City Research Online

City, University of London Institutional Repository

Citation: Najmaldeen, G. (2020). A comprehensive assessment of reading-related skills in typically developing 4;0- to 7;0-year-old Saudi Arabian children. (Unpublished Doctoral thesis, City, University of London)

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/26064/

Link to published version:

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online: http://openaccess.city.ac.uk/ publications@city.ac.uk

A Comprehensive Assessment of Reading-Related Skills in Typically Developing 4;0- to 7;0-Year-Old Saudi Arabian Children

Ghada Najmaldeen

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

City, University of London

Division of Language and Communication Science

February 2020

CONTENTS

3 CHAPTER THREE: STUDY ONE

4. CHAPTER FOUR: STUDY TWO

List of Figures

List of Tables

ACKNOWLEDGMENT

It has been a privilege to study for a PhD and what I have learned on my journey goes far beyond the contents of this thesis. Research involving fieldwork requires the help of many people and I would like to thank them for their contribution to the invaluable experience that I have gained.

First and foremost, I wish to acknowledge my debt of gratitude to my project supervisors, Professor Shula Chiat and Professor Penny Roy, for helping to guide my development throughout my research journey. Shula’s commitment and attention to detail were outweighed only by her patience in helping me to find a way through. Penny’s sincere interest, encouragement and invaluable advice on statistics were also greatly appreciated. I would also like to thank Dr. Anne Hesketh (external examiner) and Professor Victoria Joffe (internal examiner) for their helpful feedback.

My grateful thanks go to all the school principals, administrators and teachers without whose cooperation the testing for these studies could not have taken place. Special thanks also to all the children and parents who gave up their time to participate in this research. Without their invaluable contribution, this thesis would not exist.

My heartfelt thanks go to my mother, who passed away three years ago, for being a constant source of comfort. Without her prayers, I would never have made it thus far. I am also deeply indebted to my father, who right from the start, gave me the determination and ambition to strive to reach my goals. Many thanks to my children Abdullah, Majed and Jude, for without their love and encouragement I would never have finished this thesis.

Finally, many special thanks to all the other ‘special people’ in my life for the love, understanding, support and unfaltering belief they have offered me.

DEDICATION

"...رب أوزعني أن أشكر نعمتك التي أنعمت علي وعلى والدي وأن أعمل صالحا ترضاه وأدخلني برحمتك في عبادك الصالحين" آيه 19 – النمل

… My Lord! Inspire me and bestow upon me the power and ability that I may be grateful for Your Favours which You have bestowed on me and my parents, and that I may do righteous good deeds that will please you and admit me by Your Mercy among Your righteous slaves.

Holy Qur'an 27:19

DECLARATION OF ORIGINALITY

I hereby declare that this thesis is my own work and has not been submitted in any form for another degree at any university or other institute of tertiary education. Information derived from the published and unpublished work of others has been acknowledged in the text and a list of references is given in the bibliography.

ABSTRACT

This cross-sectional study involves Arabic-speaking children from Riyadh in Saudi Arabia. It aims to develop comprehensive Reading-Related Skills (RRS) using a battery of tests for the use of Arabic-speaking preschoolers in Saudi Arabia. The first study involved 384 participants as they began to learn to read at 4 years old and as they developed their skill up to 7 years old. The study investigated the extent to which children's performance in Phonological Awareness (PA), Letter Knowledge (LK) and Rapid Automatized Naming (RAN) are affected by gender, age, and socioeconomic status (SES), and also by the levels of exposure to tuition in Tajwid (Qur'anic recitation). The second study involved 60 participants to determine the predictiveness of the test battery in relation to teacher ratings of participants' reading ability.

The results from these studies confirm the validity and reliability of the test battery of test and support previous findings that PA develops from larger to smaller linguistic units. The effect of gender on RRS was found to be minimal whereas both PA and LK were found to be age sensitive and useful for discriminating between levels of ability, whilst RAN proved to be useful only for the oldest group.

The result also supports previous claims by researchers that both PA and LK play an important role in the development of reading in Arabic-speaking children and can be used as predictors of their reading ability. Similarly, the results of the RAN tasks support previous findings that the predictiveness of these tasks applies principally to older children. These points would need to be incorporated into any future RRS test. Although some socioeconomic factors were found to correlate with the RRS measured, SES did not appear to play the major role suggested in previous Western studies. Findings were inconclusive concerning the effect of levels of exposure to *Tajwid* tuition on RRS but indicate that this

area merited further investigation, particularly given the special emphasis placed on this in the Saudi Arabian system of education.

LIST OF ABBREVIATIONS USED IN THIS THESIS

ANOVA	Analysis of variance
B	Beta, type II error rate (1-power)
BPVS	British Picture Vocabulary Scale
C	Consonant
CTOPP	Comprehensive Test of Phonological Processing
DGER	Directory of General Education Report
Df	Degrees of freedom
F	variance of the group means/mean of the within group variances
g	general intelligence
G-P- C	grapheme-phoneme- correspondence
ICC	Intraclass Correlation
IALP	International Association of Logopedics and Phoniatrics
IPA	International Phonetic Alphabet
LK	Letter Knowledge
M	Mean
Med	Median
Min	Minimum
Max	Maximum

MSA	Modern Standard Arabic
n	sample size
p	Significance level
PA	Phonological Awareness
PAT	Phonological Abilities Test
QUIL	Queensland University Inventory of Literacy
PIPA	Preschool Inventory of Phonological Awareness
r	Pearson's product moment correlation coefficient
R^2	Coefficient of determination
RAN	Rapid Automatized Naming
RRS	Reading-related skills
REEL-2	Receptive-Expressive Emergent Language Test
s.d./SD	Standard Deviation
SES	Socioeconomic Status
Std. error/SE	Standard error

1 CHAPTER ONE: OVERVIEW OF RESEARCH

1.1 Introduction

The primary aim of this research was to develop and evaluate a battery of assessments of skills known to underpin the development of reading, and more specifically decoding, for young Saudi Arabic children. Currently, there are formal measures to assess the reading and pre-reading skills of Arabic language, which should be critical for identifying children with reading difficulties in their early development stages. However, these measures are limited and not standardized; resulting in poor understanding of a child's reading difficulties and the risk of providing inappropriate intervention. This research, therefore, adopts the framework originally developed by Frith (1995). Also, this research set out to explore the development of reading-related skills (hereafter referred to as RRS) and the influence of demographic factors in Arabic-speaking children in Saudi Arabia. Two studies were conducted using a sample of 4;0-7;0 year-olds in Riyadh to assess their performance on a series of RRS. The RRS which receive special attention in this thesis are phonological awareness (PA), letter knowledge (LK) and rapid automatized naming (RAN) which have been shown to predict decoding skills in many languages. These new measures of RRS were used to investigate the impact of gender, age, socioeconomic status (SES), and additional exposure to sessions which focused on *tajwid* (learning to recite the Qur'an). A key objective of this research is to produce a solid foundation for the development of a comprehensive standardization of RRS assessment to be used in Saudi Arabia.

The sections covered in Chapter One of this research include: the rationale for this research topic; the research context – contextual information about the educational setting and the geographical location in which the studies were conducted; and the linguistic context - an explanation is also given concerning *Tajwid* and its role in the Saudi curriculum. Information is also provided about the unique characteristics of Arabic as a language, and the particular challenges which it poses to those learning to read and their relevance to RRS. Chapter Two provides an overview of the reading models – the theoretical rationale for investigating RRS in children; and the framework for reading development proposed by Frith (1995) which includes both the environmental and cognitive contributions to reading. This is then followed by a review of evidence on relations between RRS and reading in English and other languages, including Arabic, and the effects of demographic factors. This leads to the aims, objectives, questions and hypotheses which the study addresses. These research aims and questions were addressed in two studies, which are presented in Chapters Three and Four. The results are summarised and discussed in Chapter Five, with concluding remarks in Chapter Six.

1.2 Rationale for the Research

Literacy is an essential part of the modern society and being able to read has proven intellectual benefits. It unlocks the riches of literature, the culture - both past and present, and emerging ideas from a wealth of knowledge articulated in scrolls, journals, articles and books, to name a few. Without reading skills, it is difficult to access the required knowledge for curriculum subjects (Saunders, 2011; Grainger, 2010). Therefore, it is no surprise that educational institutions, teachers and parents agree that proficiency in reading is a core literacy skill to develop. As part of the intellectual benefits, reading is an interactive process, conveying knowledge absorbed from texts to the reader (Cohen and Cowen, 2007; Shiotsu, 2010).

Daly et al. (2015) maintain that mastering multiple competencies needed for effective reading particularly for young learners, is a far from simple process. Reading involves more than merely decoding script and reading it aloud (Forrest-Pressley and Waller, 2013). The knowledge of alphabets, their corresponding sounds and pronunciation is only the first step towards reading successfully (Beauchat et al., 2012). Frith's (1995) Framework discussed in detail in **section 2.2.5** of this thesis, offers an insight into the critical success factors influencing a child's ability to read the Arabic language. It is arguable that these factors could be of a biological, cognitive, behavioural and environmental nature and are often interrelated in complex ways.

This study is focused on the preliminary steps of acquisition of reading skills, in particular on what are recognized as being some of the core components of readiness to read, namely, the RRS known as PA, LK, and RAN (Paris and Paris, 2007; Rathvon, 2004) which have been linked to the development of decoding skills. Research suggests that children who succeed in mastering these preliminary steps of reading, especially in their formative years, go on to become successful readers (Caravolas et al., 2012). PA, LK and RAN have been chosen because research suggests they are good predictors of later reading competences for children across various languages (Caravolas et al., 2005, Caravolas et al., 2012). Various studies have also established that these skills improve with age and can be affected by factors including gender and socioeconomic status (SES) (Moura et al., 2009; Hecht et al., 2000; Menkes et al., 2006 and Gillon, 2012). Frith's model (1995) suggests that there is the need to observe not only at these areas but also other environmentally determined factors which may influence RRS.

This thesis also explores a unique socio-cultural feature of the Saudi Arabian educational system: the emphasis placed on the study of *Tajwid* (Quránic recitation) for children at an early age. The teaching methodology explored is based on the centuries' old and much revered tradition which emphasizes recitation and rote-learning of Islamic scriptures. Children are taught to recite the Qur'an and memorize passages from it even before they can learn to read its actual text (Wagner, 1994), for which they must draw on phonological processing and memory. In this study, the focus is on the children who have received additional Qu'ran tuition.

Later, in **section 1.4**, this research explains that Arabic is a world language, spoken by millions and revered by followers of Islam as the word of God itself revealed through the Qur'an. Therefore, the Arabic language forms part of the rich cultural heritage and remain a thriving culture of Saudi Arabia. However, the section explores the challenges in developing proficient, independent readers with the necessary skills to decode Quranic messages accurately. The percentage of illiteracy is still high in Arabophone countries, standing at 35 percent across this region (Yafi, 2012: 134). The percentage of those who can read effectively is also low (Olson and Torrance, 2009). For comparison, less than 1 percent of the English population would be described as completely illiterate, although this absolute definition is not often used. Instead, the term "functionally literate" is more commonly used and approximately 16 percent, or 5.2 million adults in England, can be described as "functionally illiterate" (Literacy Trust: online).

