, (1) 331-334, 1980.

Sattler, J. <u>Assessment of Children's Intelligence.</u> Philadilphia: Saunders, (1974).

Schmida, L. C. & Keenum, D. G. <u>Education in the Middle East Mideast Education & Training Services.</u> Inc. Washington D. C., (1983).

Schwarz, P.A. & Krug,R.E. Ability Testing in Developing Countries A Handbook of Principles and Techniques. New York Praeger, (1972).

Scriven, M. "The Values of the Academy Moral Issues for American Educational Research Arising from the Jensen Case". Review of Educational Research, 40, 541-549, 1970.

Seashore, H. G. The Standardization of the Wechsler Intelligence Scale for Children. Journal of Consulting Psychology, 14, 99-110, 1950.

Sechrest, L. , Fay, T. L. & Zaidi, S. M. Problem of Translation in Cross - Cultural Research. Journal of Cross-Cultural Psychology, 3, 45- 56, 1972.

Shaw, M. C. & White, D. L. The relationship between Child - Parent Indemnification and academic Underachievement. Journal of Clinical Psychology, 21, 10-13,, 1965.

Shoer, R. E. Description of the San Rafeal Program for More able Learners as Prescribed in the California Administrative Cooed Article 23, Specially Educational Programs for Mentally Gifted Minors, (1966).

Shuey, Audrey. Social-Psychological Characteris-tics of Achieving Black Children. The Negro Educational Review 29, 80-86, 1958.

Siegler, R. & Powell, J. The Development of time, Speed, and Distance Concepts. Developmental Psycho logy, 15 (3), 288-298, 1979.

Siegler, R. & Richards, D. The Development of Intelligence In R. Stevenbery. The Handbook of Human Intelligence. New York: Cambrige, (1982).

Silverstien, A. B. Validity of a New Approach to the Design of the WAIS,WISC and WPPSI Short Forms. Journal of Consulting and Clinical Psychology, 32, 478-479, 1968.

Silverstien, A. B. Validity of the Sataz, Mogel, Yudin-Type Short Form. Journal of Consulting and Clinical Psychology, 54, 20-21, 1982.

Silverstien, A. B. Validity of the WISC-R Short Form. Journal of Clinical Psychology, 31, 696-697, (1975)

Smith, J. R. Sex Differences in Cognitive Style and Mathematical Performance. Thesis for M. Ed. University of Hull, (1984).

Solly, D. C. Comparison of WISC and WISC-R Scores of Mentally Retarded and Gifted Children. Journal of School Psychology, 15, 255-258, 1977.

Stenberg, R. I. & Powell, I. S. Comprehending Verbal Comprehension. American Psychologist, 38, 878- 893, 1983.

Stenberg, R. J. The Nature of Intelligence. New York University Education Quarterly, 12, (3), 10- 17, 1981.

Sternbery, R. & Richards, D. The Development of Intelligence. In Psychology Mussen (ed). Handbook of Child Psychology V. III, New York: John Wiley, (1983).

Swaab, D. F. & Fliers, E. Sex Differences found in Cell number. Science, Vol. 126, May 31, 1985.

Swerdlike, M.E. A Comparison of the WISC and WISC-R for Children referred to School Psychologists Michigan State University, (1976).

Swift, D. F. Social Mobility Ideology and 11 + success. British Journal of Sociology, 18,165-186,1967

Syfort, J., Srcen, F. & Lahmer, A. Human Services for Cultural Minorities. Baltimore: University Press, (1980).

Teeter, A. P. The use of Generalizability Theory with Behavioural Observation. Paper Presented at the Annual Meeting of the American Educational Research Association. New York, March, 1982.

Teeter, A., Moor, C. & Peterson, J. WISC-R Verbal and Performance abilities of Native American Students referred for School learning Problem. Psychology in the Schools, 19, 39-44, 1982.

Terman, L. M. The Measurement of Intelligence. Houghton (1921).

Terman, L. M. & Oden, M. The Gifted Group at MidLife. Stanford,CA: Stanford University Press (1959).

Terman, L. M. & Chamberlein, R. The Vocabulary test as a measurement of Intelligence. Journal of Education Psychology, 9, 452-466, 1918.

Thorndike, R. L. & Hagegen, E.P. Measurement and Evalua tion in Psychology and Education. New York: Wiley. 3ed edition, (1920).

Thorndike, E. L. The Measurement of Intelligence: Present Status.Psychological Review, 1924,31, 219-252.

Thorndike, E.L. Intelligence examinations for College Entrance. Journal of Education Research, 1, 329-337, 1920.

Thurstone, L. L. "Factorial Studies of Intelligence, Psychometr. Monogr., No. 2, (1931).

Thurstone, L.L. A Cycle - Omnibus Intelligence test for College Students. Journal of Education Researchs, 4, 265-278, 1921.

Thurstone, T.G. "Primary Mental Abilities of children".Educational Psychology Measurement.1, 105-116, 1938.

Torrance, G. P. Identifying the strength of culturally Different Children. In C. R. Reynold & T. B. Gutkings (Eds.). (1982).

Trimble,J.E.,Lonner, W. G. & Boucher, J. D. Stalking the Wiley EMIC:Alternative to Cross - Cultural Measurement In S. H. Irvine, and J. W. Berry Human Assessment and Cultural Factor (pp.259-273) New York: Plenam Press 1983 .

Tuma, J., Applebaum, A.& Bee, D. "Compatibility of the WISC and WISC-R in Normal Children of Divergent Socioeconomic Backgrounds"Psychology in Schools 15, 3, 344, July 1978.

Tyler, L. The Psychology of Human Differences. New York Appleeton-Centre Crofts, (1965).

Valentine, Charles. Culture and Poverty. Chicago: The University of Chicago Press, (1968).

Vance, H., Gaynor, P. & Coleman, M. Item Analysis of the Wechsler Intelligence Scale for Children Revised. Psychology in the School, 1977, 14, 132-139.

Vandergriff, A. & Wheaton, P. Comparison of WISC and WISC-R Scores of Highly Gifted Students in Public Schools. Psychological Reports, 43, 2, 1978.

Vandenberg, Steven G. Genetics of Sex Differences and heredity. The Encyclopedia of Education, 1972.

Wandemburg, N. L. Wechsler Intelligence Scale for Children: Arrangement Prior to the Adaptation of the Scale in Ecuador.Disciplinas Conexs 14, 53, 89-105, 1966.

Wechsler, D. Wechsler Intelligence Scale for Children Revised. British Amendments, NFER - Nelson, (1974).

Wechsler, D. The Measurement and Appraisal of Adult Intelligence (4th Ed.) Baltimore: Williams & Wilkins, (1958).

Wechsler, D. Manual for the Wechsler Intelligence Scale for Children.New York: the Psychological Corp, (1949).

Wechsler, D. The Measurement of Adult Intelligence. Baltimore: Williams & Wilkins, (1939).

Wechsler, D. The Measurement of Adult Intelligence. Baltimore: The Williams and Wilkins Co., (1944).

Weinberg, Henry. Probation Program Planning for Youthful Offenders. Intelligence, 106: 58-61, Aug. 1977.

Weiner, S. & Kaufman, Alan. WISC-R Versus WISC for Black Children Suspected of Learning or Behavioral Disabilities.Journal of Learning Disabilities. 12, 2, 103, February, 1979.

Wells, F. L. & Martin, H. A. A Method of Memory Examination Suitable for Psychotic Cases. American Journal of Psychiatric, 3,243-257, 1923.

Whipple, G. M. The National Intelligence Tests. Journal of Education Research, 4, 16-31, 1921.

Whipple, G. M. Arrange of Information Tests. Psychological Review, 16, 347-351, 1909.

Wikoff, A. S. & Parolini, K. Abbreviation of WISC and WISC-R Scores of Highly Students in Public Schools. Psychological Reports, 43, 2, 1978.

Williams, R. L. The BITCH-100: A culture-Specific Test. A Paper Presented at the Meeting of the American Psychology Association, Honolulu, 1972.

