

**DETERMINING THE RELIABILITY AND
VALIDITY OF THE ADAPTED WECHSLER
INTELLIGENCE SCALE-FOURTH EDITION
(WISC-IV) FOR LIBYAN CHILDREN AND
ADOLESCENTS**

ISMAIL ABDALLA ALMAHDI SOUAN

UNIVERSITI SAINS MALAYSIA

2017

**DETERMINING THE RELIABILITY AND
VALIDITY OF THE ADAPTED WECHSLER
INTELLIGENCE SCALE-FOURTH EDITION
(WISC-IV) FOR LIBYAN CHILDREN AND
ADOLESCENTS**

by

ISMAIL ABDALLA ALMAHDI SOUAN

**Thesis submitted in fulfillment of the requirements
for the degree of
Doctor of Philosophy**

July 2017

ACKNOWLEDGEMENT

I would like to acknowledge the contributions of several individuals who were pivotal in the successful completion of my thesis. First and foremost, I would like to thank my main supervisor Assoc. Prof. Dr. Norzarina Mohd Zaharim and my co-supervisor Assoc. Prof. Dr. Intan Hashimah Mohd Hashim for their endless insight and support as well as assistance throughout my thesis project. I would like to thank the Higher Education Ministry of Libya. I especially would like to express my gratefulness to the headmasters of Libyan schools in Zliten regarding support and assistance in completing my field study and also I would like to thank the Libyan children and adolescents who participated in this study. I would like to thank Dean of the School of Social Sciences and all staff at USM. Finally, I would like to acknowledge my parents, who always support and encourage me, and my wife as well as my kids and friends for their continual encouragement and help. Without the endless effort and patience of all the above mentioned individuals, this study project would not have been completed. Thank you!

TABLES OF CONTENTS

Acknowledgement.....	ii
Table of Contents.....	iii
List of Tables.....	ix
List of Figures.....	xii
Abstrak.....	xiii
Abstract.....	xv

CHAPTER 1 – INTRODUCTION

1.1 Overview.....	1
1.2 Background of the Study.....	1
1.3 Intelligence and Intelligence Tests in Psychology.....	4
1.4 Libya: An Introduction to the Location of the Present Study.....	6
1.5 Problem Statement.....	8
1.6 Significance of the Study.....	11
1.7 Study Questions.....	15
1.8 Objectives of the Study.....	15
1.9 Definitions of Terms.....	15

CHAPTER 2 - LITERATURE REVIEW

2.1	Introduction.....	18
2.2	Conceptualization of Intelligence.....	18
2.3	General Intelligence (g factor) and Intelligence Quotient (IQ).....	20
2.4	Theories of Intelligence.....	23
	2.4.1 Psychometric Theories.....	23
	2.4.2 Developmental Theories.....	26
	2.4.3 Information-Processing Theories.....	27
2.5	History of Intelligence Tests.....	30
2.6	Wechsler Intelligence Scales.....	32
	2.6.1 History of the Wechsler Intelligence Scales.....	32
	2.6.2 Introduction of the Wechsler Intelligence Scale for Children- Fourth Edition.....	35
	2.6.3 General Features of the WISC-IV Subtests.....	37
	2.6.4 Benefits of the WISC-IV.....	38
	2.6.5 Strengths of the WISC-IV.....	39
2.7	Translation and Adaptation Process.....	40
	2.7.1 Translation Process.....	40
	2.7.2 Adaptation Process.....	43
2.8	Test Administration.....	45
2.9	Test Scoring.....	48
2.10	Psychometric Properties of the WISC-IV.....	49
	2.10.1 Reliability.....	49

2.10.2 Validity.....	51
2.10.3 Norms.....	53
2.11 Past Studies on the Use of the WISC-IV.....	56
2.11.1 Rashed (1999): Standardization of the Wechsler Intelligence Scale for Children-Revised (WISC-R) for Students Aged 12-15 Years in Iraq.....	57
2.11.2 Naglieri and Paolitto (2005): Ipsative Comparisons of WISC-IV Index Scores	61
2.11.3 Keith et al. (2006): Higher Order, Multisample, Confirmatory Factor Analysis of the Wechsler Intelligence Scale for Children—Fourth Edition: What Does It Measure?.....	62
2.11.4 Scott (2006): Is the GAI a Good Short Form of the WISC-IV?.....	63
2.11.5 Mayes and Calhoun (2007): Wechsler Intelligence Scale for Children-Third and -Fourth Edition Predictors of Academic Achievement in Children with Attention-Deficit/Hyperactivity Disorder.....	64
2.11.6 Abdullah et al. (2009): Arabization of the Wechsler Intelligence Scale for Children-Revised (WISC-R; 1974) in Saudi Arabia.....	66
2.11.7 Lynn et al. (2009): Intelligence in Libya: Norms for the Verbal WISC-R.....	66
2.11.8 Goldbeck et al. (2010): Sex Differences on the German Wechsler Intelligence Test for Children (WISC-IV).....	67
2.11.9 Ryan et al. (2010): Stability of the WISC-IV in a Sample of Elementary and Middle School Children.....	68
2.11.10 Abu-Hilal et al. (2011): Psychometric Properties of the Wechsler Abbreviated Scale of Intelligence (WASI) with an Arab Sample of School Students.....	68
2.11.11 Ambreen and Kamal (2011): Adaptation and Validation of the Verbal Comprehension Index (VCI) Subtests of WISC-IV UK in Pakistan.....	70