1.3 The Research Context

The Kingdom of Saudi Arabia extends over most of the Arabian Peninsula (**see Figure 1.1**), with a population of 27 million inhabitants (World Factbook, 2013). It is considered as the centre of the Islamic world. Saudi Arabia's culture is predominately Islamic, and every year, millions of people visit the cities, such as the Holy cities of Makkah and Madinah, to make their pilgrimage. Saudi Arabia has a long tradition of foreign and migrant workers since the 1930s, drafting in workers from neighbouring countries and South-East Asia, thus, making its population demographically diverse. Estimates suggest that there are over five and a half million foreign workers in the Kingdom (World Factbook, 2013). The migrant workers typically include Palestinians, Yemenis, Egyptians, Pakistanis and Indians, and more recently, workers from the Philippines, Thailand, and South Korea (World Factbook, 2013).

Figure 1.1: The Kingdom of Saudi Arabia.

(Source: http://www.operationworld.org)

Although there are migrant workers in many different sectors, most of these tend to be unskilled; males tend to find employment in industries such as agriculture and construction, whilst most females end up in domestic service. Saudi Arabia's economy remains heavily dependent on these migrant workers, since Saudi citizens are usually unwilling to take these kinds of menial jobs (Al-Asmari, 2008).

The diverse composition of Saudi Arabia's population has potential implications for this study. As of 2013, the population of Riyadh (where this investigation was carried out) was 4.3 million, and nearly a third of the population were non-Saudi nationals (World Factbook, 2013). Although all of the participants (students and parents) in this study were of Saudi nationality, many of them employed non-Saudi nationals as 'live-in' domestic help, nannies or car drivers. This suggests that many of these children in the sample spoke more than one language in their home environment apart from the English language spoken amongst non-Arabic speakers. It is worth noting that the children raised under such a background would have had to learn other variations of the Arabic language, especially, since many of the drivers and school teachers in the capital are foreign, such as Egyptians and Syrians. Thus, the fact that the participants lived in a demographically diverse environment may have impacted upon their performance.

1.3.1 The Saudi Arabian education system

The Saudi education system has undergone a significant transformation in the period since 1945, when King Abdulaziz bin Abdulrahman Al-Saud began an extensive programme to establish a state school system within the country (DGER, 2010). Compulsory education in Saudi Arabia now consists of some 12 years of schooling that begins at six and ends at 18 and is free for all citizens at elementary, intermediate and secondary level.

STAGE	AGES	QUALIFICATION
Kindergarten	3-5	Not Applicable
Elementary	6-12	Elementary school certificate*
Intermediate	12-15	Intermediate school certificate*
High school	15-18	High school diploma*
Diploma	18+	Diploma
Undergraduate studies	18+	Bachelor's degree (5 years)
Postgraduate studies	24+	Masters or PhD

Table 1.1: The Saudi Arabian Education System.

*All three of these stages are now compulsory for Saudi students

Formal teaching of literacy begins at the age of six but some children may have some informal instruction in letter recognition at kindergarten. Education is completely gender segregated at all levels. Some 73.9% of male and 78% of female students are enrolled in public schools and Saudi citizens make up 75% of the teachers in both public and private schools. The remaining posts are filled by staff from other Arabic-speaking countries in the region, particularly Egypt (DGER, 2010).

As Error! Reference source not found. shows, children between the ages of three and five may attend kindergarten in Saudi Arabia. However, this is not a compulsory stage of education, and the vast majority of Saudi children are not enrolled in pre-school education. Compulsory primary education, beginning at age six with first grade, lasts for six years, leading to the General Elementary Education Certificate, after which children progress to intermediate education. This has important implications for this study, which involves children at both kindergarten and first-grade level. As of 2014, the average enrolment at kindergarten in Saudi Arabia was approximately 3% (*Al-Arabiya*, 2014).Therefore, most of the first grade pupils involved in this study will not have had the benefit of any kindergarten education and this may affect their levels of RRS.

It is important to emphasize from the outset that this research is focused on the testing of skills that are reading-*related* rather than on the skills of reading itself, since the two youngest age groups tested in this research had not yet begun formal literacy training. The RRS which are tested as part of this study are PA, LK and RAN and since these involve decoding rather than reading there is no measure of comprehension in the battery of tests.

The study of Islam forms a central part of the Saudi curriculum and, as part of this, children undertake *tajwid* (the study of Qur'anic recitation) from a very early age, a cultural factor that may have an impact on the way in which young children develop RRS and the rate at which they do this. This factor will be investigated as part of this research.

Educational policy in the Kingdom has received a large amount of attention in recent years, largely due to frustration concerning the fact that to date the Saudi education system has seemed unable to produce an effective workforce in many areas. This has forced recruiters to rely on workers from overseas, even though there are relatively high numbers of unemployed Saudi citizens (DGER, 2010).

Saudi educational policy has two key objectives: firstly, to ensure that the education system becomes more efficient at meeting the country's economic, social and religious needs and secondly, to eradicate illiteracy among Saudi adults, with some 12.8% of the population still falling into this category (World Factbook, 2013). The government agencies are responsible for the planning, administration and implementation of educational policy in the Kingdom. However, it is the Ministry of Education that sets overall standards for Saudi Arabia's educational system (both the public and the private sectors) right the way through from elementary level to higher education (DGER, 2010). The same general policies, curricula, and methods of instruction are followed by all schools, regardless of whether they are public or private.

Efforts by the Saudi Ministry of Education to reform the educational system have centred on the programme known as *Tatweer* [development], intended to shift educational methods in the Kingdom from the fact-based rote learning and repetition that has traditionally been used in much of the Arabic-speaking world, towards approaches to learning that are designed to be more analytical and problem-based (DGER, 2010). A large amount of funding has been invested in this programme, which that initially has been focused on teacher education and higher education resources. The intention is that this will have a trickle-down effect on teaching and learning strategies throughout all levels of the education system.

1.3.2 Tajwid (Qur'anic recitation)

It is useful here to add some clarification concerning what is meant by reciting the Qur'an and what takes place in these classes. Although the Qur'an now exists as a written text, it is important to remember that it was originally transmitted among the followers of the Prophet Muhammad as an oral text which had been committed to memory, as this was the usual means of cultural transmission among the tribes of the Arabian Peninsula (Boullata, 1989). As Stefan Wild notes: “The Arabic word *qur‘an* literally means ‘recitation’, and in one of the earliest *suras* [verses of the Qur‘an] the Prophet is admonished and chant the Qur‘an very distinctly’ (73.4)” (2006: 533).

This meant that from the earliest times, Qur‘anic recitation has been connected with a great dependence on memorization and on rote learning in order to ensure that the text was transmitted accurately and preserved in the exact form in which it was revealed to the Prophet Muhammad (Boyle, 2004). Along with this, those who transmitted the Qur‘an were also trained in the skills of articulation since this was judged necessary to ensure that, when Allah’s message was communicated orally, it was clearly understood and could not be misinterpreted by the audience. In addition, developing skills of auditory discrimination was important in those wishing to memorize the Qur‘an so that they could capture all the different elements that they heard in order to reproduce them faithfully.

As so-called “People of the Book”, Muslims place great emphasis on the Qur’an as a source of divine wisdom and moral guidance which also offers profound insights into the human condition (Haeri, 2011). Muslims believe that the Qur’an was verbally revealed to the Prophet Muhammad by the angel Jibril (Gabriel) and therefore the Qur’an is “revered by Muslims not only as a book of religious guidance and knowledge but also as a divine miracle of language, the very speech of God which no man can rival or surpass” (Fatani, 2006: 356-357). For these reasons, recitation and memorization of the Qur’an is believed to confer blessings on an individual, even if the individual reciting the Qur’an does not understand the exact meaning of what is being recited.

Consequently, over the course of the centuries, a whole discipline known as *‘ilm al-tajwid* (Qur‘anic recitation) evolved. As Nevad Kahteran explains, this discipline:

determines down to the most minute detail how the Qur'an should be recited, how each individual syllable is to be pronounced, the need to pay due attention to the places where there should be a pause, to elisions, where the pronunciation should be long or short, where letters should be sounded together […] and where they should be kept separate, and so on (2006: 635-636).

Clearly this level of detail is not taught to the kindergarten and primary school children who participated in this study but this explanation highlights the types of skills that the *muqri'* [tutor for Qur'anic recitation] is ultimately attempting to develop. It also makes it clear that initially, at least, less emphasis is placed on pupils learning to decode the Classical Arabic script of the written text of the Qur'an than on developing their memorization, articulation and PA skills.

Children learn Qur'anic recitation by imitating what the teacher recites and repeating this process many times with the teacher correcting them (Berglund, 2010). In the early stages, the exact meaning of the words being recited may be explained only in general terms and children are encouraged to focus on the beauty of words, "which come directly from God" (Berglund, 2010: 195). At kindergarten level, children are only required to memorise specified short passages of the Qur'an that they have listened to repeated by the tutor or heard from a cassette several times during class, sometimes without exposure to the written script. Later, at primary school, children listen and follow the text with their finger, as they hear the recitation and learn to match the sounds they hears to the letters and then words and phrases. As they grow older, they are introduced more systematically to the Classical Arabic script of the Qur'an.

Given the high regard with which Qur'anic recitation is held among Muslims, Saudi parents will often seek extra tuition for their children, either through private tuition or in lessons delivered within the mosque. It is this group of children receiving additional exposure to *tajwid* who were included as part of an exploratory study which focused on determining the extent to which the amount of exposure to *tajwid* affects performance in RRS.

A more detailed consideration of the links which may exist between exposure to *tajwid* and the development of RRS is to be found in the literature review (Chapter Two) and the discussion of the research findings in Chapter Five.

1.4 The Linguistic Context

1.4.1 Varieties Of Arabic and the phenomenon of diglossia

Arabic (العربية) belongs to a group of Semitic languages with consonantal orthographies that are known as *abjads* (Daniels and Bright, 1996). It is the official language for more than 200 million native speakers of Arabic (versteegh, 2014). **Figure 1.2** shows all of the countries with Arabic as their official or co-official language in blue colour, namely: Algeria, Bahrain, Comoros, Chad, Djibouti, Egypt, Eritrea, Iraq, Jordan, Kuwait, Lebanon, Libya, Mauritania, Morocco, Oman, the Palestinian territories, Qatar, Saudi Arabia, Somalia, Sudan, Syria, Tunisia, United Arab Emirates and Yemen.

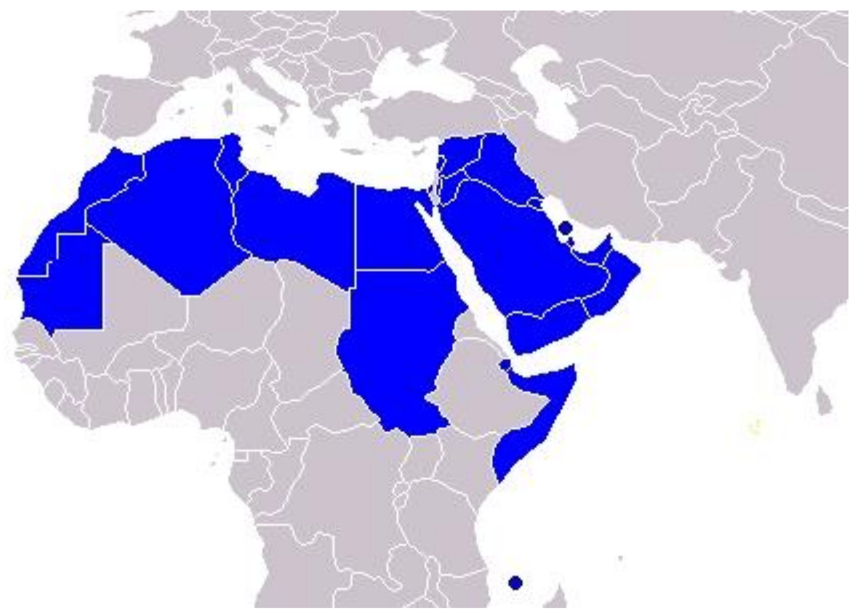


Figure 1.2: Countries where Arabic is the official language[1].