Williams, R. L. Abuses and Misuses in Testing Black chiChildren. Consulting Psychologist 2 (3), 62-73,1971

Wilson, M. L. Rural Alaska WISC-R Norms. Report, 1982, University of Alaska.

Wilson, S. R. Twins: Early Mental Development. Since,75, 1972, 414-417.

Wolfson, W. & Bachellis L. An Abbreviated Form of the WAIS Verbal Scale. Journal of Clinical Psychology, 16, 421, 1960.

Woodworth, R.S. & Wells, F.L. Association Tests. Psychology Review Monogr., No. 57, 1911.

Yoakum, E. A. & Yerkes, R. M. Army Mental Tests. New York: Holt (1920).

Young, F. W. Scaling Annual Review of Psychology, 35, 55-81, 1984.

Yudin, L.W. An Abbreviated Form of the WISC for use with Emotionally Disturbed Children. Journal of Consulting Psychology, 30, 272-275, 1966.

Yule, W., Berger, M., Rutter, M. & Yule, B. Children of West Indian Immigrants. Intellectual Performance and Reading Attainment. Journal of Child Psychology and Psychiatry, 16, 1-17, 1975.

Zajonc, R. B. & Others. Annual Report. University of Michigan in Arbor. (1984)

## Appendix (A)

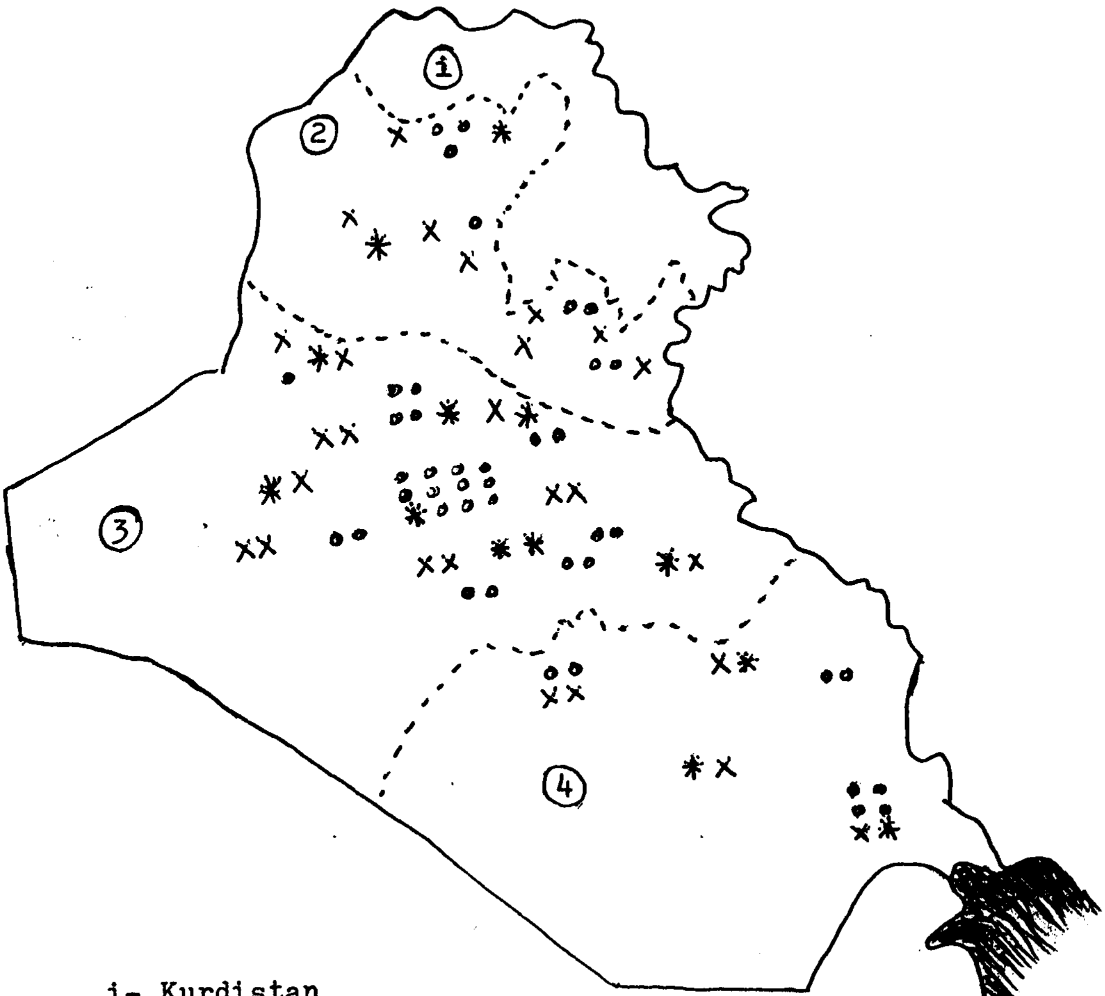
**The New Vocabulary Test.**

Sight
Hard
Heart
Uncover
Universal
Ulcer
Ram
Whip
Wheat
Inadvertent
Harm
Prohibited
Response
Conform
Cut-Off
Moist
Earnest
Ice
Amass
Importunity
Heap-up
Postpone
Noun
Rope
Collect
Support
Solemnity
Wean
Head
Camel
Unconscious
Kneel
Assassin
Bow
Body
Drought
Spear
Nightingale
Superior
Squander

Appendix (B)

i- Map of Iraq and the Three Geographical Regions, North, Middle, and South and the Autonomous Region for Kurdistan .

2- Location of the Schools Sample.




i- Kurdistan

2- North Iraq

3- Middle =

4- South =

●- Large Schools

X- Small Schools

✱- Mixed Schools

APPENDIX (C)

Scaled Score Equivalents of Raw Scores
of Boys and Girls from Main Study

yrs.12 mos. 0 days 0
through
yrs.12 mos. 3 days 30

| Scaled Score | Verbal: Information | Similarities | Arithmetic | Vocabulary | Coprehension | Scaled Score | Performance: Picture Comple. | Picture Arrang. | Block Design | Object Assembly | Coding | Scaled Score |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0-6 | 0-2 | 0-4 | 0-12 | 0-6 | 1 | 0-6 | 0-3 | 0-5 | 0-5 | 0-18 | 1 |
| 2 | 7 | 3 | 5 | 13-15 | 7 | 2 | 7 | 4-5 | 6 | 6-7 | 19-21 | 2 |
| 3 | 8 | 4 | 6 | 16-19 | 8 | 3 | 8 | 6-7 | 7-8 | 8-9 | 22-24 | 3 |
| 4 | 9 | 5-6 | 7 | 20-21 | 9 | 4 | 9 | 8-9 | 9-10 | 10 | 25-28 | 4 |
| 5 | 10 | 7 | 8 | 22-23 | 10-11 | 5 | 10 | 10-13 | 11-13 | 11-12 | 29-31 | 5 |
| 6 | 11-12 | 8-9 | 9 | 24-25 | 12 | 6 | 11-12 | 14-15 | 14-16 | 13 | 32-34 | 6 |
| 7 | 13 | 10-11 | 10 | 26-27 | 13-14 | 7 | 13 | 16-17 | 17-19 | 14-15 | 35-37 | 7 |
| 8 | 14 | 12 | 11 | 28-30 | 15-16 | 8 | 14 | 18-20 | 20-22 | 16-17 | 38-40 | 8 |
| 9 | 15 | 13 | 12 | 31-32 | 17-18 | 9 | 15 | 21-23 | 23-25 | 18 | 41-44 | 9 |
| 10 | 16-17 | 14-15 | 43 | 33-35 | 19-20 | 10 | 16 | 24-25 | 26-29 | 19 | 45-47 | 10 |
| 11 | 18 | 16-17 | 14 | 36-38 | 21-22 | 11 | 17 | 26-28 | 30-33 | 20-21 | 28-50 | 11 |
| 12 | 19 | 18 | 15 | 34-41 | 23-24 | 12 | 18 | 29-31 | 34-37 | 22 | 51-53 | 12 |
| 13 | 20-21 | 19-20 | 15 | 42-44 | 25 | 13 | 19 | 32-35 | 38-41 | 23-24 | 54-56 | 13 |
| 14 | 22-23 | 21-22 | 15 | 45-46 | 26-27 | 14 | 20-21 | 36-37 | 42-45 | 25 | 57-58 | 14 |
| 15 | 24-25 | 23 | 16 | 47-48 | 28 | 15 | 22 | 38-39 | 46-49 | 26 | 59-60 | 15 |
| 16 | 26 | 24-25 | 17 | 49-51 | 29-30 | 16 | 23 | 40-41 | 50-52 | 27-28 | 61-62 | 16 |
| 17 | 27-28 | 26 | 18 | 52-54 | 31 | 17 | 24 | 42 | 53-56 | 29-30 | 63-64 | 17 |
| 18 | 28 | 27 | 18 | 55-57 | 32 | 18 | 25 | 43-44 | 55-57 | 31 | 65-66 | 18 |
| 19 | 29-30 | 28 | 18 | 58-60 | 33-34 | 19 | 26 | 45-48 | 58-62 | 32-33 | 67-75 | 19 |