2.11.12	Watkins and Smith (2013): Long-Term Stability of the Wechsler Intelligence Scale for Children-Fourth Edition.....	73
2.11.13	Yang et al. (2013): Wechsler Intelligence Scale for Children 4th Edition-Chinese Version Index Scores in Taiwanese Children with Attention-Deficit/Hyperactivity Disorder.....	73
2.11.14	Ambreen and Kamal (2011): Development of Norms of the Adapted Verbal Comprehension Index Subtests of WISC-IV for Pakistan Children.....	74
2.12	Summary of the Past Studies.....	75
2.13	Conceptual Framework for the Present Study.....	77
2.14	Theoretical Framework for the Present Study.....	78

CHAPTER 3 – METHOD

3.1	Introduction.....	81
3.2	Sample of the Study.....	81
3.2.1	Population.....	81
3.2.2	Participants.....	82
3.3	Study Design.....	82
3.4	Instrument.....	83
3.5	Procedures of Test Translation and Adaptation Process.....	84
3.5.1	Translation of Instrument.....	84
3.5.2	Adaptation of Process.....	88

3.6	Pilot Study.....	93
	3.6.1 The Purpose of the Pilot Study.....	93
	3.6.2 Participants of the Pilot Study.....	94
	3.6.3 Results and Discussion of the Pilot Study.....	94
3.7	Main Study.....	100
	3.7.1 Contacting the Schools and the Participants.....	100
	3.7.2. Time for Administering the Wechsler Intelligence Scale for Children-Fourth Edition (WSIC-IV).....	100
	3.7.3 Administration of the WISC-IV.....	101
	3.7.4 Scoring of the WISC-IV.....	104
3.8	Data Analysis.....	104
	3.8.1 Computing Reliability of the WISC-IV.....	105
	3.8.2 Computing Validity of the WISC-IV.....	105
	3.8.3 Statistical Norms.....	106

CHAPTER 4 – RESULTS

4.1	Introduction.....	107
4.2	Sample Scores.....	107
4.3	Objective 1: To Evaluate the Reliability of the Translated and Adapted WISC-IV for Libyan Children and Adolescents.....	108
4.4	Objective 2: To Evaluate the Validity of the Translated and Adapted WISC-IV for Libyan Children and Adolescents.....	116
4.5	Objective 3: To Determine the Score Norms for Libyan Children and Adolescents Based on the Translated and Adapted WISC-IV.....	133

4.6	Summary of the Results.....	143
-----	-----------------------------	-----

CHAPTER 5 – DISCUSSION

5.1	Introduction.....	144
5.2	Objective 1: To Evaluate the Reliability of the Translated and Adapted WISC-IV for Libyan Children and Adolescents.....	145
5.3	Objective 2: To Evaluate the Validity of the Translated and Adapted WISC-IV for Libyan Children and Adolescents.....	147
5.4	Objective 3: To Determine the Score Norms for Libyan Children and Adolescents Based on the Translated and Adapted WISC-IV.....	158
5.5	Time for Administering the Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV)	159
5.6	Theoretical Implications of the Present Study.....	161
5.7	Practical Implications of the Present Study.....	162
5.8	Limitations of the Present Study.....	164
5.9	Recommendations for Future Research.....	166
5.10	Conclusions.....	167
	REFERENCES.....	168

APPENDICES

LIST OF TABLES

			Page
Table	2.1	Nine Independent Types of Intelligence in Gardner’s Theory.....	29
Table	2.2	Four Main Indices of the WISC-IV and their Measurements.....	37
Table	3.1	The Original and Adapted Items of Arithmetic Subtests of the WISC-IV.....	90
Table	3.2	Number and Presentage of Modified Items for the WISC-IV Arabic Adapted Version.....	92
Table	3.3	Suitability of Factor Analysis and Sample Adequacy KMO and Bartlett’s Test.....	99
Table	3.4	Means, Standard Deviations and Standard Errors of Measurement by Age ($N = 28$)	99
Table	4.1	Descriptive Statistics for the WISC-IV Subtests, Subscales, and FSIQ ($N = 210$).....	108
Table	4.2	Reliability Coefficients for the WSIC-IV Subtests, Subscales, and FSIQ ($N = 210$): Split-Half Method.....	111
Table	4.3	Reliability Coefficients for the WSIC-IV Subtests, Subscales, and FSIQ by Age Group Using Split-Half Method.....	111
Table	4.4	Reliability Coefficients for the WSIC-IV FSIQ Using Cronbach’s Alpha Method.....	113
Table	4.5	Descriptive Statistics for the Processing Speed Subtests ($N = 210$).....	113
Table	4.6	Correlation Coefficients between First Testing and Second Testing for the Total Sample ($N = 210$).....	114
Table	4.7	Stability Coefficients for the Processing Speed Index by Age Group Using Pearson Correlations.....	115