(Source http://www.google.maps.com)

Arabic also serves as the liturgical language of over a billion Muslims in communities all over the globe, making it one of the most widely spoken world languages (versteegh, 2014). Arabic is also one of the six official languages that are used within the United Nations and all of its sister organizations.

This geographical spread of the language from Northwest Africa to Southwest Asia has meant that over the course of time, Arabic has evolved into a great variety of regional dialects (see **Figure 1.3**) that differ very considerably from one Arabic-speaking country to another and often, even within the same nation state, different communities may use their own linguistic variant as a result of socioeconomic, cultural, or geographical differences (Holes, 1995).

The general form of Saudi Arabic forms part of the Eastern Arabic area and has remained closer to the Classical Arabic of the Qur'an, which had its origins in the dialects spoken by the tribes of the Arabian Peninsula. As noted above, due to the large number of migrant workers in Saudi Arabia, children are often exposed to different varieties of Arabic, particularly Egyptian with an estimated 300,000 speakers of this variant in the Kingdom (Ethnologue: online).

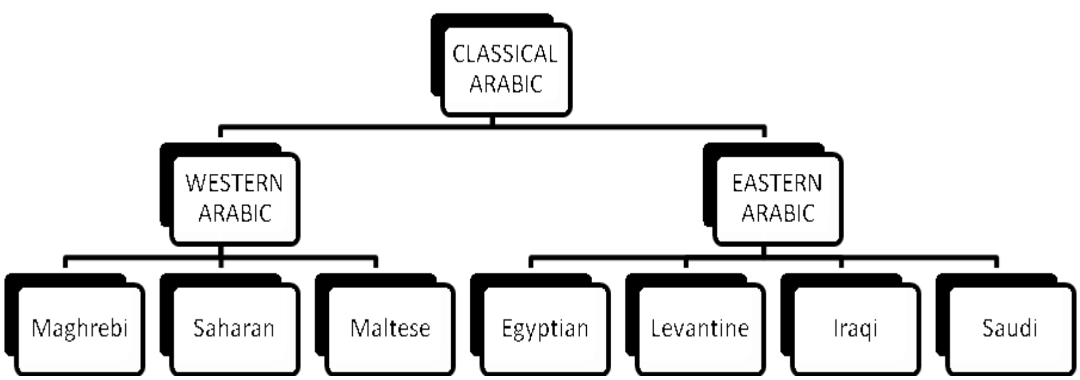


Figure 1.3: Varieties of Arabic

(Source: Al-Shboul adapted from Frías Conde, 2000).

As Saiegh-Haddad (2004) notes, in sociolinguistic terms, Arabic is known as a Diglossic language, a term that was first introduced by the sociolinguist Charles Ferguson in 1959. 'diglossia' is used to refer to a linguistic situation that includes four features:

(a) a differentiation between the written and the oral modes;
(b) a rigid socio-functional complementarity of two separate sets of functions performed by two different linguistic codes;
(c) a rich and dominant written literary tradition;
(d) linguistic relatedness between the two linguistic codes: the written and the spoken (Saiegh-Haddad, 2004:1).

Diglossia essentially means that there are two distinct varieties of a language and that each of these tends to be employed in different situations. Importantly, one of the two forms is considered to have a higher socio-cultural prestige than the other (Al-Batal, 1996; Haeri, 2000). The form of Arabic commonly referred to as Modern Standard Arabic (MSA), as the name suggests, is a formal written language, a uniform variant used as the official language in all the Arab states (Al-Toma, 1969). Most Arabic speakers learn this prestigious form through formal education and it is not generally employed for everyday conversation but used for most written and formal spoken purposes, for example by the media, in literary and academic discourse (versteegh, 2014).

Most Arabs would agree that MSA serves as "the rallying point for the intellectual, spiritual, literary and social activities incarnated in [...] Arab Islamic civilization" (UNDP, 2003:122). However, for most native speakers of Arabic, it is the non-prestigious, colloquial variant of the language which will form their principal means of everyday spoken communication and act as their "language of cordial, spontaneous expression, emotions, and feelings" (UNDP, 2003: 121).

In practical terms, in the case of Saudi Arabia, diglossia means that all educated Saudis will not only speak a colloquial form of Arabic (which will depend on the region of the Arabian Peninsula they inhabit) but in addition, they will also have learned MSA, which is the medium used for teaching in the education system. Moreover, as is the case for all Muslims, Saudis will also have been exposed from their earliest years to the language of the Qur'an, Classical Arabic. These multiple verities of Arabic have implications for children's development of RRS. The considerable differences that often exist, in term of morphology, vocabulary and phonology, between the spoken variant of the language that children are exposed to in the home environment prior to starting formal schooling (in this case Saudi Arabic) and MSA, the written form of the language that they will be exposed to and will be required to use when learning to read and write in the formal school setting, complicates the link between phonology and orthography. For example in MSA the word <*qamar*> is pronounced /*qamar(r*)/ (moon), while for speakers of dialectal variants this might be /*gamar(r)*/ or /*gumar(r)*/, suggesting a different orthography. Some phonemes found in spoken Arabic moreover, have no corresponding orthographic representation in MSA while other phonemes exist only in MSA (Saiegh-Haddad et al., 2011). MSA, thus, effectively acts as a second language for Arabic-speaking children and they may have poor PA when starting school, which in turn impacts on their phonological processing skills and their reading acquisition (Abu-Rabia, 2000). At a later stage, decoding skills can also be effected; Saiegh-Haddad (2003) found that when phonemic and syllabic structures were tested using words occurring in two different forms of Arabic, children made more decoding errors when presented with those that are unique to MSA.

As discussed later **section 3.8.3**, diglossia can pose some difficulties when devising a test battery for use with very young Saudis. It also highlights some of the potential problems to be taken into account in developing a standard Arabic reading test.

1.4.2 The Arabic alphabet and relations to phonology

This section introduces some of the key elements of the Arabic alphabet and writing system to provide an insight into some areas that may create problems for those at the early stages of learning to read and write this language, as were all of the participants in this research. MSA uses a 28-letter alphabet in order to represent in written form the 34 phonemes of the consonants and long vowels (/a:/ /i:/ /u:/) that are necessary for the written codification of the language(Taha, 2013). To be more precise, the script is made up of a total of 17 characters and dot-like diacritics, known as *i'jām* (اعجام), are added to these characters to create the 28 letters of the alphabet. These diacritics play a crucial role since they are employed to distinguish between characters, as seen in /ب/-/b/- bah and / ت/-/t/ teh and / ث /- /Ɵ/ theh as shown in **Table 1.2.**

ا a	ب b	د d	ذ d	ض ḍ
ف f	غ g	ه h	ح ḥ	ى i
خ j	ك k	ل l	م m	ن n
ق q	ر r	س s	ص ṣ	ش š
ت t	ث ṯ	ط ṭ	و u	ج y
ز z	ظ ẓ	ع .		

Table 1.2: Arabic alphabet with transliteration. (Source: Gent 2002:3)

In written Arabic, the three short vowels /a/, /i/ and /u/ do not exist as independent graphemes, being represented instead as a set of diacritics that appear above or below the character and are known as *harakat* (حركات).[2] Thus, in the case of /b/:

[2] A number of other diacritics are also used to indicate the pronunciation of consonants.

ب with diacritic *fat-h'a* pronounced as /ba/	بَ
ب with diacritic *kasra* pronounced as /bi/	بِ
ب with diacritic *dhamma* pronounced as /bu/	بُ

Table 1.3: The short vowel diacritics with */b/- /ب/*

Like other Semitic languages, Arabic is written horizontally from right to left, and unlike English, for example, it does not have any equivalent of lower or upper case letters. In addition, Arabic has a cursive script, meaning that many of its letters are linked together by means of ligatures. Of the total of 28 letters forming the Arabic alphabet, 22 of these must be joined to both the letters which precede them and those which follow them within a word. The remaining six letters must be joined to the letters that precede them only. The cursive nature of Arabic writing and this system of connection can initially make it difficult for inexperienced readers to distinguish the boundaries between words in Arabic text.

In addition, an Arabic letter can vary its written form depending on the letters that surround it and on the position that it occupies, whether at the start of the word (initial), in the middle of the word (medial) or at the end of the word (final).

- In the **Initial** position the letter links only to the letter which follows it.
- In the **Medial** position it may link to both to the letter that precedes it and the one that follows it
- In the **Final** position the letter links only to the letter which precedes it.

As a result, some Arabic letters can take on distinctive forms which to the untrained eye look very different (as illustrated in **Table 1.4**). The non-connecting or isolated form of the letter is usually the one which is first presented to children learning the Arabic alphabet and this is the form that was used with participants in this study for the tasks that involved letter recognition (i.e. LK and RAN).

POSITION IN WORD			
Final	**Medial**	**Initial**	**Isolated**
ـغ	ـغـ	غـ	غ
LETTER FORMS			

Table 1.4: The four written forms of the Arabic letter ghayn/ɣ/

1.4.3 Transparency of orthography in Arabic

Languages differ in the degree of consistency in grapheme/ phoneme relationship. Writing symbols that represent just one phoneme are known as transparent or shallow orthographies, whereas those containing graphemes that are pronounced differently according to the context of the word in which they occurs are deemed to be opaque or deep. This has implications for inexperienced readers. According to Snowling (2000: 63), the fact that writing systems vary in the "transparency of their orthographies" means that they also differ in the inherent level of difficulty that they present for young readers. There is some evidence that children learning to read languages with transparent orthographies demonstrate more rapid development in reading skills (Patel et al. 2004; Juel 1988; Fumes and Samuelsson 2011; Caravolas et al. 2012).

On the one hand, the fact that there is generally a one-to-one correspondence between the letters/diacritics and phonemes suggests that Arabic script is relatively transparent, which might help Arabic-speaking children who are learning to read. In theory, it should prove easier for them than it does for children learning to read languages with deeper (more opaque) orthographies, such as English. On the other hand, the ligatures in cursive writing described above mean that correspondences between letters and phonemes are not entirely consistent.