APPENDIX(C)

Scaled Score Equivalents of Raw Scores
of Boys and Girls from Main Study

12 yrs. 4 mos. 0 days
through
12 yrs. 7 mos. 30 days

| Scaled Score | Verbal: Information | Verbal: Similarities | Verbal: Arithmetic | Verbal: Vocabulary | Verbal: Comprehension | Scaled Score | Performance: Picture Comple. | Performance: Picture Arrang. | Performance: Block Design | Performance: Object Assembly | Performance: Coding | Scaled Score |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0- 6 | 0 - 2 | 0 - 4 | 0 - 12 | 0 - 6 | 1 | 0 - 6 | 0 - 3 | 0 - 6 | 0 - 5 | 0 - 18 | 1 |
| 2 | 7 | 3 | 5 | 13 - 15 | 7 | 2 | 7 | 4 - 6 | 7 - 8 | 6 - 7 | 19- 21 | 2 |
| 3 | 8 | 4 | 6 | 16 - 19 | 8 - 9 | 3 | 8 | 7 - 8 | 9 -10 | 8 - 9 | 22- 24 | 3 |
| 4 | 9 | 5 - 6 | 7 | 20 - 22 | 10 | 4 | 9 | 9 -11 | 11 -12 | 10 | 25- 27 | 4 |
| 5 | 10 | 7 - 8 | 8 | 23 - 25 | 11 | 5 | 10 | 12 -14 | 13 -15 | 11 -12 | 28- 30 | 5 |
| 6 | 11-12 | 8 - 9 | 9 | 26 - 27 | 12 - 13 | 6 | 11 -12 | 15 -16 | 16 -17 | 13 | 31- 33 | 6 |
| 7 | 13 | 10 - 11 | 10 | 28 - 29 | 14 | 7 | 13 | 17 -18 | 18 -20 | 14 -15 | 34- 36 | 7 |
| 8 | 14 | 12 | 11 | 30 - 31 | 15 - 16 | 8 | 14 | 19 -20 | 21 -23 | 16 -17 | 37 -39 | 8 |
| 9 | 15 | 13 | 12 | 32 - 33 | 17 - 18 | 9 | 15 | 21 -23 | 24 -26 | 18 | 40- 43 | 9 |
| 10 | 16- 17 | 14 - 15 | 13 | 34 - 35 | 19 - 20 | 10 | 16 | 24 -26 | 27 -30 | 19 | 44 -46 | 10 |
| 111 | 18 | 16 - 17 | 13 | 36 - 38 | 21 - 22 | 11 | 17 | 27 -29 | 31 -34 | 20 - 21 | 47 -49 | 11 |
| 12 | 19 | 18 | 14 | 39 - 41 | 23 - 24 | 12 | 18 | 30 -32 | 35 -38 | 22 | 50 -52 | 12 |
| 13 | 20- 21 | 19- 20 | 15 | 42 - 44 | 25 | 13 | 19 | 33 -35 | 39 -42 | 23 - 24 | 53 -555 | 13 |
| 14 | 22- 23 | 21- 23 | 15 | 45 - 47 | 26 - 27 | 14 | 20-21 | 36 -37 | 43 -46 | 25 | 56 - 57 | 14 |
| 15 | 24- 25 | 24- 25 | 16 | 48 - 49 | 28 | 15 | 22 | 38 -39 | 47 -50 | 26 | 58 - 59 | 15 |
| 16 | 26 | 26 | 17 | 50 - 51 | 29 - 30 | 16 | 23 | 40 -41 | 51 -52 | 27 - 28 | 60 - 61 | 16 |
| 17 | 27 | 27 | 17 | 52 - 54 | 31 | 17 | 24 | 42 | 53 -54 | 29 - 30 | 62 - 63 | 17 |
| 18 | 28 | 28 | 18 | 55 - 57 | 32 | 18 | 25 | 43 -44 | 55 -57 | 31 | 64 - 65 | 18 |
| 19 | 29- 30 | 29- 30 | 18 | 58 - 64 | 33 - 34 | 19 | 26 | 45 -48 | 58 -62 | 32 - 33 | 66 - 75 | 19 |

## Scaled Score Equivalents of Raw Scores

12 yrs. 8 mos. 0 days
through
12 yrs. 11 mos. 30 days

| Scaled Score | Verbal: Information | Verbal: Similarities | Verbal: Arithmetic | Verbal: Vocabulary | Verbal: Comprehension | Scaled Score | Performance: Picture Completion | Performance: Picture Arrangement | Performance: Block Design | Performance: Object Assembly | Performance: Coding | Scaled Score |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 - 6 | 0 - 3 | 0 - 5 | 0 - 12 | 0 - 6 | 1 | 0 - 6 | 0 - 3 | 0 - 6 | 0 - 5 | 0 - 18 | 1 |
| 2 | 7 | 4 | 6 | 13 - 16 | 7 | 2 | 7 - 8 | 4 - 6 | 7 - 9 | 6 - 7 | 19 - 21 | 2 |
| 3 | 8 | 5 | 7 | 17 - 19 | 8 - 9 | 3 | 9 | 7 - 8 | 10 - 11 | 8 - 9 | 22 - 24 | 3 |
| 4 | 9 | 6 | 8 | 20 - 23 | 10 | 4 | 10 | 9 - 10 | 12 - 14 | 10 | 25 - 27 | 4 |
| 5 | 10 | 7 - 8 | 9 | 24 - 26 | 11 | 5 | 11 | 11 - 14 | 15 - 16 | 11 - 12 | 28 - 30 | 5 |
| 6 | 11 - 12 | 9 | 10 | 27 - 28 | 12 - 13 | 6 | 12 - 13 | 15 - 16 | 17 - 19 | 13 | 31 - 33 | 6 |
| 7 | 13 | 10 - 11 | 11 | 29 - 30 | 14 | 7 | 14 | 17 - 18 | 20 - 21 | 14 - 15 | 34 - 36 | 7 |
| 8 | 14 | 12 | 12 | 31 - 32 | 15 - 16 | 8 | 15 | 19 - 20 | 22 - 24 | 16 - 17 | 37 - 39 | 8 |
| 9 | 15 | 13 | 13 | 33 - 34 | 17 - 18 | 9 | 16 | 21 - 23 | 25 - 27 | 18 | 40 - 43 | 9 |
| 10 | 16 - 17 | 14 - 15 | 14 | 35 - 36 | 19 - 20 | 10 | 17 | 24 - 26 | 28 - 30 | 19 - 20 | 44 - 46 | 10 |
| 11 | 18 | 16 - 17 | 14 | 37 - 38 | 21 - 22 | 11 | 18 | 27 - 29 | 31 - 34 | 21 | 47 - 50 | 11 |
| 12 | 19 | 18 - 19 | 14 | 39 - 41 | 23 - 24 | 12 | 19 | 30 - 32 | 35 - 38 | 22 | 51 - 53 | 12 |
| 13 | 20 - 21 | 20 | 15 | 42 - 44 | 25 - 26 | 13 | 20 | 33 - 35 | 39 - 42 | 23 - 24 | 54 - 56 | 13 |
| 14 | 22 - 23 | 21 - 22 | 15 | 45 - 47 | 27 | 14 | 21 - 22 | 36 - 37 | 43 - 46 | 25 | 57 - 58 | 14 |
| 15 | 24 - 25 | 23 | 16 | 48 - 49 | 28 - 29 | 15 | 23 | 38 - 39 | 47 - 50 | 26 - 27 | 59 - 60 | 15 |
| 16 | 26 | 24 - 25 | 17 | 50 - 52 | 30 - 31 | 16 | 24 | 40 - 41 | 51 - 53 | 28 | 61 - 62 | 16 |
| 17 | 27 | 26 - 27 | 18 | 53 - 54 | 32 | 17 | 25 | 42 | 54 - 55 | 29 - 30 | 63 - 64 | 17 |
| 18 | 28 | 28 | 18 | 55 - 58 | 33 | 18 | -- | 43 | 56 - 57 | 31 | 65 - 66 | 18 |
| 19 | 29 - 30 | 29 - 30 | 18 | 59 - 64 | 34 | 19 | 26 | 45 - 48 | 58 - 62 | 32 - 33 | 67 - 75 | 19 |