Table	4.8	Descriptive Statistics for the WISC-IV Subscales and FSIQ.....	117
Table	4.9	Correlation Coefficients between Subscales and FSIQ for the Total Sample.....	118
Table	4.10	Correlation Coefficients between Subscales and FSIQ for the Total Sample.....	118
Table	4.11	Correlation Coefficients between the VCI Subscale and Its Subtests for the Total Sample.....	119
Table	4.12	Correlation Coefficients between the PRI Subscale and Its Subtests for the Total Sample.....	119
Table	4.13	Correlation Coefficients between the WMI Subscale and its Subtests for the Total Sample.....	120
Table	4.14	Correlation Coefficients between the PSI Subscale and Its Subtests for the Total Sample.....	120
Table	4.15	Pearson Correlation Coefficients between Subtests and FSIQ ($N = 210$).....	121
Table	4.16	Correlation Coefficients between Subtests and Subscales for the Total Sample.....	123
Table	4.17	Correlation Matrix of the 15 WISC-IV Subtests for the Total Sample ($N = 210$).....	127
Table	4.18	Suitability of Factor Analysis and Sample Adequacy KMO and Bartlett's Test.....	128
Table	4.19	Total Variance Explained.....	129
Table	4.20	Exploratory Factor Pattern Loadings of Core and Supplemental Subtests for the Total Sample ($N=210$)....	130
Table	4.21	Means and Standard Deviations of the WISC-IV FSIQ Scores by Age.....	132
Table	4.22	Post-Hoc Multiple Comparisons (Scheffe ^a) for the Relationships between IQ and Age using One-Way ANOVA.....	133

Table	4.23	Scaled Score Values for the WISC-IV Subtests, Subscales, and FSIQ.....	135
Table	4.24	Distribution of the Full Scale IQ Index Scores by IQs Category, Age Group, and the Total Sample.....	136
Table	4.25	Distribution of the Verbal Comprehension Index Scores by IQs Category, Age Group, and the Total Sample.....	137
Table	4.26	Distribution of the Perceptual Reasoning Index Scores by IQs Category, Age Group, and the Total Sample.....	138
Table	4.27	Distribution of the Working Memory Index Scores by IQs Category, Age Group, and the Total Sample.....	139
Table	4.28	Distribution of the Processing Speed Index Scores by IQs Category, Age Group, and the Total Sample.....	140
Table	4.29	Means and Standard Deviations for the FSIQ and Subscale Scores by Age Group.....	141
Table	4.30	Norm Scores for the FSIQ by Age Group.....	141
Table	4.31	Norm Scores for the Verbal Comprehension Index by Age Group.....	141
Table	4.32	Norm Scores for the Perceptual Reasoning Index by Age Group.....	142
Table	4.33	Norm Scores for the Working Memory Index by Age Group.....	142
Table	4.34	Norm Scores for the Processing Speed Index by Age Group.....	142

LIST OF FIGURES

			Page
Figure	2.1	Historical Chart of the Development of the Wechsler Intelligence Scales.....	34
Figure	2.2	Conceptual Framework for the Present Study.....	77
Figure	2.3	Theoretical Framework for the Present Study.....	80

**PENENTUAN KEBOLEHPERCAYAAN DAN KESAHAN SKALA
KECERDASAN WECHSLER–EDISI KEEMPAT (WISC-IV) YANG TELAH
DIADAPTASI UNTUK KANAK-KANAK DAN REMAJA LIBYA**

ABSTRAK

Kajian ini bertujuan untuk menterjemahkan dan mengadaptasikan Skala Kecerdasan Wechsler bagi Kanak-Kanak-Edisi Keempat (WISC-IV; Wechsler et al., 2004) di Libya. Sampel terdiri daripada 210 orang peserta yang berumur dalam lingkungan 6 hingga 15 tahun (umur 6-7: $n = 42$; umur 8-9: $n = 42$; umur 10-11: $n = 42$; umur 12-13: $n = 42$; umur 14-15: $n = 42$). Sampel dipilih menggunakan pensampelan rawak berstrata berdasarkan umur dan jantina. WISC-IV dalam versi Bahasa Arab terdiri daripada sepuluh sub-ujian teras dan lima sub-ujian tambahan, yang boleh dijumlahkan ke dalam empat indeks dan satu skala penuh IQ (FSIQ). Keputusan menunjukkan bahawa instrumen mempunyai kebolehpercayaan (split-half: $r = 0.95$; ujian-ujian semula: $r = 0.81$; dan alfa Cronbach: $r = 0.92$) dan kesahan tinggi. Kesahihan terjemahan dan adaptasi WISC-IV dalam versi bahasa Arab diperiksa melalui ujian korelasi antara sub-ujian. Keputusan menunjukkan bahawa subskala WISC-IV mempunyai korelasi yang tinggi dengan FSIQ ($r = 0.93, 0.90, 0.86$, and 0.86 bagi *Verbal Comprehension Index*, *Perceptual Reasoning Index*, *Working Memory Index*, dan *Processing Speed Index*) dan semua sub-ujian WISC-IV mempunyai korelasi yang tinggi (seperti Vocabulary, $r = 0.90$) dengan FSIQ

kecuali bagi sub-ujian *Symbol Search*, *Digit Span* dan *Coding* yang mempunyai korelasi yang sederhana dengan FSIQ ($r = 0.66$; $r = 0.67$; $r = 0.69$). Selanjutnya, korelasi antara sub-ujian dan subskala berjalut daripada 0.46 (*Verbal Comprehension Index* dan *Symbol Search*) hingga 0.97 (*Verbal Comprehension Index* dan *Vocabulary*). Kesahan juga telah ditentukan oleh faktor analisis. Kaedah Pemfaktoran Axis Utama digunakan untuk faktor pengekstrakan dan kaedah *Direct Oblimin* (putaran *Oblique*) digunakan untuk faktor putaran. Keputusan menunjukkan bahawa dua faktor iaitu *Verbal Comprehension Index* (VCI) dan *Processing Speed Index* (PSI) telah diekstrak. Kesahan daripada perbezaan umur menunjukkan bahawa skor FSIQ meningkat dengan usia. Sebagaimana yang dijangka, terdapat perbezaan statistik yang signifikan dalam IQ antara umur ($F(4, 205) = 102.20, p = .000, \eta^2 = 0.67$), kecuali bagi dua kumpulan umur, iaitu 12-13 dan 14-15. Norma daripada sampel menunjukkan bahawa purata skor skala (IQ) daripada WISC-IV Skala Penuh IQ (FSIQ) adalah 99.91 bagi kesemua sampel ($N = 210$). Keputusan menunjukkan bahawa terjemahan dan adaptasi WISC-IV dalam versi bahasa Arab mempunyai kebolehpercayaan dan kesahan yang baik. Justeru, ujian ini boleh digunakan ke atas kanak-kanak dan remaja Libya pada masa hadapan. Selanjutnya, ujian ini mempunyai potensi yang baik untuk digunakan ke atas kanak-kanak dan remaja Arab yang lain dengan syarat norma tempatan dalam populasi tersebut ditetapkan.