Even more significant for transparency is the fact that diacritics are not usually represented in written Arabic meaning that many Arabic words that appear in common everyday texts have identical written forms and if they are presented out of context, they are semantically and phonologically ambiguous (see examples in **Table 1.5**). Diacritics are thus added in those cases where ambiguity may distort understanding in texts that require careful reading, for example, religious texts. They are also marked on books and teaching materials aimed at young learners but as the learning process progresses (i.e. after 4th grade), short vowels are largely unmarked in most texts (for a thorough review see Abu-Rabia and Siegel 2002). Linguists refer to these two types of orthography in Semitic languages such as Arabic as being "shallow" or "deep". Shallow orthography is also known as vocalized or vowelized text and refers to the use of *harakat* to represent short vowels. When these diacritics are not added, this is referred to as "deep", non-vocalized or non-vowelized Arabic (Bentin and Ibrahim, 1996, Abu Rabia, 2001). As shown in **Table 1.5.**

Shallow Arabic	Deep Arabic	Transliteration	Meaning
عَلِمَ	علم	*a'alima*	He knew
عُلِمَ	علم	*a'ulima*	It has been known
عَلّمَ	علم	*a'allama*	He taught
عِلّم	علم	*a'ilm*	Knowledge

Table 1.5: Examples of shallow and deep Arabic orthography
(Source: Al-Shboul, 2010)

The progression of reading acquisition in Arabic, therefore, follows two stages: the first in which children can quickly learn to read a transparent orthography, and the second in which they must draw on a variety of other tools and information in order to establish the word in use. For the purposes of this study, however, Arabic is considered to have a transparent orthography, as the age range covered in this study (4;0-7;0 year-olds) means that the participants will not yet have been exposed to the orthography in its opaque form.

1.4.4 Word Formation in Arabic: The root system

Another characteristic feature of languages that belong to the Semitic family is the way they form all their verbs and most of their nouns by using consonantal roots, most of which are made up of a group of two or three letters. The examples shown in ***Table****1.6*: **1.6** illustrates how the consonantal root, in this case, /k-t-b/ /ب- ت-ك/forms the core of three words with different patterns of sets of vowels being inserted between these to form the individual items which are connected in meaning (Prunet et al., 2000).

كـتـب	كتـــــاب	كـتـب
katab **= 'he wrote'**	***kitaab*** **= 'a book'**	***kutub*** **= 'books'**

Table1.6: An illustration of a consonantal root.

(Source: Al-Shboul, 2010).

The consonantal root can be used to provide a vast range of meanings by employing two linguistic procedures. The first is to vary the way in which the simple root is vocalized (i.e. by adding different vowels), while the second involves using prefixes, suffixes and infixes.[3] Every single variation in this pattern of the same consonantal root is capable of producing a different meaning. It should be noted, however, that the individual components of the consonantal root remain in exactly the same order in any word that has been produced from the original root. Thus, for example, for the root k-t-b in addition to the words illustrated in **Table: 1.6** one could cite *maktab* (office); *maktaba* (library) and its plural *maktabat* (libraries). In other words, in an Arabic word, the consonantal root conveys a core semantic concept but each variation in the vowel pattern relates to a grammatical function, and is used to indicate, for instance, negation, verb tense and aspect, person, number, gender, case, and definite or indefinite (Holes, 1995).

[3] Arabic is a highly agglutinative language, meaning that suffixes, prefixes and infixes can be added to the word root. Affixed and suffixed pronouns, possessives, prepositions and conjunctions are also frequently used in Arabic.

This root system means that although the morphology of Arabic is complex, for the most part it is predictable. Once learners understand this is the case, identifying the root letters of a word and understanding the patterns they produce means they are able to form different structures following these patterns. Knowing this helps readers to pronounce words correctly and also to guess their meaning. However, inexperienced readers still need to put in a lot of effort to arrive at the exact meaning of an Arabic word. The regularity of the morphology helps to support this process of understanding and morphological knowledge is very important when learning to read Arabic as knowledge of related words (derivations) and of different forms of the same words (inflections) often provide clues to orthographically correct spelling. Elbeheri and Everatt (2007) note that, for this, reason morphological/orthographic knowledge may make an important contribution to the reading development of Arabic-speaking children and may also act as an additional source of individual differences in reading ability, and this has been confirmed in recent research by Tibi and Kirby, 2019 (see **Section 2.3**).

However, while non-vowelized script is clearly opaque and reliant on morphosyntactic and semantic knowledge, young children are exposed to shallow (vowelized) orthography when learning to read, so morphosyntactic and semantic knowledge are not required. The assessments developed in this study (PA, LK and RAN) target links between phonology and orthography which are the key skills for the early stages of reading single words and relatively transparent script. Theoretical models and evidence supporting this are considered in the next chapter.

2 CHAPTER TWO: LITERATURE REVIEW

2.1 Assessment Of Reading-Related Skills: Rationale and Evidence

It is well established that reading requires the mapping of orthographic units (the conventional spelling system of a language) and phonological units (working with the sounds of spoken language), a process known as decoding. However, theoretical models vary in their characterization of the decoding process, the nature of the orthographic and phonological skills involved and the nature of connections between these.

Languages also vary in the nature of connections between orthography and phonology. These differences have implications for the skills that underpin the decoding process. This chapter reviews different models of reading, including the simple view of reading model, dual route cascaded (DRC) model, triangle model, interactionist models and Frith's model. The chapter also considers the implications for foundation skills in reading development. It then critically reviews cross-linguistic research on relations between these foundation skills and decoding in children, and the influence of demographic factors. The languages focused on in this literature review include English with reference to other European languages, in which most studies have been conducted, and Arabic, which is the focus of this study.

2.2 Theoretical Rationale For Assessing RRS: Models Of Reading

2.2.1 The Simple View Of Reading Model

The Simple View of Reading (SVR) was developed by Gough and Tunmer (1986). As shown in **Figure 2.1** the components in the SVR model are two interdependent processes which, when put together, result in reading. The first component of reading is decoding. This is a key skill for learning how to read. Decoding can be defined as the ability to apply knowledge of letter-sound relationships in order to produce the phonological forms that are represented by a group of letters. It is possible, therefore, to decode these groups of letters whether they represent real words already known to the reader, words that the reader has not encountered previously or even invented words. The second component of reading, linguistic comprehension, refers to the ability to grasp the meanings that underpin written forms of language, rather than mere phonological processing (Hoover and Gough, 1990). This study focused on some of the RRS that are said to predict decoding skills only.

Although it is acknowledged that decoding and linguistic comprehension are not entirely independent processes, the SVR has proven to be valuable because it "has allowed a set of testable predictions" to be made (Hoover and Gough, 1990:157). This view postulates that reading consists of more than decoding, since someone may be able decode a text in a foreign language but still not understand what it actually means. Similarly, linguistic comprehension alone is also insufficient, because a five-year-old child may be able to comprehend spoken language fluently but still not be able to read texts in the same language. This approach focuses solely on decoding and comprehension and is not concerned with other skills and abilities that may also shape the acquisition of reading skills in young children, such as the role played by phonological skills.

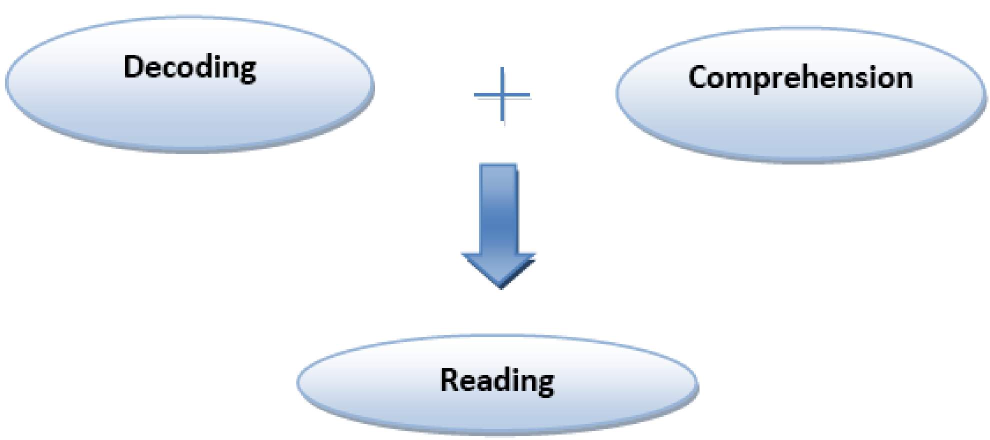


Figure 2.1: Simple View of Reading Model illustration
(Source: based on Gough and Tunmer, 1986: 6)

Figure2.1 illustrates the assumptions that decoding is simply not sufficient for reading; hence thus the need for linguistic comprehension. That is, in order for readers to be successful, linguistic comprehension must take place. In other words, reading (R) equals the product of decoding (D) and linguistic comprehension (LC), or R = D x LC. This naturally leads to the argument that the term comprehension has been selectively used in the SVR model to address certain aspects of reading engagement and the model is therefore limited in its application. This is because, comprehension has not been used in its broadest sense but recognized as the quality component in reading skills acquisition. It is, therefore, arguable that a child can satisfy the context of reading simply by decoding alone (pseudoword reading) in the beginning stages and gradually progress to comprehension in its broadest sense at later stage.

Decoding happens when children use their knowledge of letter-sound relationships to pronounce written words. This process involved in decoding consists of three distinct elements: the first is orthography, which is the characters that are used to create written form of a language; the second is phonological awareness, which concerns the sound and sound patterns formed by words within a language; and the third element combines both of these two other elements, with an understanding of how visual characters correspond to units of sound. In the English language, the orthography refers to units known as graphemes, while phonology is concerned with units known as phonemes.

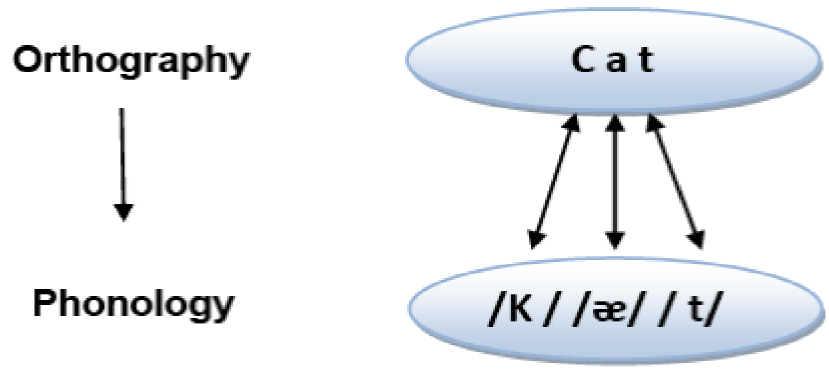


Figure 2.2: An adapted version of the word decoding model

(Source: based on Hulme and Snowling, 2009:47)

Figure 2.2 illustrates that the word "Cat" consists of three graphemes, each representing a particular phoneme. In order to read this word, first the reader must recognize the orthographic units, and by means of the mapping process, the orthography leads the reader to the phoneme sequence, namely, /k/ /æ/ /t/, and then onto the semantics or meaning of this word. This example suggests that decoding takes place in a number of stages and requires all the following components to be available:

1. Knowledge of letters or of their written shape (letter knowledge)

2. Knowledge of and access to spoken word representations, needed in order to understand how this word is pronounced e.g. /kæt/

3. Ability to identify the units or phonemes within spoken words that are represented by the written letters, in this case /k/, /æt/, /t/ (phonological awareness)

4. Ability to link letter knowledge to corresponding units of phonology using phonological awareness and to combine these to form the spoken word.