APPENDIX (C) Continued

Scaled Score Equivalents of Raw Scores

13 yrs. 0 mos. 0 days through 13 yrs. 3 mos. 30 days

of Boys and Girls from Main Study

| Scaled Score | Verbal: Information | Similarities | Arithmetic | Vocabulary | Comprehension | Scaled Score | Performance: Picture Compl. | Picture Arrang. | Block Design | Object Assembly | Coding | Scaled Score |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 - 7 | 0 - 3 | 0 - 5 | 0 - 13 | 0 - 6 | 1 | 0 - 7 | 0 - 4 | 0 - 7 | 0 - 6 | 0 - 19 | 1 |
| 2 | 8 | 4 | 6 | 14 - 17 | 7 - 8 | 2 | 8 | 5 - 7 | 8 - 10 | 7 - 8 | 20 - 22 | 2 |
| 3 | 9 | 5 | 7 | 18 - 20 | 9 | 3 | 99 | 8 - 9 | 11 - 12 | 9 - 10 | 23 - 25 | 3 |
| 4 | 10 | 6 | 8 | 21 - 24 | 10 - 11 | 4 | 10 | 10 - 12 | 13 - 15 | 11 | 26 - 28 | 4 |
| 5 | 11 | 7 - 8 | 9 | 25 - 27 | 12 | 5 | 11 | 13 - 14 | 16 - 17 | 12 - 13 | 29 - 30 | 5 |
| 6 | 12 | 9 | 10 | 28 - 29 | 13 | 6 | 12 - 13 | 15 - 17 | 18 - 20 | 14 | 31 - 34 | 6 |
| 7 | 13 - 14 | 10 - 11 | 11 | 30 - 31 | 14 - 15 | 7 | 14 | 18 - 19 | 21 - 22 | 15 - 16 | 35 - 37 | 7 |
| 8 | 15 | 12 - 13 | 12 | 32 - 34 | 16 - 17 | 8 | 15 - 16 | 20 - 21 | 23 - 25 | 17 - 18 | 38 - 40 | 8 |
| 9 10 | 16 | 14 | 13 | 35 - 36 | 18 | 9 | 17 | 22 - 24 | 26 - 28 | 19 - 20 | 41 - 44 | 9 |
| 10 | 17 | 15 - 16 | 14 | 37 - 38 | 19 - 20 | 10 | 18 | 25 - 26 | 29 - 31 | 21 | 45 - 47 | 10 |
| 11 | 18 | 17 - 18 | 14 | 39 - 40 | 21 - 22 | 11 | 19 | 27 - 29 | 32 - 34 | 22 | 48 - 51 | 11 |
| 12 | 19 | 19 - 20 | 14 | 41 - 42 | 23 - 24 | 12 | 20 | 30 - 32 | 35 - 38 | 23 | 52 - 54 | 12 |
| 13 | 20 - 21 | 21 - 22 | 15 | 43 - 44 | 25 - 26 | 13 | 21 | 33 - 34 | 39 - 42 | 24 - 25 | 55 - 57 | 13 |
| 14 | 22 - 23 | 23 | 15 | 45 - 47 | 27 | 14 | 22 | 36 - 38 | 43 - 46 | 26 | 58 - 59 | 14 |
| 15 | 24 - 25 | 24 | 16 | 48 - 50 | 28 - 29 | 15 | 23 | 39 - 40 | 47 - 50 | 27 - 28 | 60 - 61 | 15 |
| 16 | 26 - 27 | 25 | 17 | 51 - 53 | 30 - 31 | 16 | 24 | 41 | 51 - 53 | 29 | 62 - 63 | 16 |
| 17 | 28 | 26 - 27 | 18 | 54 - 55 | 32 | 17 | 25 | 42 | 54 - 56 | 30 - 31 | 64 - 65 | 17 |
| 18 | 29 | 28 - 29 | 18 | 56 - 58 | 33 | 18 | - | 43 - 44 | 57 - 60 | 32 | 66 - 67 | 18 |
| 19 | 30 | 30 | 18 | 59 - 64 | 34 | 19 | 26 | 45 - 48 | 61 - 62 | 33 | 68 - 75 | 19 |

APPENDIX (C) Continued

Scaled Score Equivalents of Raw Scores

of Boys and Girls from Main Study

13 yrs. 4 mos. 0 days
through
13 yrs. 7 mos. 30 days

| Scaled Score | Verbal | | | | | Scaled Score | Performance | | | | | Scaled Score |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Information | Similarities | Arithmetic | Vocabulary | Comprehension | | Picture Comple. | Picture Arrang. | Block Design | Object Assembly | Coding | |
| 1 | 0 – 7 | 0 – 3 | 0 – 5 | 0 – 13 | 0 – 7 | 1 | 0 – 7 | 0 – 5 | 0 – 7 | 0 – 6 | 0 – 20 | 1 |
| 2 | 8 | 4 | 6 | 14 – 17 | 8 – 9 | 2 | 8 | 6 – 8 | 8 – 10 | 7 – 8 | 21 – 24 | 2 |
| 3 | 9 | 5 | 7 | 18 – 20 | 10 | 3 | 9 | 9 – 10 | 11 – 12 | 9 – 10 | 25 – 26 | 3 |
| 4 | 10 | 6 – 7 | 8 | 21 – 24 | 11 | 4 | 10 – 11 | 11 – 13 | 13 – 15 | 11 | 27 – 29 | 4 |
| 5 | 11 | 8 | 9 | 25 – 27 | 12 – 13 | 5 | 12 | 14 – 15 | 16 – 17 | 12 – 13 | 30 – 32 | 5 |
| 6 | 12 | 9 | 10 | 28 – 30 | 14 | 6 | 13 | 16 – 18 | 18 – 20 | 14 | 33 – 35 | 6 |
| 7 | 13 – 14 | 10 – 11 | 11 | 31 – 32 | 15 – 16 | 7 | 14 | 19 – 20 | 21 – 22 | 15 – 16 | 36 – 38 | 7 |
| 8 | 15 | 12 – 13 | 12 | 33 – 35 | 17 – 18 | 8 | 15 – 16 | 21 – 22 | 23 – 25 | 17 – 18 | 39 – 41 | 8 |
| 9 | 16 | 14 – 15 | 13 | 36 – 38 | 19 | 9 | 17 – 18 | 23 – 25 | 26 – 28 | 19 – 20 | 42 – 44 | 9 |
| 10 | 17 | 16 | 14 | 39 – 40 | 20 – 21 | 10 | 19 | 26 – 27 | 29 – 31 | 21 – 22 | 45 – 47 | 10 |
| 11 | 18 | 17 – 18 | 14 | 41 – 42 | 22 – 23 | 11 | 20 | 28 – 30 | 32 – 34 | 23 | 48 – 51 | 11 |
| 12 | 19 | 19 – 20 | 15 | 43 – 44 | 24 – 25 | 12 | 21 | 31 – 33 | 35 – 38 | 24 | 52 – 54 | 12 |
| 13 | 20 – 21 | 21 – 22 | 16 | 45 – 46 | 26 | 13 | 22 | 34 – 35 | 39 – 42 | 25 | 55 – 57 | 13 |
| 14 | 22 – 23 | 23 | 16 | 47 – 48 | 27 – 28 | 14 | 23 | 36 – 38 | 43 – 46 | 26 | 58 – 59 | 14 |
| 15 | 24 – 25 | 24 | 17 | 49 – 51 | 29 – 30 | 15 | 23 | 39 – 40 | 47 – 50 | 27 – 28 | 60 – 62 | 15 |
| 16 | 26 – 27 | 25 | 18 | 52 – 53 | 31 | 16 | 24 | 41 – 42 | 51 – 53 | 29 | 63 – 64 | 16 |
| 17 | 28 | 26 – 27 | 18 | 54 – 56 | 32 | 17 | 25 | 43 | 54 – 56 | 30 – 31 | 65 – 66 | 17 |
| 18 | 29 | 27 – 28 | 18 | 57 – 58 | 33 | 18 | 26 | 44 | 57 – 60 | 32 | 67 – 68 | 18 |
| 19 | 30 | 30 30 | 18 | 59 – 64 | 34 | 19 | 26 | 45 – 48 | 61 – 62 | 33 | 69 – 75 | 19 |