**DETERMINING THE RELIABILITY AND VALIDITY OF THE ADAPTED
WECHSLER INTELLIGENCE SCALE-FOURTH EDITION (WISC-IV) FOR
LIBYAN CHILDREN AND ADOLESCENTS**

ABSTRACT

The present study aimed to translate and adapt the Wechsler Intelligence Scale for Children–Fourth Edition (WISC-IV; Wechsler et al., 2004) in Libya. The sample consisted of 210 Libyan children and adolescents aged 6 to 15 years (age 6-7: $n = 42$; age 8-9: $n = 42$; age 10-11: $n = 42$; age 12-13: $n = 42$; age 14-15: $n = 42$). The sample was selected using stratified random sampling according to age and gender. The WISC-IV Arabic version consists of ten core subtests and five supplemental subtests that can be summed into four indices and one Full Scale IQ (FSIQ). The results revealed that the instrument was reliable (split-half; $r = 0.95$, test-retest; $r = 0.81$, and Cronbach's alpha; $r = 0.92$) and valid. The validity of the translated and adapted WISC-IV Arabic version was examined by inter-subtest correlations. The results indicated that WISC-IV subscales correlated highly with the FSIQ (r s of 0.93, 0.90, 0.86, and 0.86 for Verbal Comprehension Index, Perceptual Reasoning Index, Working Memory Index, and Processing Speed Index, respectively) and that all WISC-IV subtests had high correlations (such as Vocabulary, $r = 0.90$) with the FSIQ except for Digit Span, Coding, and Symbol Search subtests which had moderate correlations with the FSIQ ($r = 0.66$; $r = 0.67$; r

= 0.69, respectively). Further, the correlations between subtests and subscales ranged from $r = 0.46$ (Perceptual Reasoning Index and Coding) to $r = 0.97$ (Verbal Comprehension Index and Vocabulary). Validity was also determined by factor analysis. The Principle Axis factoring method was used for factor extraction and Direct Oblimin method (Oblique rotation) was applied for factor rotation. The results showed that two factors were extracted, the Verbal Comprehension Index (VCI) and the Processing Speed Index (PSI). Validity by age differentiation showed that the FSIQ scores increased with age. As expected, there were statistically significant differences in IQ by age ($F(4, 205) = 102.20, p = .000, \eta^2 = 0.67$), except for two age groups: 12-13 and 14-15. The norms derived from the sample indicated that an average of scaled scores (IQs) of the WISC-IV Full Scale IQ (FSIQ) was 99.91 for the total sample ($N = 210$). The results showed that the translated and adapted WISC-IV Arabic version was a reliable and valid instrument. Therefore, the test can be used for Libyan children and adolescents in the future. Further, the test has good potential to be used with other Arab children and adolescents as long as local norms are established in those populations.

Chapter 1

Introduction

1.1 Overview

This chapter describes the background of the study, conception of intelligence, intelligence tests in psychology, general information about Libyan society, problem statement, the significance of the study, study questions, objectives of the study, and definitions of terms.

1.2 Background of the Study

Over the years, philosophers, researchers, policy makers, academicians and students both within and outside educational institutions, have been grappling with the issue of human intelligence. Many leading scientists and theorists of psychological measurement, such as Raymond B. Cattell, John Horn, J. P. Guilford, Charles Spearman, Louis Thurstone, Alfred Binet and Theophile Simon, Lewis Terman and Dived Wechsler, and Howard Gardner have constructed some tests to measure the strengths, weaknesses, functions and levels of human intelligence as well as cognition that was written, visual or verbal methods (Aiken, 1997; Bukatko & Daehler, 2004; Cattell, 1987; Cook & Cook, 2009; Cummings, 1995).

Due to the vital place of intelligence in cognitive abilities and creativity among individuals, researchers have paid a great deal of attention to it. Owing to the complexities in the genetic, environmental and historical factors affecting human intelligence in the modern age, it has become imperative for tests to be adapted across societies, particularly in the developing countries of Africa, like Libya.

Some intelligence tests conducted in Libya were non-verbal tests and administered in a group (e.g., Al-Shahomee & Lynn, 2012). Al-Shahomee and Lynn's study (2012) was conducted to standardize the Raven test on a representative sample of 520 Libyan adults (260 men and 260 women) aged between 38 and 50 years in order to measure their general cognitive abilities. The results revealed that the intended sample had a mean IQ of 79. As shown, the reliability tested by Cronbach's alpha (KR-20) for the RSPM for the total sample was 0.92, while split-half reliability for the total sample was 0.88. The Principal components analysis method results showed that only one loaded factor was extracted (with an eigenvalue greater than 1) that accounted for 58.56% of the total variance explained. According to Al-Shahomee and Lynn (2012), the lower scores of the Libyan sample on the RSPM test were attributed to various reasons: An emphasis in schools on memorization at the expense of problem solving skills; the human development report in 2002 on Libya, which stated that the teaching skills of many teachers were deficient in this regard; poorer schools, such as the average class size being 30 or more students per teacher and school buildings and facilities being out-dated and inappropriate for effective teaching. Moreover, up-to-date computer programs are not available in 89% of schools, as well as poor nutrition, which has been shown to have an adverse effect on intelligence.