This SVR model implies that both awareness of phonological units (phonemes) and the knowledge of correspondence with orthographic units (letters) are critical for reading. The SVR model does not, however address the reading of words whose phonology deviates from grapheme-phoneme correspondence rules (i.e. irregular words), nor does it address reading in languages with non-alphabetic scripts. Both of these aspects of orthographic systems have implications for the contribution of phonological awareness. Invariably, there has been much contributed to the understanding of decoding since Gough and Tunmer's framework in 1986 and the ambiguity of decoding defined as successful word recognition has been recognized. Kirby and Savage (2008) addressed this issue by examining two separate measures of decoding as predictors of reading comprehension in SVR: (a) pseudoword reading, and (b) word identification. They discovered that when each was put together with listening comprehension, word recognition was a better predictor than pseudoword reading of reading comprehension, accounting for more obvious variance.

Furthermore, pseudoword reading indicates that one aspect of the decoding process relies on vocalising the orthography (i.e., letters and syllables) rather than whole-unit recognition. This indicates that decoding consists of earlier developing skills including PA in addition to the more traditional skills associated with decoding. Therefore, it is important to consider how decoding is defined and the fact that SVR may be incomplete if decoding is defined based on real word recognition, particularly for "less able readers" who are dependent on the earlier developmental skills of decoding. In addition, the decoding component of reading is more intense in younger children regardless of their varying abilities, which may decrease as they grow older (e.g Garcia and Cain, 2014; Vellutino, Tunmer, Jaccard, and Chen, 2007).

2.2.2 Dual Route Cascaded (DRC) Model of Reading

The DRC model is a combined model of visual word recognition and reading aloud (Coltheart et al., 2001). The model elaborates the decoding route outlined above, and includes an alternative route from orthography to phonology that does not involve awareness of sublexical phonological units and addresses the decoding of irregular words. The DRC model proposed by Coltheart et al. (2001), was developed to account for the reading of regular and irregular words and new words or non-words. As illustrated in **Figure 2.3**, readers can apply specific linguistic rules that govern the sound associated with individual graphemes (such as letters) and use these grapheme-phoneme correspondence rules to convert graphemes into their matching phonemes and construct the pronunciation (the indirect non-lexical or phonological route). This enables readers to process both regular words and non-words. Alternatively, they can employ a more direct system and retrieve the pronunciation for the whole word stored in their mental lexicon (the lexical or whole word route).

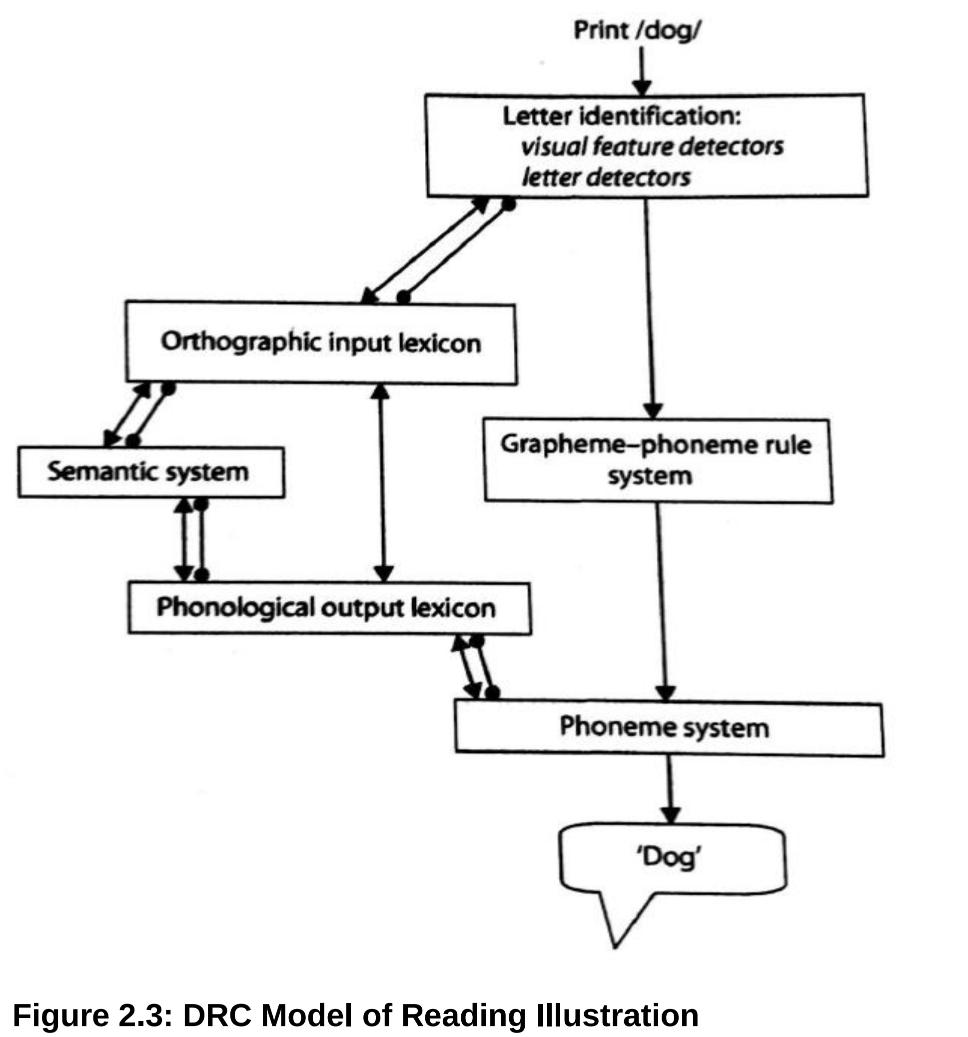


Figure 2.3: DRC Model of Reading Illustration

(Source: Cain, 2010 adapted from Coltheart et al., 2001, p.213)

The DRC model shares this fundamental assumption that the correct pronunciation of a word can be determined in two ways. In addition, the second fundamental assumption of the DRC model relates to the nature of lexical representation. According to Coltheart et al. (2001), both forms of words (the orthographic and the phonological) are represented holistically, functioning as discrete processing units within the mental lexicon. This means that the pronunciation of words that are already known can be obtained by mapping its graphemes onto the orthographic unit that offers the closest correspondence. This orthographic unit can then be used to directly activate the phonological unit that matches the pronunciation for that word.

In contrast to some of the other dual-route models, in the DRC model the indirect (or assembled) and direct routes operate in parallel, meaning that in most cases the pronunciation of a word is determined by the combined products of the two routes. In this way, Coltheart et al. (2001) attempt to offer an explanation for the degree of difficulty involved in reading different categories of words. Since the two routes operate in parallel, words classed as having regular pronunciation are pronounced more quickly and more accurately than those classed as being irregular because the assembled and the direct routes work together to produce reliable pronunciations for regular but not irregular words. Similarly, because the routes are activated more efficiently for graphemes found in frequently encountered words, they are likely to be pronounced more rapidly and with greater accuracy than infrequently used words.

Children are able to construct partial representations of words through a rudimentary phonological re-codification that allows them to recognize unfamiliar or pseudo-words, while rudimentary visual strategies help them to recognize those which are familiar. When facing new or unknown terms, children resort to the indirect or phonological route to translate the constituent letters of a word into sounds and apply phoneme-grapheme correspondence rules to identify the word and grasp its meaning (Bjaalid, et al, 1997, Ehri, 1992).

The dual route model implies that phonological awareness and letter-sound knowledge are critical for reading new words and non-words, and may play a role in reading real words. An important precursor of learning to read new words is described as implicit awareness, which is the ability to analyse words into their basic sounds at the level of the syllable or sub-syllabic unit. This ability can be distinguished from explicit phonological awareness or phonemic awareness, which is the ability to detect and manipulate phonemes within words, an ability that develops as a result of learning to read, and which enables the child to read non-words.

In contrast to new words or non-words, real words may also be read by the direct route, and this is critical for the reading of irregular words. Using the direct route to read words does not require phonological awareness. It does, however, require knowledge/representation of the word's phonology and access to this, so should be affected by the strength and speed of access to word phonology.

Similar to the SVR of model, the dual-route model also has its weaknesses. It is heavily influenced by English orthography which is alphabetic, with letters mapping onto phonemes, but has many sub-regularities and irregularities in the mapping between letters and sounds. English orthography is therefore highly opaque (see **section 1.4.3** on transparency of orthography), with complex mappings between sounds (phonemes) and letters (graphemes). Although some general 'rules' appear to be applicable (e.g. the digraph *ea* as in *seat*, is pronounced as a long /i:/), there are many exceptions to this (for example, *bear* with diphthong /ɛə/ and *fear* with diphthong /ɪə/). Conversely, a single sound such as /i/ may be represented by different digraphs, as in *team* and *teem*. While *g* is pronounced /g/ in isolation, in particular lexical contexts it may be pronounced /dʒ/ as in *giant,* or in combination with *h* may be pronounced /f/ as in *rough*, *cough*, or silent as in *high, bough*. The direct route therefore plays a critical role in decoding, even if it supported by the phonological route. Turning to Arabic, as noted above (**section 1.4.3**), the vowelized Arabic script to which children are exposed is highly transparent. The phonological (indirect) route is therefore sufficient for decoding during the early stage of learning to read (i.e. when vowelized script is used), and we would expect phonological awareness and letter knowledge to relate strongly to decoding. Given the parallel operation of the direct route, the strength and speed of access to word phonology may contribute to reading, even when the script is fully transparent. We might, therefore, expect both sets of phonological skills to be important for early literacy in Arabic. The indirect route is not, however, sufficient in later literacy when children are required to read the opaque, non-vowelized script (**see section1.4.3**). When the word is read in isolation, filling in of vowels requires direct access to the lexical item. Indeed, research with adults has found that decoding of single words is quicker for non-vowelized than vowelized script, suggesting that they are using a direct route to lexical phonology and that the presence of vowels delays this process (Schiff, 2012). In text, as pointed out above (see **section 1.4.3**), the vowels depend on the syntactic and semantic context of the word, so we would expect awareness at these other linguistic levels to influence decoding in the reading of non-vowelized text.

Languages also differ with respect to the size of the phonological unit involved in the link between phonology and orthography. In contrast to alphabetic languages, orthography in some languages is syllabic, mapping characters onto syllables (e.g. Japanese) or logographic, mapping characters onto whole words (e.g. Mandarin) (Defior, 2004). Theoretically, segmentation at the level of the phoneme is not necessary for learning grapheme-phoneme correspondences, and we might expect phoneme segmentation to be less important for decoding. Decoding skills might then be more dependent on the strength of phonological representations and access (see section 2.3). This would not apply to the vowelized script in Arabic, which is alphabetic with mappings at the level of the phoneme vowel or diacritic-consonant in vowelized text. In the case of non-vowelized script, letter-consonant correspondences provide some phonological information about the root but this must be supplemented by lexical and morphosyntactic information to access the phonology of the word fully.

2.2.3 Triangle Model of Reading

The triangle model of reading was developed to simulate how a child's brain learns and processes reading. This approach presented the researchers with an understanding of how the process of translating graphemes into their corresponding phonemes takes place and how the meanings of these words can be accessed as a result of interaction with written words and feedback on this. Various versions of this model have been developed with a focus on the connections between orthography and phonology and the accurate pronunciation of written words. The models developed by researchers such as Harm and Seidenberg (1999, 2004); Plaut et al. (1996), and Seidenberg and McClelland (1989) are usually referred to as the triangle models due to the form which these models take to represent this process.