APPENDIX (C)

Scaled Score Equivalents of Raw Scores

of Boys and Girls from Main Study

yrs. 13 mos. 8 days 0
through
yrs. 13 mos. 11 days 30

| Scaled Score | Verbal: Information | Verbal: Similarities | Verbal: Arithmetic | Verbal: Vocabulary | Verbal: Comprehension | Scaled Score | Performance: Picture Comple. | Performance: Picture Arrang. | Performance: Block Design | Performance: Object Assembly | Performance: Coding | Scaled Score |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0-8 | 0-3 | 0-6 | 1-14 | 0-7 | 1 | 0-7 | 0-5 | 0-7 | 0-7 | 0-21 | 1 |
| 2 | 9 | 4 | 7 | 15-17 | 8-9 | 2 | 8 | 6-8 | 8-10 | 8-9 | 22-25 | 2 |
| 3 | 10 | 5 | 8 | 18-20 | 10-11 | 3 | 9 | 9-10 | 11-12 | 10-11 | 26-27 | 3 |
| 4 | 11 | 6-7 | 9 | 21-24 | 12 | 4 | 10-11 | 11-13 | 13-15 | 12 | 28-30 | 4 |
| 5 | 12 | 8 | 10 | 25-27 | 13-14 | 5 | 12 | 14-15 | 16-18 | 13-14 | 31-33 | 5 |
| 6 | 13 | 9 | 11 | 28-30 | 15 | 6 | 13 | 16-18 | 19-21 | 15 | 34-36 | 6 |
| 7 | 14-15 | 10 | 12 | 31-33 | 16-17 | 7 | 14 | 19-20 | 22-23 | 16-17 | 37-39 | 7 |
| 8 | 16 | 12-13 | 13 | 34-36 | 18-19 | 8 | 15-16 | 21-22 | 24-26 | 18-19 | 40-42 | 8 |
| 9 | 17 | 14-15 | 14 | 37-38 | 20 | 9 | 17-18 | 23-25 | 27-29 | 20 | 43-45 | 9 |
| 10 | 18 | 16 | 15 | 39-40 | 21-22 | 10 | 19 | 26-28 | 30-32 | 21-22 | 46-48 | 10 |
| 11 | 19 | 17-18 | 15 | 41-42 | 23-24 | 11 | 20 | 29-30 | 33-35 | 23 | 49-50 | 11 |
| 12 | 20 | 19-20 | 16 | 43-44 | 25-26 | 12 | 21 | 31-33 | 36-39 | 24 | 53-55 | 12 |
| 13 | 21 | 21-22 | 16 | 45-47 | 27 | 13 | 22 | 34-35 | 40-43 | 25-26 | 56-57 | 13 |
| 14 | 22-23 | 23 | 16 | 48-49 | 28 | 14 | 23 | 36-38 | 44-47 | 27 | 58-59 | 14 |
| 15 | 24 | 24 | 17 | 50-51 | 29-30 | 15 | 23 | 39-40 | 48-50 | 28 | 60-62 | 15 |
| 16 | 26-27 | 25 | 18 | 52-53 | 31 | 16 | 24 | 41-42 | 51-53 | 29-30 | 63-64 | 16 |
| 17 | 28 | 26-27 | 18 | 54-56 | 32 | 17 | 25 | 43 | 54-56 | 31 | 65-66 | 17 |
| 18 | 29 | 28-29 | 18 | 57-59 | 33 | 18 | 26 | 44 | 57-60 | 32 | 67-68 | 18 |
| 19 | 30 | 30 | | 60-64 | 34 | 19 | 26 | 45-48 | 61-62 | 33 | 69-75 | 19 |

*APPENDIX (C)*

Scaled Score Equivalents of Raw Scores

of Boys and Girls from Main Study

yrs. 14 mos. 0 days 0
through
yrs. mos. .. days

| Scaled Score | Verbal: Information | Similarities | Arithmetic | Vocabulary | Comprehension |
|---|---|---|---|---|---|
| 1 | 0-8 | 0-4 | 0-6 | 0-14 | 0-8 3 |
| 2 | 9 | 5 | 7 | 13-17 | 9-10 |
| 3 | 10 | 6 | 8 | 18-20 | 11-12 |
| 4 | 11 | 7 | 9 | 21-24 | 13-14 |
| 5 | 12 | 8-9 | 10 | 25-27 | 15 |
| 6 | 13 | 10 | 11 | 28-30 | 16 |
| 7 | 14 | 14-15 | 12 | 31-33 | 17-18 |
| 8 | 16 | 13-14 | 13 | 34-36 | 19-20 |
| 9 | 17 | 15-16 | 14 | 37-39 | 21 |
| 10 | 18 | 17-18 | 15 | 40-41 | 22 |
| 11 | 19 | 19 | 15 | 42-43 | 23-24 |
| 12 | 20 | 20-21 | 16 | 44-46 | 25-26 |
| 13 | 21-22 | 22 | 16 | 47-48 | 27 |
| 14 | 33 | 23-24 | 16 | 49-50 | 28 |
| 15 | 24-25 | 25 | 17 | 51-52 | 29-30 |
| 16 | 26-27 | 26-27 | 18 | 53-54 | 31 |
| 17 | 28 | 28 | 18 | 55-56 | 32 |
| 18 | 29 | 29 | 18 | 57-59 | 33 |
| 19 | 30 | 30 | 18 | 60-64 | 34 |

| Scaled Score | Performance: Picture Comple. | Picture Arrang. | Block Design | Object Assembly | Coding | Scaled Score |
|---|---|---|---|---|---|---|
| 1 | 0-8 | 0-5 | 0-7 | 0-7 | 0-21 | 1 |
| 2 | 9 | 6-8 | 8-10 | 8-9 | 22-25 | 2 |
| 3 | 10 | 9-11 | 11-12 | 10-11 | 26-27 | 3 |
| 4 | 11-12 | 12-13 | 13-15 | 12 | 28-31 | 4 |
| 5 | 13 | 14-15 | 16-18 | 13-14 | 32-34 | 5 |
| 6 | 14 | 16-18 | 19-21 | 15 | 35-37 | 6 |
| 7 | 15 | 19-20 | 22-24 | 16-17 | 38-40 | 7 |
| 8 | 16 | 21-22 | 25-27 | 18-19 | 41-44 | 8 |
| 9 | 17-18 | 23-26 | 28-30 | 20-21 | 44-46 | 9 |
| 10 | 19 | 27-29 | 31-33 | 22 | 47-49 | 10 |
| 11 | 20 | 30-31 | 34-36 | 23-24 | 50-53 | 11 |
| 12 | 21 | 32-33 | 37-40 | 25 | 54-56 | 12 |
| 13 | 22 | 34-35 | 41-44 | 26-27 | 57-58 | 13 |
| 14 | 23 | 36-38 | 45-48 | 28 | 59-60 | 14 |
| 15 | 23 | 39-41 | 49-51 | 29 | 61-62 | 15 |
| 16 | 24 | 42-43 | 52-54 | 30 | 63-64 | 16 |
| 17 | 25 | 44 | 55-57 | 31 | 65-66 | 17 |
| 18 | 26 | 45 | 58-60 | 32 | 67-68 | 18 |
| 19 | | 46-48 | 61-62 | 33 | 69-75 | 19 |