Another study by Bony and Al-Majdoub (1999) entitled "Translated and Standardized Cattell's Culture Fair Intelligence Test, Scale Three (CCFIT-III) for Libyan Adults". This study was conducted in a group and standardized on a sample of 2600 Libyan high school students (males and females). The results showed that the reliability coefficient was 0.60 using the Equivalent Forms method and 0.78

using the split-half method. Concurrent validity was used to examine the validity of the test by identifying the correlation coefficient between the Cattell's Culture Fair Intelligence Test, Scale Three and the Raven's Standard Progressive Matrices Test. The results indicated that the correlation coefficient was 0.58 for the total sample.

To the best of researcher's knowledge, there has been one past study of intelligence in Libya related to the present study, which was the study by Lynn et al. (2009). It was conducted to administer the Wechsler Intelligence Scale for the Children-Revised verbal subscale (WISC-R; Wechsler, 1974) to a sample of Libyan children and adolescents to measure their general cognitive abilities. However, the non-verbal subscale of the WISC-R was not administered, as this study was the second version of the Wechsler tests. Nevertheless, the WISC-R as an individual test overall consists of both verbal and non-verbal subscales that should be administered together as a complete test to measure general cognitive ability.

In addition to those mentioned above, most of these past studies, which were conducted in Libya, used non-verbal tests. Furthermore, the WISC-IV has not been adapted and applied in Libya (see Appendix A regarding Letter from the National Library for Scientific Research). Thus, no suitable test can be used to assess cognitive ability effectively, even though some studies were conducted in Libya. In the present study, the verbal and non-verbal subscales of the Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV) were translated, adapted and administered to a sample of Libyan children and adolescents aged from six to 15 years old. The WISC-IV is an individually administered intelligence test and is broadly used to assess the general cognitive ability of children and adolescents (Wechsler et al., 2004). It consists of 15 verbal and non-verbal subtests that are

summed into four subscales/indices and Full Scale IQ (i.e., general cognitive ability). An individual administration (e.g., WISC-IV) using a test manual gives a better chance to observe examinees than group administration (Aiken & Groth-Marnat, 2006).

1.3 Intelligence and Intelligence Tests in Psychology

There are various conceptions and definitions of intelligence. David Wechsler (1944) defined intelligence as “the aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with the environment.” (p. 3). Intelligence is a general intellectual ability that includes the capability to think abstractly, to solve problems, to understand complex ideas, learn quickly and learn from experience (Fletcher & Hattie, 2011; Pässler & Beinicke, 2015). According to Sternberg (2014), intelligence is the ability to adapt to, shape, and select environments. Further, there are many different definitions of intelligence, but ultimately, intelligence is a person's interactions with the environment in which he/she lives.

Intelligence is measured by individual and group intelligence tests, for example, the Binet-Simon Intelligence Scale (BSIS) (Flanagan & Kaufman, 2009); the Wechsler-Bellevue Intelligence Scale for Children (WISC; Wechsler, 1939, as cited in Wechsler et al., 2004), to name a few. These tests are individually administered and commonly used intelligence tests on school children around the world and have been applied on several occasions to evaluate the cognitive abilities of children and adults in different societies (Aiken, 1997). Furthermore these tests (BSIS & WISC) have introduced, applied, modified and even standardized by

different researchers and students, whose research interests span through psychology, especially educational psychology and psychometric studies.

The first formal intelligence test was created in 1905 by Binet and Simon (Bukatko & Daehler, 2004; Rathus, 2006; Scott & Spencer, 1998). The Binet-Simon scales were revised by other researchers and also translated into other languages. For instance, English, German, Swiss, Italian, Russian, Chinese, Japanese, and Turkish (Binet & Simon, 1916). There have been three English translations and adaptations of the Binet-Simon scale in the United States. The most popular of these versions, the Stanford-Binet Intelligence Scale was published by Lewis Terman in 1916 at Stanford University in the United States, and this revision became known as the Stanford-Binet Intelligence Scale (SBIS). The version that is used today is the Stanford-Binet Intelligence Scale–Fifth Edition. The Stanford-Binet intelligence scale is an uniform test to evaluate cognitive ability for children and adults aged 2 to 23 years (Aiken, 1997; Rathus, 2006; Scott & Spencer, 1998).

The Wechsler intelligence tests are available in three forms: The Wechsler Intelligence Scale for Children (WISC), the Wechsler Adult Intelligence Scale (WAIS), and the Wechsler Preschool and Primary Scale of Intelligence (WPPSI; Wechsler et al., 2004), which are the most broadly used to measure intelligence and for neuropsychological assessments (Aiken, 1997; McIntire & Miller, 2000). The Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV) is one of the most widely used in measuring the intelligence of children and adolescents, as well as an instrument commonly used to determine the intellectual and cognitive capabilities of children, aged 6 to 16 years (Cheramie et al., 2008; Wechsler et al., 2004). The fourth edition of this popular instrument was created in 2003, revised

from its predecessor the WISC-III, but updated with structural and research-based changes (Kaufman et al., 2006).