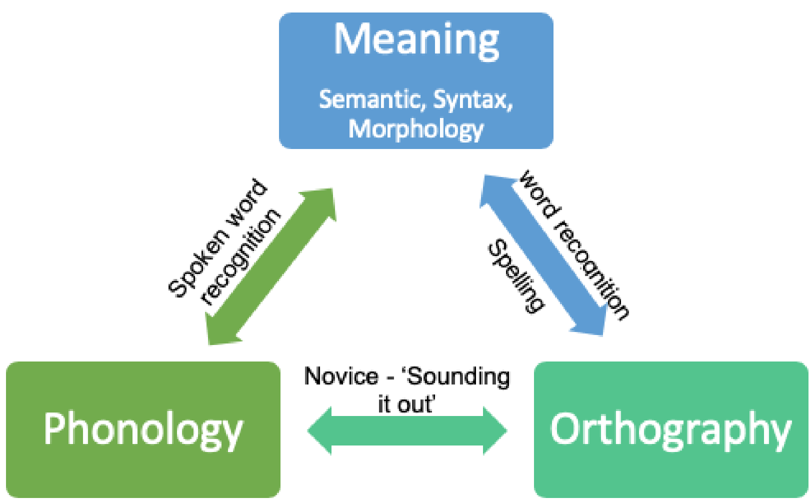


Figure 2.4: Illustration of the Triangle Model of Reading

Figure 2.4 illustrates Plaut et al.'s (1996) 'triangle' model in which three cognitive facilities interact to enable various language skills. This model is based on the idea that one of the most important early skills a child must master is "decoding", the ability to map grapheme to sound. The model shows that there are two perceptual paths to reading; the visual path or the sound path to the meaning (known as the whole word memorization method). The model is an attempt to explain reading, how reading should be done, and how children can be successful at reading. For example; to pronounce a string of letters, Seidenberg (2005) suggests that, when children read by decoding; they link orthography to phonology and can then use spoken word recognition.

Thus, decoding is important in early stages for much more fluent reading later, and there is evidence that proves its importance (Erhi, Nunes, Stahl and Willows, 2001). However, these connections between sound and spelling are of a complex nature in the English language. For example, /i/ is used to pronounce *ea*, as in *bead* and *neat*; however, there are many exceptions to the rule, including words such as *head and bear*. While the DRC model operates at the level of single word representations represented by discrete processing units in the lexicon, in the triangle model, orthographic input can take the form of representations of single letters or visual features of letters. Likewise, phonological input could mean a phoneme or a specific feature of a phoneme. In the latter model, these representations take the form of a pattern of activity that is distributed over multiple units and relies on the connections among these. Thus, this model hypothesizes that there is no one-to-one correspondence between an orthographic or phonological representation and an individual word; rather, lexical information is stored in a distributed manner as a pattern of activity across the whole network of units and contained in the connections that mediate between these.

Due to the fact that repeated experience with words not only creates connections but actually serves to strengthen them, triangle models hypothesize that frequently used words will be pronounced more quickly and with a greater degree of accuracy than their less frequently encountered counterparts. In a similar fashion, words with regular pronunciation patterns are likely to be pronounced more rapidly and accurately than those displaying irregular patterns, because the features that are common to the pronunciation of multiple words, create stronger shared associations than is the case for those mediating the pronunciation of irregular words which share fewer characteristics.

Since consistent links occur at different levels of phonology, from phonological features through to syllables or sequences of syllables, triangle models can account for decoding in languages with alphabetic, syllabic and logographic orthographies. They can also account for decoding of sub-regularities in orthography-phonology pairings, and exceptional pairings (i.e. irregular words) through connections between sub-lexical combinations from phoneme sequences through syllables to whole words. Triangle models imply that all levels of phonological processing will contribute to decoding.

With regards to the implications of single word decoding models for RRS assessment, the different models of decoding exemplified above all entail links between phonological and orthographic structures, but differ in the nature of the structures involved and the nature of the relations between them. Most notable are the differences between the DRC model and the triangle model. While the DRC model states that there is a route that processes information in a serial way, the triangle model posits that all processing takes place in a parallel way. The second difference is that DRC adopts the concept of 'entry' to describe the stored information about a word in the lexicon; triangle models do not represent words as a whole. According to triangle models, words are stored in the form of a distributed representation of their orthography, pronunciation and meaning and features that are common between words are shared. This might lead us to expect all levels of phonological structure with shared connections to orthography to be important in decoding. The DRC model, explicitly including two routes to a word's pronunciation, would lead us to expect letter-sound knowledge and strength of lexical phonological representations of words to be important for decoding.

However, the models do not themselves provide a rationale for the tasks used to assess the relevant skills, whether these target implicit or explicit awareness (see

section 2.2.2), and related to this, whether the skills assessed contribute to the bottom-up (sensory) or the top-down (cognitive) strategy of information processing. The motivation for the assessment of RRS, including the phonological structures targeted and the methods employed, stems from empirical evidence of relations between performance on these tasks and performance on decoding tasks. Before considering this empirical evidence (see **section 2.3**), we turn to the motivation for investigating the influence of environmental factors as well as the cognitive skills highlighted by models of single word decoding

2.2.4 Interactionist Models Of Reading (Environmental Factors)

Interactionist models emphasize that early childhood is a critical period for language acquisition. This model theorizes that, if children are not exposed to language at an early age, then they will be unable to develop full linguistic fluency in any language. This theory leads to the consideration of the role of the socio-environment in language development. From this perspective, it is not just the case that a child has a genetic predisposition for a language, which enables it to emerge; the child needs the correct environment to realize the full potential. Such an environment is created when children interact with their peers and adults in the community in which they are situated, with teachers playing a key role in this. The model "assumes that language acquisition is influenced by the interaction of a number of environmental factors such as, physical, linguistic, cognitive, and social," (Cooter and Reutzel, 2004). Therefore, the interactionist model recognizes that children have a predisposition for language acquisition but that socio-environmental factors play an important role in a child's literacy development. The model views literacy as being socially acquired and requiring specific skills and verbal interaction. The socio environment also includes specific strategies of teacher intervention in the classroom, based on verbal interaction.

According to Rogoff's model of guided participation (1993), literacy is a sociocultural instrument that requires specific skills gained by taking part in everyday experiences or actively observing these. Learning happens in contexts in which the individual interacts with their community by means of guided participation. Borzone and Signorini's (2002) socio-cultural framework proposes that teaching interventions should be based on the Vygotskian concept of potential development defined as the distance between the actual developmental level, i.e what the child can achieve working alone and the level of potential development, i.e. what the child can acheive under adult guidance, or in collaboration with more capable peers' (Vygotsky, 1991[1935]: 133). Vygotsky's socio-cultural model posits that children's cultural development is visible in two stages. First, the child observes the behaviour and interaction between other people. Secondly, a child learns by interacting with those around him/her. The child moves from baby talk to correct sentence formation through teacher intervention. Teachers and the strategies they use for verbal interaction thus play a prominent role in literacy acquisition. The models of Nelson (1996) and Tomasello (2000) both draw on two central Vygotskian concepts:

a) mental development is the result of a natural and cultural development;

b) the process of formation of higher mental functions develops most during early childhood, that is between the ages of two and six years. Nelson (1996) also stresses that social interaction plays a fundamental role in the early stages of the process of acquiring the ability to recognize written words.

As noted above, the interactionist model emphasizes the importance of children's social and verbal interaction with those in the family and their immediate environment for acquiring literacy skills. Beginning with the home, there are two key aspects important for the acquisition of reading skills. Firstly, the educational level of family members, particularly parents, since they transfer their knowledge of language, linguistic functions and reading to other family members, influencing children's abilities, literacy strategies and cognitive skills. Secondly, the home literacy environment including the quantity and quality of reading material in the home, what different types of written discourse children are exposed to (Scribner and Cole, 1981) and by whom (Porta, 2008). The socioeconomic status on child's family can also impact on children's cognitive and linguistic child development. Children of lower SES display lower levels of phonological awareness, letter naming, word writing, word recognition, and vocabulary (Bowey, 1995; Korat et al., 2007; Lundberg, Larsman, and Strid, 2012).

Interactionist models are also interested in education policies, since they influence pedagogical strategies and styles of verbal interaction in the classroom and have major potential to impact on literacy acqusition by using particular types of interventions, such as stimulating PA and letter-sound and letter-name knowledge for kindergarteners and first graders (Bombini, 2008).

2.2.5 Frith's Biological, Cognitive And Behavioural Perspectives

After a decade focusing on cognitive aspects of reading development, as a clinical psychologist with a special interest in areas such as dyslexia and autism, Uta Frith (1995) set out to develop a more comprehensive account of the various factors that can affect the development of reading, including biological, cognitive, behavioural and environmental factors. Since, as Frith observed, "the centrality of phonology in the acquisition process is asserted by all current models of cognitive psychology" (Frith, 1995: 10), her model was also intended "to provide the beginnings of a systematic tool for assessing phonological skills" (Frith, 1995: 14). It is argued here that this cohesive framework provides a more integrated and complete theoretical underpinning for understanding phonological and reading-related phenomena.

Frith's (1995) theoretical framework is intended to describe how PA skills develop and how an individual's performance in phonological processing tasks can be affected by a diverse range of internal or external factors. It also serves to locate any underlying difficulties with these skills within a particular level and forms the conceptual basis for the PA battery, known as PhAB (Frederickson et al. 1997). See **Figure 2.5.**