APPENDIX (C)

## Scaled Score Equivalents of Raw Scores

of Boys and Girls from Main Study

yrs. 14 mos. 4 days 0
through
yrs. 14 mos. 7 days 30

Verbal

| Scaled Score | Information | Similarities | Arithmetic | Vocabulary | Comprehension | Scaled Score |
|---|---|---|---|---|---|---|
| 1 | 0-8 | 0-4 | 0-6 | 0-14 | 0-8 | 1 |
| 2 | 9 | 5 | 7 | 15-18 | 9-10 | 2 |
| 3 | 10 | 6 | 8 | 19-21 | 11-12 | 3 |
| 4 | 11 | 7 | 9 | 22-24 | 13-14 | 4 |
| 5 | 12 | 8-9 | 10 | 25-27 | 15-16 | 5 |
| 6 | 13-14 | 10 | 11 | 28-30 | 17 | 6 |
| 7 | 15 | 11-12 | 12 | 31-33 | 18-19 | 7 |
| 8 | 16 | 13-14 | 13 | 34-36 | 20-21 | 8 |
| 9 | 17-18 | 15-16 | 14 | 37-39 | 22 | 9 |
| 10 | 19 | 17-18 | 15 | 40-41 | 23 | 10 |
| 11 | 20- | 19 | 15 | 42-43 | 24-25 | 11 |
| 12 | 21 | 20-21 | 16 | 44-46 | 26 | 12 |
| 13 | 22 | 22 | 16 | 47-48 | 27 | 13 |
| 14 | 23 | 23-24 | 17 | 49-51 | 28 | 14 |
| 15 | 24 | 25 | 17 | 52-53 | 29-30 | 15 |
| 16 | 25-26 | 26-27 | 18 | 54-55 | 31 | 16 |
| 17 | 27-28 | 28 | 18 | 56-57 | 32 | 17 |
| 18 | 29 | 29 | 18 | 58-59 | 33 | 18 |
| 19 | 30 | 30 | 18 | 60-64 | 34 | 19 |

Performance

| Picture Comple. | Picture Arrang. | Block Design | Object Assembly | Coding | Scaled Score |
|---|---|---|---|---|---|
| 0-8 | 0-6 | 0-7 | 0-7 | 0-22 | 1 |
| 9 | 7-9 | 8-10 | 8-9 | 23-25 | 2 |
| 10 | 10-12 | 11-12 | 10-11 | 26-28 | 3 |
| 11-12 | 13-14 | 13-15 | 12 | 29-32 | 4 |
| 13 | 15-16 | 16-18 | 13-14 | 33-35 | 5 |
| 14-15 | 17-18 | 19-21 | 15 | 36-38 | 6 |
| 16 | 19-20 | 22-24 | 16-17 | 39-41 | 7 |
| 17-18 | 21-23 | 25-27 | 18-19 | 42-44 | 8 |
| 19 | 24-26 | 28-30 | 20-21 | 45-47 | 9 |
| 20 | 27-29 | 31-33 | 22 | 48-50 | 10 |
| 21 | 30-31 | 34-37 | 23-24 | 51-53 | 11 |
| 22 | 32-33 | 38-41 | 25-26 | 54-56 | 12 |
| 23 | 34-36 | 42-45 | 27 | 57-58 | 13 |
| 23 | 37-38 | 46-49 | 28 | 59-60 | 14 |
| 24 | 39-41 | 50-52 | 29 | 61-62 | 15 |
| 24 | 42-43 | 53-55 | 30 | 63-64 | 16 |
| 25 | 44 | 56-58 | 31 | 65-66 | 17 |
| 26 | 45 | 59-60 | 32 | 67-68 | 18 |
| 26 | 46-48 | 61-62 | 33 | 69-75 | 19 |

APPENDIX (C)

## Scaled Score Equivalents of Raw Scores of Boys and Girls from Main Study

yrs. 14 mos. 8 days 0
through
yrs. 14 mos. 11 days 0

### Verbal

| Scaled Score | Information | Similarities | Arithmetic | Vocabulary | Comprehension | Scaled Score |
|---|---|---|---|---|---|---|
| 1 | 0-8 | 0-4 | 0-6 | 0-15 | 0-8 | 1 |
| 2 | 9 | 5 | 7 | 16-18 | 9-10 | 2 |
| 3 | 10 | 6 | 8 | 19-21 | 11-12 | 3 |
| 4 | 11 | 7-8 | 9 | 22-25 | 13-14 | 4 |
| 5 | 12 | 9 | 10 | 26-28 | 15-16 | 5 |
| 6 | 13-14 | 10-11 | 11 | 29-31 | 17-18 | 6 |
| 7 | 15 | 12 | 12 | 32-33 | 19-20 | 7 |
| 8 | 16 | 13-14 | 13 | 34-36 | 21-22 | 8 |
| 9 | 17-18 | 15-16 | 14 | 37-39 | 23 | 9 |
| 10 | 19 | 17-18 | 15 | 40-41 | 24 | 10 |
| 11 | 20 | 19-20 | 15 | 42-43 | 25 | 11 |
| 12 | 21 | 21 | 16 | 44-46 | 26 | 12 |
| 13 | 22 | 22 | 16 | 47-48 | 27-28 | 13 |
| 14 | 23 | 23-24 | 17 | 49-51 | 29 | 14 |
| 15 | 24 | 25 | 17 | 52-54 | 30 | 15 |
| 16 | 25-26 | 26-27 | 18 | 54-56 | 31 | 16 |
| 17 | 27-28 | 28 | 18 | 57-58 | 32 | 17 |
| 18 | 29-30 | 29 | 18 | 59-60 | 33 | 18 |
| 19 | 29-30 | 30 | 18 | 61-64 | 34 | 19 |

### Performance

| Picture Comple. | Picture Arrang. | Block Design | Object Assembly | Coding | Scaled Score |
|---|---|---|---|---|---|
| 0-9 | 0-6 | 0-8 | 0-8 | 0-22 | 1 |
| 10-11 | 7-9 | 9-11 | 9-10 | 23-25 | 2 |
| 12 | 10-12 | 12-13 | 11-12 | 26-28 | 3 |
| 13 | 13-14 | 14-16 | 13 | 29-32 | 4 |
| 14 | 15-16 | 17-19 | 14-15 | 33-35 | 5 |
| 15 | 17-18 | 20-22 | 16-17 | 36-38 | 6 |
| 16-17 | 19-21 | 23-25 | 18 | 39-41 | 7 |
| 18 | 22-24 | 26-28 | 19-20 | 42-45 | 8 |
| 19 | 25-27 | 29-31 | 21-22 | 46-48 | 9 |
| 20 | 28-30 | 32-34 | 23 | 49-50 | 10 |
| 21 | 31-32 | 35-38 | 24-25 | 51-53 | 11 |
| 22 | 33-34 | 39-42 | 26-27 | 54-56 | 12 |
| 23 | 35-36 | 43-45 | 28 | 57-59 | 13 |
| 23 | 37-38 | 46-49 | 29 | 60-61 | 14 |
| 24 | 39-41 | 50-52 | 30 | 62-63 | 15 |
| 25 | 42-43 | 53-55 | 31 | 64-65 | 16 |
| 25 | 44-45 | 56-59 | 32 | 65-67 | 17 |
| 26 | 46 | 60 | 32 | 68-69 | 18 |
| 26 | 47-48 | 61-62 | 33 | 70-75 | 19 |