The Stanford-Binet and Wechsler tests are considered to be the most famous in the field of intelligence measurement from among individual intelligence tests, and may be applied on one individual at a time (Coon & Mitterer, 2007). The tests require considerable time and effort in their application, correction, and interpretation of results. As a result, psychologists have developed tests suitable for measuring a wide range of individuals, and which fall under the so-called collective tests and are able to be applied to a group of individuals at one time. Among these tests are the Cattell's culture fair intelligence tests (Coon & Mitterer, 2007).

1.4 Libya: An Introduction to the Location of the Present Study

As the present study was conducted in Libya, information pertaining to this country serves as a context for the present study. Libya is one of the most prosperous countries in the North African region, although it is sometimes considered to be a middle-Eastern country in the Arab world (Thompson, 2011). The capital of the country is Tripoli. Libya has a population of 6,546,000 citizens and foreigners who live in different parts of the country (Population Reference Bureau, PRB, 2011). The United Nations Population Fund (UNFPA, 2011) reported that the female share of the population was 3.2 million, while the male accounted for the remaining 3.2 million.

The United Nations Development Program (UNDP, 2011) recently indicated in its Human Development Report (HDR) that in 2011, about 78.1% of the Libyan population lived in urban areas. This implies that less than one-fourth of the

population resided in places considered to be rural areas. In a related version, the United Nations Human Settlement Program (UN-HABITAT, 2011) reported that out of the 6.546 million people in Libya in 2011, urban residents accounted for 5.098 million, while the number of rural residents was 1.447 million. The capital city of the country, Tripoli, had 1.108 million inhabitants in 2011.

Education in Libya is made up of four levels; pre-school, primary, preparatory, and secondary. Preschool includes pupils who are between the ages of 3 and 6 years, primary school includes pupils whose ages range from 6 to 12 years, and those attending preparatory schools are young students between 13 and 15 years old. The secondary school is made up of adolescent students who are aged between 16 and 18 years.

Primary and preparatory levels are under the basic educational system. This is probably one of the main reasons that made the government decide to make basic education compulsory for all Libyan children. According to Clark (2004), the educational system “from primary school right up to university and post-graduate study” is not only freely provided to all citizens, but it is also made compulsory for those that fall between the ages of 6 and 15 years, as this is regarded as basic education.

The people of Libya are predominantly Muslims, although they belong to different ethnic groups, which primarily include Arabs, Arab-Berber, Berbers, black African groups like Touareg and Tabu tribes, who are mostly found in the southern part of the country. According to the *Population Reference Bureau* (2011, p. 6), 31% of the Libyan population are less than 15 years old, while 4% are 65 years and above in mid-2011. Further, the United Nations Development Program (UNDP, 2011)

reported that the adult literacy rate (15 years and older) between 2005 and 2010 was 88.9%, while the gross enrollment ratio was 110.3% in the primary level and 93.5% in secondary, as well as 55.7% in the tertiary level between 2001 and 2010 in Libya.

The education system in Libya has been affected by the recent revolution, which spread across the country from February to September 2011. A large number of children, adolescents and their families were destabilized owing to many bombs that were detonated across different conflict-ridden communities in the country. Aside from the thousands of innocent people, including schoolchildren and adolescents who lost their lives in the revolution, quite a number of young children sustained physical, mental, and psychological injuries, which posed a serious threat to their mental abilities. This was coupled with the fact that many schools were destroyed during the war, and consequently a lot of schoolchildren were not able to attend classes again, particularly, those whose parents and guardians were killed in the revolution and those whose means of sustenance, including housing were wrecked.

1.5 Problem Statement

Various problems exist and justify the present study. Firstly, there is no current study in Libya to individually measure overall verbal and non-verbal abilities. In an earlier study by Lynn et al. (2009), the non-verbal subscale of the WISC-R was not administered; thus, the test was not complete. This is due to the fact that both verbal and non-verbal subscales should be administered. By administering only the verbal subscale of the WISC-R, it may not be valid/sufficient in determining general cognitive ability for Libyan children and adolescents or in

obtaining adequate and satisfactory clinical information about their cognitive abilities. According to the study by Lynn et al. (2009), the verbal subtests of the Wechsler Intelligence Scale for Children-Revised (WISC-R) were administered to a sample of Libyan children and adolescents.

Secondly, there is no suitable test to place Libyan children and adolescents at appropriate academic levels according to cognitive abilities. Stakeholders in educational guidance and counselling and policy makers of education with regard to placement in schools seldom used standardized intelligence tests to place students in a proper place to study. Academic report scores are usually used for this purpose. According to the study by Salem (2013), most standards for the placement process in schools or specific specializations are invalid/unacceptable because they are solely based on the academic achievement scores of a student. Instead, they should be based on both the academic achievement scores of a student and results of standardized intelligence tests that measure general cognitive ability.

Thirdly, there is no suitable test, which could be used to identify disabilities, such as attention-deficit hyperactivity disorder (ADHD). There were no tests that could be administered individually. This may lead this group of children to be left without any help. According to the study by Ashir (2013) on Attention Deficit Hyperactivity Disorder (ADHD) among kindergarten children in Tripoli, Libya, the source of the complaint and worry among teachers and mothers is ADHD. This led to the need for providing a reliable and valid test to identify it. Among intelligence tests that can identify ADHD is the WISC-IV.

Fourthly, there are outdated studies that have been used to determine the cognitive ability for Libyan children and adolescents; for example, the study by

Bony and Al-Majdoub (1999), which was entitled “Translated and Standardized the Cattell's Culture Fair Intelligence Test, Scale Three (CCFIT-III) for Libyan Adults.”