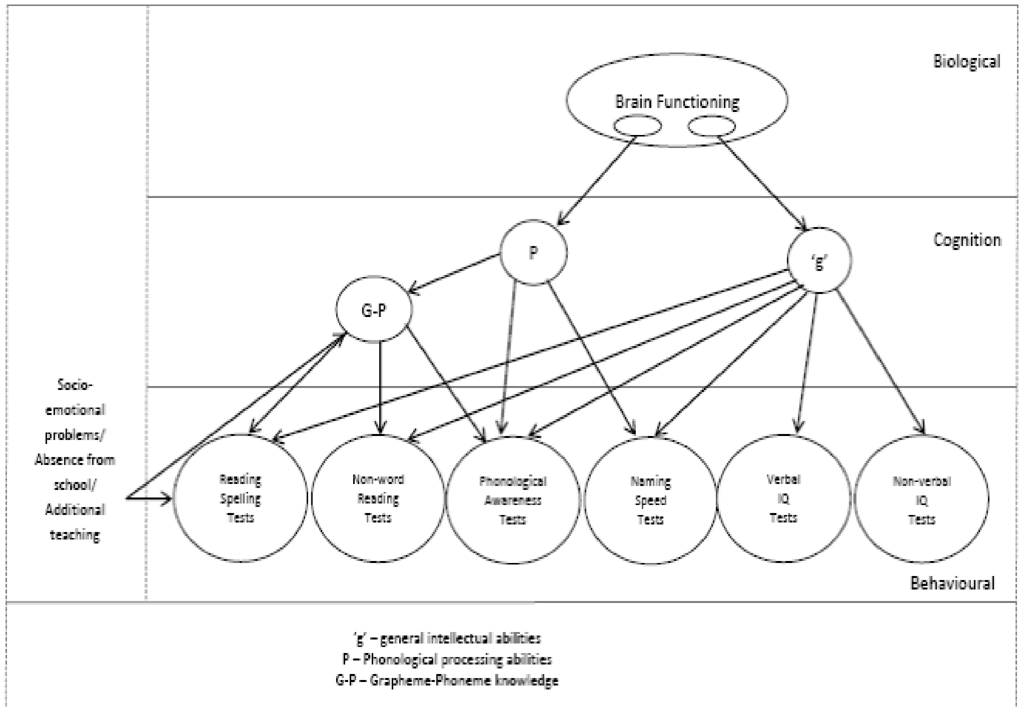
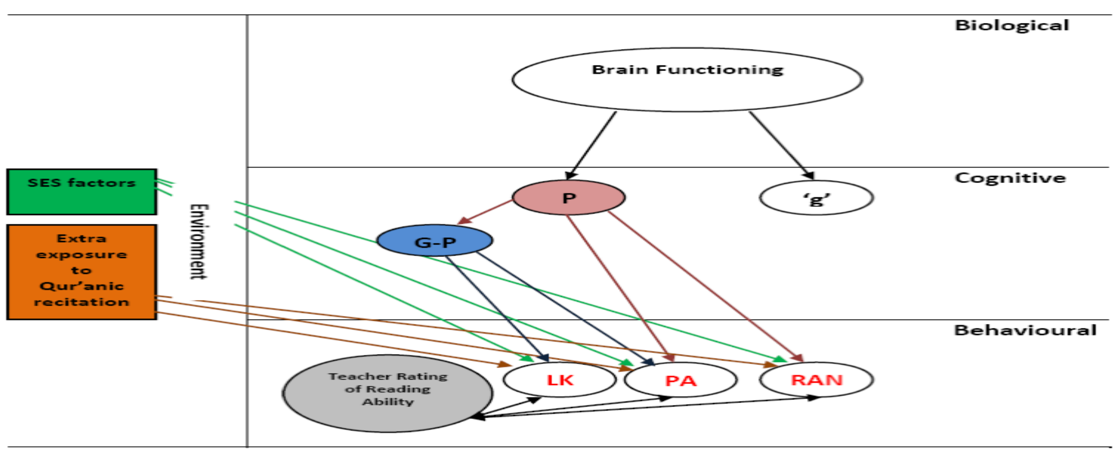


Figure 2.5: Frith's (1995) theoretical framework and conceptual basis for the Phonological Assessment Battery (PhAB).

Frith's model considers three cognitive modalities, which she identifies as general intelligence ('g'), phonological processing awareness (P), and grapheme-phoneme knowledge (G-P). These cognitive modalities are considered to act as mediators between the biological (i.e. genetic) basis for reading aptitude, on the one hand, and the behavioural aspects or observed reading capability as measured by various forms of tests, on the other hand. Frith's model suggests how different types of tests (including PA and other RRS) tap into different cognitive level skills, allowing these to be observed indirectly from an individual's performance at the behavioural level.

Prior to Frith's framework, measures of general intelligence levels were rarely incorporated into the scientific understanding of the development of reading skills with other, more specifically reading-related factors usually being highlighted, However, she argues that it is necessary "to have a gross measure of general ability at the very least in order to help us to interpret performance" (Frith, 1995: 15). This may be an important consideration when developing the final version of a test battery for identifying Saudi children at risk of developing reading-related difficulties. However, since verbal and non-verbal IQ tests already exist for Arabic-speaking children, the focus in this research is on developing and assessing age-sensitive materials that can be used to test individual performance in specific reading-related skills at the behavioural level and to tap into these skills at the cognitive level.

Frith's model also takes into consideration the impact that environmental influences can have on the biological, cognitive, and behavioural dimensions of reading ability, as well as the relationships between these dimensions, as indicated by the arrows in the model (**Figure 2.5**). Frith thus signals that environmental factors can play a significant role in the measurement of reading capabilities, because even if a child has no problems at the level of biological functioning and has properly trained cognitive capabilities, he or she may have an emotionally disturbing home environment or a socially oppressive school environment that transforms the act of learning to read into a difficult task.



KEY:
P = Phonological processing abilities
'g' = general intellectual abilities
PA = Phonological awareness
SES = Socio-economic status
G-P = Grapheme-phoneme knowledge
LK = Letter knowledge
RAN = Rapid automatized naming

Figure 2.6: Model for this research adapted from Frith (1995).

It is argued here that one of the key strengths of Frith's conceptual model lies in its holistic approach to identifying a wide range of possible variables that may influence individual progress, alerting teachers and practitioners to the multiple interconnected factors that may impact on children's literacy development, including environmental factors. This makes this conceptual framework applicable to different cultures. It should be noted that as part of this research, data were originally collected on a wide range of environmental factors that have been highlighted as affecting development of literacy skills, including participants' exposure to languages other than Arabic and their home literacy environment. Decisions needed to be made to limit the scope of this work to a smaller number of these areas, as shown in **Figure 2.6**. Social and economic background covered multiple variables that have been investigated in previous research on reading-related skills in largely Western countries but very little work had been done in this area in Saudi Arabia or its neighboring states (see **section 2.5.3 and 2.5.4**). The decision was also taken to focus on a culturally specific aspect of the Saudi curriculum, namely, Qur'anic recitation, since some claims had also been made regarding the role that this may play in developing particular reading-related skills (see **section 2.5.5**).

A further advantage of Frith's model is that it links together PA and reading capability, a connection that has not been explored adequately in previous models. This is accomplished in Frith's model (1995) by pointing out that PA, or the familiarity with different sounds, is distinct from grapheme-phoneme awareness, or knowledge of written syllables. In fact, she argues that these capabilities are modular, meaning that one can be trained without significantly influencing the other.

It should be emphasized here that the purpose of this study is not to evaluate Frith's model. Rather the aim here is to provide a framework/context that captures key cognitive and environmental influences in the development of reading, as a foundation for developing an assessment of these skills in Arabic. Ample evidence of the importance of these factors is provided in the next section.

2.3 Rationale For Assessing Reading Related Skills (RRS): Empirical Evidence And Methods

Much of the research literature since the late 1970s, concerned with literacy and the maturation of reading ability in young children, mainly studied the significance of RRS in the early stages of reading skill development. These studies were particularly interested in determining whether the use of RRS could identify in the early stages the children at risk of developing reading difficulties[4].

This section begins by reviewing relevant studies on reading-related skills (RRS) – phonological awareness (PA), letter knowledge (LK) and rapid automatized naming (RAN)– examining how these three RRS relate to reading ability in different orthographic systems including Arabic, and the possible causal or reciprocal relationships amongst them. The focus then shifts to studies which have attempted to determine the possible links between demographic factors, PA and reading development. These factors are gender, socioeconomic status and age. Consideration is given to the reasons for the relatively limited amount of research in the Arab world concerning the influence of environmental factors on PA and reading development, and research on one such factor, Qur'anic recitation, is examined.

2.3.1 Phonological Awareness (PA) and reading Skills

PA has been defined as "the ability to reflect on the phonological properties and structure of words independent of meaning" (Stackhouse and Wells, 1997; Hatcher, Hulme, and Ellis, 1994). It includes the awareness of speech units

[4] Although this thesis is primarily concerned with the RRS of typically developing children, the relationship between PA and reading ability also has important implications for identifying children at risk of developing literacy and language-related problems, specifically dyslexia. According to the British Dyslexia Association (2007: online) dyslexia "is characterized by difficulties with phonological processing, rapid naming, working memory, processing speed and the automatic development of skills that may not match up to an individual's cognitive abilities". Phonological theory has informed debates about the potential causes of this learning disability (Snowling, 2000).

(sounds), referred to as phonemes, as well as other units such as syllables. Phonological awareness is a process by which children understand the sound structure of spoken words and observe those elements of spoken language without the meaning of the word (Mattingly, 1972).

Gral-Azulay (2015) argues that the units in phonological awareness form a hierarchical structure termed 'Supra-phonemic'. This term refers to the consciousness of segments (i.e. syllables, onset-rime) larger than phonemes. Notwithstanding, all three units of phonological awareness are widely accepted and researched (Goswami and Bryant, 1990; Treiman and Zukowski, 1991). As shown in **Figure 2.7**, the levels include: the syllable level, which refers to the ability to recognise, for example, that the word 'target' consists of two syllables: /tɑ/ and /gɪt/ (Treiman, 1993); the onset-rime level in which the onset comprises the initial consonant(s) and the rime comprises the vowel and any following consonant(s); for example, children understand that the syllable /gɪt/ can be further divided into an onset /g/ and rime unit /It/ (Goswami and Bryant, 1990): the phoneme level, which is the individual sounds in words, for example, children grasp the concept that all words consist of sequences of phonemes (Treiman and Zukowski, 1991; Gillion, 2012), for example, /t/-/ɑ/.

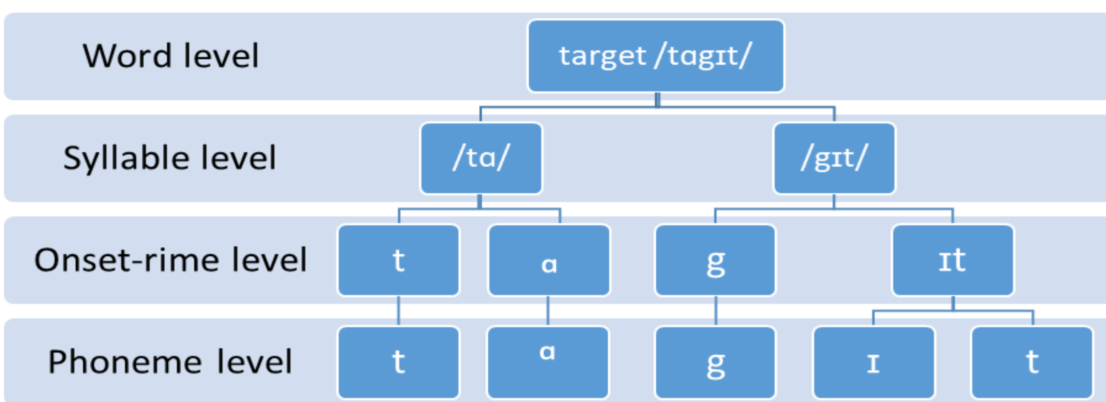


Figure 2.7: Hierarchical structure of the word 'target'

When discussing the levels of PA, it is important to distinguish between the two terms: 'rime' and 'rhyme'. Ziegler and Goswami (2005) explained that the term 'rhyme' is "used to refer to judgements about phonology, and to the phonological unit in any word following the onset (e.g., r-abbit, t-opic)"; and the term 'rime' is used when the phonological unit refers specifically to the division of a single syllable (e.g., s-eam, str-eam)."

Studies have shown that while awareness of supra-phonemic segments appears to develop spontaneously, this is not the case with phonemic awareness. Gullberg and Granholm (2010) and Xhafaj (2011) all agree that phonemic awareness seems to involve a more complex set of skills which are closely related to word recognition skills that underpin reading ability (Perfetti et al., 1987).This is because isolated syllables and larger segments are manifested as discrete units of speech, while phonemes are not (Kawamoto et al., 2014). Children need to be explicitly instructed about the rules of alphabetic writing to identify individual phonemes. As discussed in **section 2.2.2** on the DRC model, the concepts of implicit and explicit skills are directly linked to phonological awareness and play a significant role in the development of writing and reading skills. Several studies have confirmed that children who are good at making judgments about similarities in rhyme and initial sound (often referred to as rime-onset) make better-reading progress than those who are not (Hoien and Lundberg, 1988; Lundberg et al., 1988). These studies demonstrated that the ability to analyse words into their basic sounds at the level of the syllable or sub-syllabic unit could be considered an essential sign of learning to read.