# Scaled Score Equivalents of Raw Scores

of Boys and Girls from Main Study

15 yrs. 0 mos. 0 days
through
15 yrs. 3 mos. 30 days

## Verbal

| Scaled Score | Information | Similarities | Arithmetic | Vocabulary | Comprehension |
|---|---|---|---|---|---|
| 1 | 0-9 | 0-5 | 0-6 | 0-15 | 0-8 |
| 2 | 10 | 6 | 7 | 16-18 | 9-10 |
| 3 | 11 | 7 | 8 | 19-22 | 11-12 |
| 4 | 12 | 8-9 | 9 | 23-25 | 13-14 |
| 5 | 13 | 10 | 10 | 26-28 | 15-16 |
| 6 | 14 | 11-12 | 11 | 29-31 | 17-18 |
| 7 | 15-16 | 13-14 | 12 | 32-33 | 19-20 |
| 8 | 17 | 15 | 13 | 34-36 | 21-22 |
| 9 | 18 | 16 | 14 | 37-39 | 23-24 |
| 10 | 19 | 17-18 | 15 | 40-41 | 25 |
| 11 | 20-21 | 19-20 | 15 | 42-44 | 26 |
| 12 | 22 | 21-22 | 15 | 45-47 | 27 |
| 13 | 23 | 23 | 16 | 48-49 | 28 |
| 14 | 24 | 24 | 17 | 50-51 | 29 |
| 15 | 25-26 | 25-26 | 17 | 52-54 | 30 |
| 16 | 27 | 27 | 18 | 55-56 | 31 |
| 17 | 28 | 28 | 18 | 57-58 | 32 |
| 18 | 29-30 | 29 | 18 | 59-60 | 33 |
| 19 | 30 | 30 | 18 | 61-64 | 34 |

## Performance

| Scaled Score | Picture Compla. | Picture Arrang. | Block Design | Object Assembly | Coding | Scaled Score |
|---|---|---|---|---|---|---|
| 1 | 0-9 | 0-6 | 0-8 | 0-8 | 0-23 | 1 |
| 2 | 10-11 | 7-10 | 9-11 | 9-10 | 24-26 | 2 |
| 3 | 12-13 | 11-13 | 12-13 | 11-12 | 27-29 | 3 |
| 4 | 14 | 14-15 | 14-16 | 13 | 30-33 | 4 |
| 5 | 15 | 15-16 | 17-19 | 14-15 | 34-36 | 5 |
| 6 | 16 | 18-19 | 20-22 | 16-17 | 37-39 | 6 |
| 7 | 17 | 20-21 | 23-25 | 18 | 40-42 | 7 |
| 8 | 18 | 22-24 | 26-28 | 19-20 | 43-45 | 8 |
| 9 | 19 | 25-27 | 29-31 | 21-22 | 46-48 | 9 |
| 10 | 20 | 28-30 | 32-34 | 23 | 49-51 | 10 |
| 11 | 21 | 31-33 | 35-38 | 24-25 | 52-54 | 11 |
| 12 | 22 | 34-35 | 39-42 | 26-27 | 55-57 | 12 |
| 13 | 23 | 36-37 | 43-46 | 28 | 58-60 | 13 |
| 14 | 24 | 38-39 | 47-50 | 29 | 61-62 | 14 |
| 15 | 24 | 40-41 | 51-53 | 30 | 63-64 | 15 |
| 16 | 25 | 42-43 | 54-56 | 31 | 65-66 | 16 |
| 17 | 25 | 44-45 | 57-60 | 32 | 67-68 | 17 |
| 18 | 26 | 46 | 61 | 32 | 69-70 | 18 |
| 19 | 26 | 47-48 | 62 | 33 | 71-75 | 19 |

APPENDIX (C)