Fifthly, in Libya, most of the tests are non-verbal tests. Therefore, verbal abilities of Libyans have not been tested. The non-verbal intelligence tests whether translated or standardized were used with Libyan children and adolescents. For example, the study by Bony and Al-Majdoub (1999) was entitled “Translated and Standardized the Cattell's Culture Fair Intelligence Test, Scale Three (CCFIT-III) for Libyan Adults”, and the study by Souan (2006) was also entitled “Standardization of the Cattell's Culture Fair Intelligence Test, Scale two (CCFIT-II) for Libyan Children.” Furthermore, the study by Al-Shahomee and Lynn (2012) was entitled “A Standardisation of the Raven’s Standard Progressive Matrices Test for Libyan Adults Aged 38 to 50 years” were all based on non-verbal tests.

Sixthly, a study conducted in Libya using the Raven’s Standard Progressive Matrices Test (RSPM) showed that the test is not suitable to measure cognitive ability. The RSPM is considered one of the non-verbal intelligence tests. According to the study by Gignac (2015) on estimating the association between Raven's test and general intelligence factor, the results showed that Raven’s test is a poor indicator of general intelligence, *g*. According to Wechsler (1944), any test is either verbal or non-verbal cannot alone fully measure a person’s general intelligence.

Seventhly, in contrast to culturally fair intelligence tests, such as group tests that were administered in Libya society, the Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV) is an individual measure, and widely used across the world to evaluate general cognitive ability, yet has not been adapted and applied in Libya. To date, no study has ever been conducted using the WISC-IV in

Libya (see Appendix A regarding Letter from the National Library for Scientific Research). Given its many benefits and functions, it is very important to apply the WISC-IV in determining the intelligence capabilities of Libyans, especially children and adolescents aged between 6 and 15 years.

Eighthly, since there are no appropriate tests, mental capability and retardation cannot be determined. Therefore, a lack of usable tests, such as uniform intelligence tests exists in related centres and schools to identify these categories. This may lead to minimal chances of discovering and caring for students in schools.

Ninthly, some norms used in Libya were based on norms that had been derived from other Arab countries, such as Egypt and Tunisia, and were also based on non-verbal intelligence tests. Additionally, even though there are Libyan norms, there are no suitable norms which can be used to measure the general cognitive abilities of Libyans. For example, the norms of the study by Lynn et al. (2009), which was related to the present study, were only derived from six verbal subtests of the WISC-R; namely Similarities, Vocabulary, Comprehension, Information, Arithmetic, and Digit Span. However, there are nonverbal subtests of the WISC-R to obtain norms for the complete test.

1.6 Significance of the Study

Firstly, the present study is important because the WISC-IV Arabic version was administered completely, i.e., verbal and non-verbal subtests of the WISC-IV as Full Scale IQ, which can be used to measure the general cognitive ability for Libyan children and adolescents.

Secondly, the present study is important because the translated and adapted WISC-IV Arabic version can be used to place Libyan children and adolescents in school according to their cognitive abilities. The results of the present study will benefit educational policy makers and stakeholders in educational guidance and counselling, including those involved in activities related to the placement of children in schools and career choice. Additionally, the WISC-IV Arabic version as general cognitive ability can be used to correlate and compare with school performance in order to identify the kind of correlation, and whether there is any gap in an individual's performance according to this relation.

Thirdly, the present study is important because the WISC-IV Arabic version can be used to identify disabilities, such as attention-deficit hyperactivity disorder (ADHD) and learning disabilities (LD). According to the WISC-IV Technical and Interpretive Manual, the Wechsler scales have proved their clinical usefulness for purposes, such as the identification of intellectual or, learning disabilities, placement in specialized programs and so on (Wechsler et al., 2004). According to the study by Mayes and Calhoun (2007), the results showed that the Working Memory Index (WMI) and the Processing Speed Index (PSI) are the most powerful predictors of LD in children with ADHD. Furthermore, the results of the study by Zhu and Chen (2013) showed that children with individual disabilities, LD, ADHD, and mathematics disorder had significant deficits in performance on the cancellation subtest (i.e., Cancellation random and structured) of the WISC-IV. These results showed the usefulness of the cancellation subtest in clinical assessment.

Fourthly, the present study is important because the WISC-IV Arabic version can be used to identify Cognitive abilities for Libyan children and adolescents. The

purpose of the this study was to demonstrate that the adapted version of the Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV) is a reliable and valid instrument of the WISC-IV for use with Libyan children and adolescents. The present study seeks to provide the educational system in Libya with an adapted test of intelligence, i.e., the WISC-IV, for school children aged from 6 years to 15 years. The test can be used to aid psychologists and educators to identify children who may experience further difficulties in different areas of the school curriculum (Rashed, 1990).

Fifthly, the present study is important because the WISC-IV Arabic version can be used in the field of education to assist stakeholders, particularly teachers, in identifying the strengths and weaknesses of school students. The WISC-IV has 15 subtests (Wechsler et al., 2004). These subtests of the WISC-IV scaled scores (IQs) can be used to give an accurate estimate of an individual regarding their strengths and weaknesses (Gregory, 2000). The WISC-IV not only provides an overall IQ score, it also yields indices/subscales (such as the Verbal Comprehension Index), which allow the examiner to quickly determine the areas in which the child is strong or weak (Santrock, 2011). Groth-Marnat (2003) mentioned that intelligence tests, specifically the WAIS-III and WISC-III, provide valuable information regarding an individual's cognitive strengths and weaknesses.