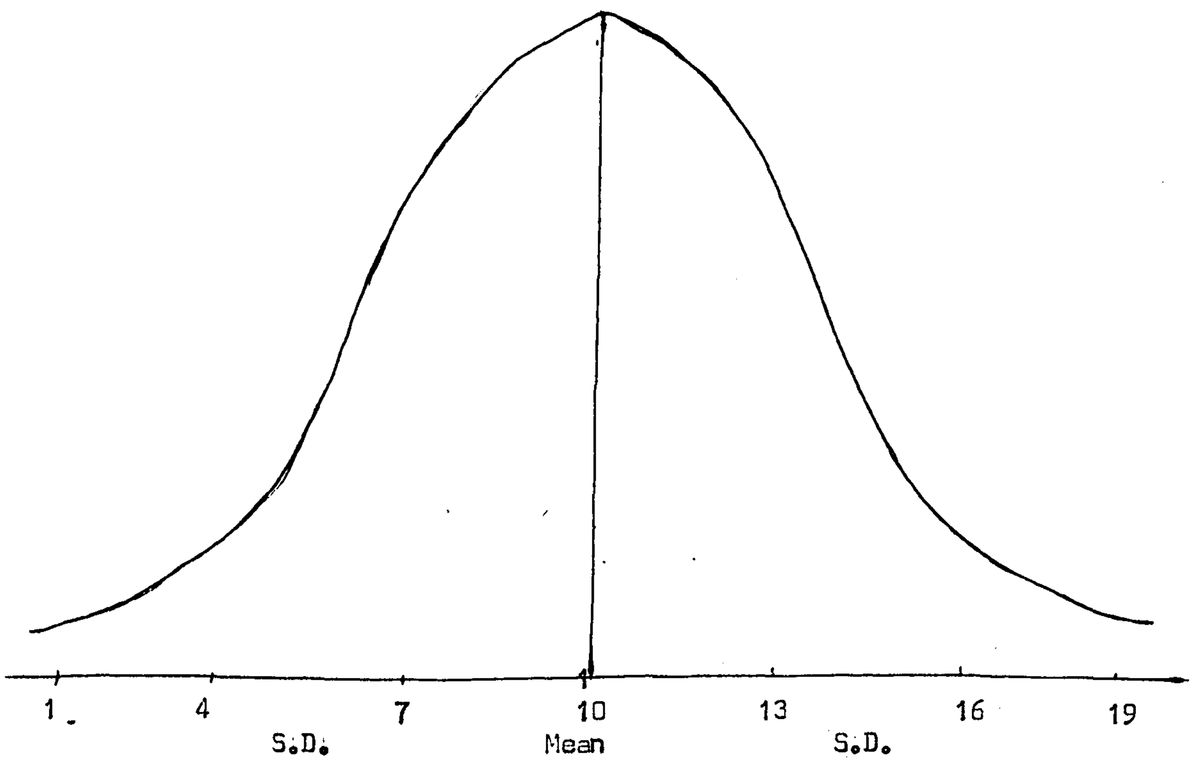
Scaled Score Equivalents of Raw Scores

of Boys and Girls from Main Study

15 yrs. 8 mos. 0 days
through
15 yrs. 11 mos. 30 days

| | Verbal | | | | | | Performance | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Scaled Score | Information | Similarities | Arithmetic | Vocabulary | Coprehension | Scaled Score | Picture Compla. | Picture Arrang. | Block Design | Object Assembly | Coding | Scaled Score |
| 1 | 0-9 | 0-5 | 0-7 | 0-16 | 0-9 | 1 | 0-10 | 0-7 | 0-8 | 0- 8 | 0-24 | 1 |
| 2 | 10 | 6 | 8 | 17-19 | 10-11 | 2 | 11-12 | 8-10 | 9-12 | 9-10 | 25-27 | 2 |
| 3 | 11 | 7-8 | 9 | 20-23 | 12-13 | 3 | 13 | 11-13 | 13-14 | 11-12 | 28-30 | 3 |
| 4 | 12 | 9 | 10 | 24-26 | 14-15 | 4 | 14-15 | 14-16 | 15-17 | 13-14 | 31-34 | 4 |
| 5 | 13 | 10-11 | 11 | 27-29 | 16-17 | 5 | 16 | 17-18 | 18-20 | 15-16 | 35-37 | 5 |
| 6 | 14 | 12 | 12 | 30-32 | 18-19 | 6 | 17 | 19-20 | 21-23 | 17-18 | 38-40 | 6 |
| 7 | 15-16 | 13-14 | 13 | 33-35 | 20 | 7 | 18 | 21-22 | 24-26 | 19 | 41-43 | 7 |
| 8 | 17 | 15 | 14 | 36-37 | 21-22 | 8 | 19 | 23-24 | 27-29 | 20-21 | 44-46 | 8 |
| 9 | 18 | 16 | 15 | 38-40 | 23-24 | 9 | 20 | 25-27 | 30-32 | 22-23 | 47-49 | 9 |
| 10 | 19 | 17-18 | 16 | 41-42 | 25 | 10 | 21 | 28-30 | 33-35 | 24 | 50-52 | 10 |
| 11 | 20-21 | 19-20 | 16 | 43-45 | 26 | 11 | 22 | 31-33 | 36-39 | 25-26 | 53 -55 | 11 |
| 12 | 22-23 | 21-22 | 16 | 46-47 | 27 | 12 | 23 | 34-35 | 40-42 | 27 | 56- 58 | 12 |
| 13 | 24 | 23 | 17 | 48-50 | 28 | 13 | 23 | 36-37 | 43-46 | 28 | 59-61 | 13 |
| 14 | 25 | 24 | 17 | 51-52 | 29 | 14 | 24 | 38-39 | 47-50 | 29 | 62-63 | 14 |
| 15 | 26 | 25-26 | 17 | 53-55 | 30 | 15 | 24 | 40-41 | 51-53 | 30 | 64-65 | 15 |
| 16 | 27 | 27 | 18 | 56-57 | 31 | 16 | 25 | 42-43 | 54-56 | 31 | 66-67 | 16 |
| 17 | 28-29 | 28 | 18 | 58-59 | 32 | 17 | 25 | 44-45 | 57-60 | 31 | 68-69 | 17 |
| 18 | 30 | 29 | 18 | 60 | 33 | 18 | 26 | 46 | 61 | 32 | 70-71 | 18 |
| 19 | 30 | 30 | 18 | 61-64 | 34 | 19 | 26 | 47-48 | 62 | 33 | 72-75 | 19 |

Appendix (D)




The Diagram Wechsler put for his Tests a Mean of 10 and S.D. of 3 in each side (Normal Distribution).

## Appendix(E)

**Schools in the North Region.**

**Urban Schools (Large Schools).**

1- Al-Zuhoure Secondary School for Girls.

2- Al-Tammim = = = = .

3- Al-Sharkia = = = = .

4- Kudija = = = = .

5- 1 June = = = Girls.

6- Abd Al-Rahman Al-Kaffiky Secondary School for Boys.

7- Sentiral Secondary School for Boys

8- Al-Markazia = = = =.

**Rural Schools (Small)**

9- Kara-Koush Intermediate School for Boys.

10- = = = = = Girls 11- Hawija = = = = .

12- = = = = = .

13- Al-Hamdania = = = = .

14- = = = = Boys.

15- Al-Zahra = Mixed = .

16- Al-Mansoure = = = .

17- Balttila = = for Boys.

18- Al-Yakda = = = = .

**Schools at the Middle Region**

**Urban Schools (Large Schools).**

19- Sallah Al-Deen Secondary School for Boys.

20- Al-Shab = = = = .

21- Al-Kudomia = = = = .

22- Al-Ahrrar = = = = .

23- Al-Kalliej = = = = .

24- 1 June Comprehensive Mixed School .

25- Al-Heala Secondary School =.

26- Al-Zahawy Secondary School for Boys .

27- Alwehda = = = = .

28- Bakuba = = = = .

29- Al-Tejara = = = = .

30- Al-Zahra = = = = .

31- Samera = = = = .

32- Al-Kutt = = = = .

33- Al Jamhuria = = = = .

34- Balkiess = = = Girls.

35- Al-Kahira = = = = .

36- Palastin = = = = .

37- Samara = = = = .

38- Wassitt = = = = .

39- Al-Kanssa = = = = .

40- Al-Urouba = = = = .

41- Al- Kutt = = = = .

42- Kadija = = = = .

43- Al-Zuhoure = = = = .

44- Al-Kadissia = = = = .

45- Al-Wassity = = = = .

**Rural Schools (Small Schools).**

46- Ballad Intermediate School for Girls.

47- Al-Tarrmia = = = = .

48- Al-Mahmoudia = = = = .

49- Al-Hamza = = = = .

50- Al-Rashdia = = = = .

51- Al-Jamhuria = = = = .

52- Wassit = = = = .

53- Dabeass = = = = .

54- Al-Massiab = = = = .

55- al-Mashkub = = = = .

56- Al-Azizia = = = = .

57- Al-Kalis = = = = .

58- Al-Jamhuria = = = = .

59- Al-Abassia = = = = .

60- Al-Shaheed = = = = .

61- Fatima = = = = .

62- Al-Ahrrar = Mixed School.

63- Al-Mansure Secondary Mixed School.

64- Al-Sallam = Mixed School.

**Schools in the South Region**

**Urban Schools (Large Schools).**

65- Al-Hussinia Secondary School for Girls.

66- Al-Shatra = = = = .

67- Souke Al-Shiouke = = = Boys .

68- Thee-Karr = = = = .

69- Al-Jamhouria = = = = .

70- Dar Al-Mallimat = = = Girls .

71- Al-Teajara = = = = .

72- Meesan = = = = .

**Rural Schools (Small Schools)**

73- Ali Al-Garrbi Intermediate School for Girls.

74- Al-Sallam = = = Boys.

75- Al-Antissar = = = = .

76- Al-Nassiria = = = Girls.

77- Al-Makal = = = Boys .

78- Al-Th-Thowra = = = Girls.

79- Aumar Al-Mouktar = Mixed School .

80- Fateama = = = .

81- Al-Shaheed = = = .

82- Al-Farabi = = = .

Appendix (F)

بسم الله الرحمن الرحيم
الجمهورية العراقية

وزارة التربيـــــة

مركز البحوث والدراسات التربوية

العدد / ٢٨٧٢٠

التاريخ / ٨ / ١١ / ١٩٨٧

الى / المديرية العامة لتربية مدينة بغداد / الكرخ

م / تسهيل مهمة طالب الدكتوراه

اشارة الى مذكرة المديرية العامة للعلاقات الثقافية المرقمة بلا في ٧/١١/١٩٨٧ •
يرجى التفضل بتسهيل مهمة السيد عدنان غائب راشد طالب الدكتوراه في المطالعـــة المتحدة للقيام بتطبيق (( اختبارات واسلوب لقياس ذكاء طلبة المرحلة المتوسطة والثانويــة في العراق )) كجزء من متطلبات الحصول على شهادة الدكتوراه في علم النفس •

مع التقديـــــر ...

محمود عبد الله احمد

ع / رئيس مركز البحوث والدراسات التربوية

نسخة منه الى /

المديرية العامة لتربية مدينة بغداد / الرصافة
المديرية العامة لتربية نينوى
المديرية العامة لتربية التأميم
المديرية العامة لتربية ذي قار
المديرية العامة لتربية الانبار
المديرية العامة لتربية اربيل
المديرية العامة لتربية القادسية

لنفس الغرض اعلاه /

مركز البحوث والدراسات التربوية / البحوث

شامة / ٨/١١