Sixthly, the present study is important because the WISC-IV Arabic version can be used as an appropriate instrument to assist practitioners and clinical researchers in assessing children's intelligence. The test can be provided with valuable information during neuropsychological evaluations (Wechsler, n.d.). With regard to brain dysfunction, if there are large differences in verbal and non-verbal

intelligence, this may indicate specific types of brain damage (Wechsler Intelligence Scale, n.d.).

Seventhly, the present study is important because the WISC-IV Arabic version can be used to distinguish between gifted and mentally retarded children. It can be used to identify these categories so as to design different programs for promoting and evaluating developmental abilities among individuals (Groth-Marnat, 1997; Kline, 1999). Thus, they can be given special care and education.

Eighthly, the present study is important because the WISC-IV Arabic version can be used as a reliable and valid instrument when conducting future studies. It can be used by the national library for scientific research as a uniform tool to assist researchers and practitioners in understanding a method of adapting tests and profiting from the results and recommendations of the present study, especially educators and teachers, etc. The study by Mayes and Calhoun (2007) showed that the results from both the WISC-III and the WISC-IV revealed that the Full Scale IQ (FSIQ) was the strongest single predictor of achievement in all areas.

Ninthly, the present study is important because the norms of the WISC-IV Arabic version were derived from overall 15 verbal and non-verbal subtests, which can be used to determine general cognitive ability for Libyan children and adolescents. Furthermore, it can be also used to identify scaled scores (IQs) for Libyan children and adolescents to classify them according to their IQs; namely, whether a child's IQs is an average, below average, or above average. This in turn determines whether a child is within a category, such as gifted or mentally retarded, to name but two.

1.7 Study Questions

The present study attempted to provide answers to the following questions:

- 1) Is the translated and adapted WISC-IV a reliable measure for assessing Libyan children's and adolescents' intelligence?
- 2) Is the translated and adapted WISC-IV a valid measure for assessing Libyan children's and adolescents' intelligence?
- 3) What are the score norms for Libyan children and adolescents based on the translated and adapted WISC-IV?

1.8 Objectives of the Study

The objectives of the study were as follows:

- 1) To evaluate the reliability of the translated and adapted WISC-IV for Libyan children and adolescents
- 2) To evaluate the validity of the translated and adapted WISC-IV for Libyan children and adolescents
- 3) To determine the score norms for Libyan children and adolescents based on the translated and adapted WISC-IV

1.9 Definitions of Terms

1.9.1 Intelligence

Intelligence is “an individual's ability to derive information, learn from experience, and adapt to the environment while correctly understanding and utilizing acquired thoughts and reasons” (VandenBos, 2007, p. 488). According to David

Wechsler, intelligence is “the global capacity of the individual to act purposefully, to think rationally and to deal effectively with his environment” (Wechsler, 1944, p. 3).

The operational definition of intelligence is an individual’s performance in the adapted Arabic version of the Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV).

1.9.2 Intelligence measurement

Intelligence is measured as an IQ score. Intelligence quotient (IQ) is “a score obtained from an intelligence test by dividing the mental age (MA) obtained on the test by the actual/chronological age (CA) and multiplying by 100, i.e. $IQ = MA / CA \times 100$ ” (Statt, 1998, p. 74). This form (i.e., classic intelligence quotient) has been replaced by the deviation intelligence quotient/scaled score (IQs), which is defined “by one’s place in a uniform curve of scores on an intelligence test which has a mean of 100 and a standard deviation of either 15 or 16, depending on the test” (Matsumoto, 2009, p. 260). The deviation intelligence quotient (IQs) is a measure of intelligence, which can be obtained by administering uniform intelligence tests (e.g., WISC-IV) by using the formula as follows: $IQ_s = (Z \times SD) + M$ (Urbina, 2014). Furthermore, Wechsler (1944) defined intelligence quotient (IQ) as something which “merely tells us how much better or worse, or how much above or below the average any individual is, when compared with persons of his own age” (p. 41).

1.9.3 Translation

“Translation refers to “a language processing and text production task involving two different languages: the source language (SL) and the target language

(TL). It requires linguistic competence in both the SL and the TL” (Dimitrova, 2005, p. 2). A bilingual person translates items from the source language to the target language (VandenBos, 2007, p. 954).

1.9.4 Adaptation

“Adaptation is a process of revising items, concepts, words, and expression so that these are equivalent culturally and linguistically in a target language and culture” (Hambleton et al., 2006, p. 4). It is a process that entails careful planning concerning its content maintenance, psychometric properties of the new version of the instrument, and general validity for the intended population (Borsa et al., 2012).

1.9.5 Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV)

The WISC-IV is an instrument to assess the cognitive ability of children, young persons and early, mid and late adolescents aged between 6 and 17 years. It has four major indices, the Verbal Comprehension Index, the Perceptual Reasoning Index, the Working Memory Index, and the Processing Speed Index. These indices are supported by the Full Scale IQ (Kaufman et al., 2006). The WISC-IV has 15 subtests (core and supplemental subtests) are as follows: Similarities (SI), Vocabulary (VC), Comprehension (CO), Block Design (BD), Picture Concepts (PCn), Matrix Reasoning (MR), Digit Span (DS), Letter-Number Sequencing (LN), Coding (CD), and Symbol Search (SS), Information (IN), Word Reasoning (WR), Picture Completion (PCm), Arithmetic (AR), and Cancellation (Wechsler, 2003b) (for more details, refer to Section 3.4 in Chapter 3).