Furthermore, in terms of reading skills, a recent study by Al-Sulaihimi and Theo (2017) has focused on defining the link between phonological awareness, emergent literacy, and reading success. The study discusses explicitly the building of PA among children who speak Arabic. The study was based on a longitudinal design where the results suggest that literacy training plays a significant role in improving PA skills. Another study by Elsaad, Ali and El-Hamid

(2016) has also assessed the development of PA skills among Arabic students. The study investigated the relationship between PA and word reading abilities. The results of showed a positive and robust relation between PA and proficiency in word reading abilities. Schiff and Saiegh-Haddad (2018) sought to determine the development and relationships between PA, morphological awareness and word reading in spoken and standard Arabic. This study briefly discussed the development of foundational metalinguistic skills and word reading skills in Arabic, and their relationship. The study compared Arabic children's PA, voweled and unvoweled reading skills and morphological awareness reading skills. The findings showed that PA directly affected the voweled and unvoweled reading among Arabic children, whereas PA also contributed towards reading fluency. Another study conducted by Assad and Eviatar (2014) explored the contribution of visual abilities, phonological awareness, accessibility of letter names and text reading speed and accuracy of Arabic. Assad and Eviatar (2014) recruited participants from the first grades (mean age= 7.02 years), third grades (mean age= 8.94 years) and fifth grades (mean age= 10.88 years) to explore the connection among the aforementioned factors. The results showed that phonological awareness contributes greatly to reading accuracy (Assad and Eviatar, 2014). The results from various studies included indicate the impact and importance of PA in the development and improvement of reading skills.

2.3.1.1 PA and developmental progression

The degree of difficulty in PA skills differs in nature and appears to follow a developmental progression. Many studies have been undertaken to establish the developmental sequence of PA in children, which begins with awareness of the larger identifiable units of words, syllables and onset-rimes, and then proceeds onto smaller phonemic units (Goswami and Bryant, 1990; Gillon, 2012). Lieberman et al. (1974) investigated whether children progress from the larger unit to the smaller unit using a sample of preschool, kindergarten and first grade children (N=135) who had to listen to a word, repeat it and then tap out the number of syllables or phonemes. Results indicated that all three age groups found it easier to identify syllables than phonemes. Moreover, none of the preschoolers were able to successfully complete the phoneme task, compared with 70 percent of the first graders. This suggests that syllable awareness develops before phoneme awareness. Research findings reported by Maclean Bryant and Bradley (1987) revealed similar outcomes with young children demonstrating PA knowledge at the rhyme level. The research was on British children (n=60, average CA: 3;4) using a rhyming task (three stimulus items). All children were assessed on their ability to carry out the rhyming task - to name the words and then state which of those words did not rhyme. In line with previous findings on the earliest stages of phonological awareness, Lonigan et al. (1998) found that approximately 25% of participating children in their two-year-old age group performed at higher than chance level on a rhyme task, evidencing that some very young children develop PA knowledge of large phonological units.

Further insight into developmental progression in PA was provided by Yopp's (1988) study. Yopp's study investigated the relative difficulty of a range of PA tasks among kindergarten children (N=96). Participants were tested on rhyme recognition, auditory discrimination, phoneme blending, segmentation, reversal and deletion, word-to-word matching, sound isolation, and counting tasks. The results from Yopp's investigation suggested that there was a hierarchy of complexity in PA tasks, in which rhyming was the easiest and phoneme deletion the most difficult. However, Yopp's study did not control for the level of linguistic complexity of the tasks, which was an important source of variability.

Stahl and Murray (1994) used rhyme recognition, syllable and phoneme manipulation tasks with a sample of kindergarteners and first graders (N=113) to re-examine Yopp's findings. However, they also assigned a weighting for the level of linguistic complexity, based on the number of syllables and type of stress pattern involved. Results for the PA tasks indicated that syllable and rhyme awareness preceded phoneme awareness, confirming previous findings on PA that children distinguish larger units before smaller ones. PA of onsets and rimes was found to be essential for more complex levels of phonemic analysis and for word reading.

Stackhouse and Wells (1997) argued that PA development should be viewed as a multilevel construct with two dimensions, one representing the size of the linguistic unit being processed and the other the level of explicitness of the task to be performed (see **Figure: 2.8**).

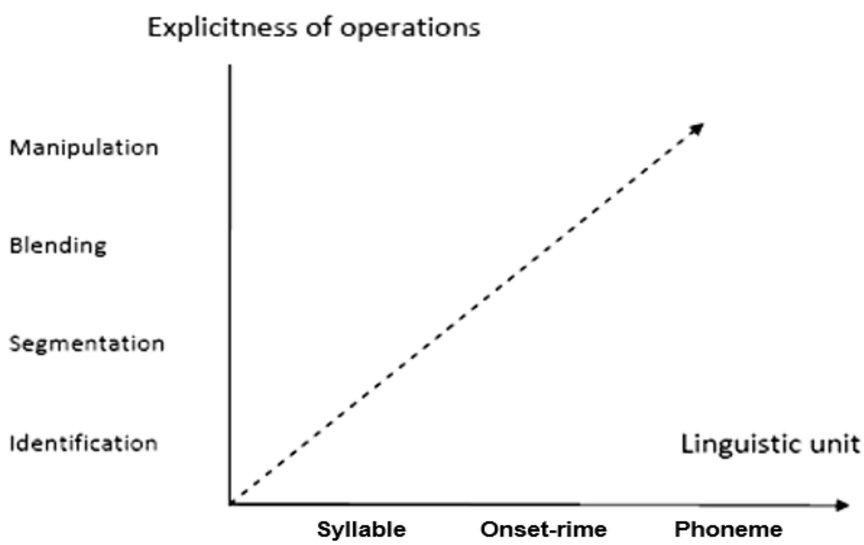


Figure 2.8: Development of PA in preschool children.

(Source: Adapted from Stackhouse and Wells, 1997).

The figure illustrates how awareness begins with the largest sub-lexical unit, the syllable, progressing to the onset-rime and finally the phoneme. Explicitness refers to the type of metalinguistic reflection needed to complete a particular PA task and is also closely linked to the level of cognitive complexity involved: the more explicit a task is, the greater the degree of cognitive processing it entails. The least explicit PA operation is considered to be identification, while the most complex involves manipulation of linguistic units. In **Figure 2.8** the diagonal arrow, shown as a broken line, represents the typical pattern of PA skill development in pre-school children who progress from the less cognitively demanding PA skills e.g. identifying syllables to the more demanding tasks of manipulating phonemes.

There are some studies which show similarities in empirical evidence for a universal sequence of PA development across languages, particularly languages using alphabetic writing systems (e.g. Cossu, Shankweiler, Liberman, Katz, and Tola, 1988, Schaefer, 2009).

Arab-Moghaddam and Sénéchal (2001) divided PA skills into shallow- and deep-level tasks. The former consisted of activities such as syllable segmentation, identifying word boundaries, and generating rhymes, while the latter included the ability to isolate and manipulate individual sounds, skills which typically develop later in children. They argued that while training in shallow-level skills can provide the foundation for development of deep-level skills, this is not sufficient to impact on reading or spelling abilities. In the context of the Arabic language, Tibi (2010) examined a developmental hierarchy of four PA tasks. Four different PA tasks were administered to children from grade one to grade three. Results indicated that the four phonological awareness tasks ranged from easy to difficult in the following order: rhyme, initial sound identification, syllable deletion and phoneme segmentation. The findings of previous studies on the English language were supported by their study with regard to the hierarchical order of phonological awareness development. This means that when phonological awareness tasks are trained, they should be considered in a non-random order. The phonological developmental hierarchy of Arabic phonological awareness skills is complicated and places demands on cognitive processes and, therefore, these should not be considered randomly.

2.3.1.2 The concurrent and longitudinal / predictive relationship between PA and reading skills

Studies have shown that PA is a major factor in early reading achievement (Ehri et al., 2001). Thus, when a child has a deficit in PA, it is likely that the deficit will cause reading disabilities (Lyon, Shaywitz, and Shaywitz, 2003). The report prepared by the National Institute of Literacy also supports this prediction (Kennedy et. al., 2008), and there exists extensive evidence to support the relationship between PA and children's word reading skills. Many studies have highlighted the association between PA and the acquisition of literacy. These studies assess PA by means of a variety of tests, with particular emphasis on tests of rhyme and phoneme awareness. The evidence often appears in the form of *concurrent* correlations. In these studies, PA and reading were measured at the same time, whereas studies on predictive correlations have assessed PA at an early stage then assessed reading skills later. However, there are very few studies that have examined the relationship between syllable awareness and word reading. A study by Engen and Hoien (2002) tested 1300 Norwegian children that were in their first grade (age range: 7;5 – 8;5 years) on measures of PA, word recognition, and reading comprehension. Though Engen and Hoien (2002) found a unique variance of syllable awareness in word recognition and reading comprehension, the impact of syllable awareness on reading skills was low in comparison to the impact of phoneme awareness. However, in a review study on Spanish-speaking children, dealing with a language whose orthography is regular and has a well-defined syllabic structure, unlike English whose orthography is irregular, Denton et al. (2000) suggest that syllable awareness may have a noticeable effect and serve as a better predictor of language reading ability. However, they acknowledge the diverse findings in other studies relating to the impact of syllable awareness on reading.

There are also studies that discuss the extent to which rime awareness and phonemic awareness differ in relationship to the development of word reading skills. In this area, a number of concurrent correlational studies agree that phonemic awareness is a better predictor of children's word reading skills than rime awareness (See Bowey, Cain, and Ryan (1992) on English-speaking children from ages 7.36 to 9.3 years; Nation and Hume (1997) on English-speaking students in 2^{nd} – 4^{th} grade and Mann and Foy (2003) on English-speaking children aged 4-6 years old, Bowey, Cain and Ryan's (1992) 3-group (n=48) reading-level design confirmed the findings of the above-mentioned studies. Their study probed the relationship between PA and reading performance of fourth grade readers. They assessed children’s level of sensitivity to syllabic and phonemic units and pseudo-word reading. The study found that English-speaking children with underdeveloped skills in reading performed more poorly than older children.

Furthermore, another study on 99 American pre-school children aged between 4 and 6 conducted by Mann and Foy (2003) found that phoneme awareness showed a stronger correlation with early reading measures than rhyme awareness.

Beside studying the concurrent correlations of PA and reading skills, other studies have explored the existence of longitudinal predictive relationships. These studies examined the ability levels in tasks that measure the individual components of PA and the development of reading skills in various language contexts. In the study conducted by Hogan et al. (2005), a sample of 570 English speaking kindergarten children were given measures of PA and letter identification. Measures of phonetic decoding (non-word reading) and word reading were subsequently introduced in grades two and four. The findings of this study revealed that PA assessment predicted grade 2 word reading. However, by that level, PA and word reading were highly correlated, such that PA does not add information about later grade 4 reading. Instead, it was the non-word and word reading tasks which were identified as higher predictors of grade 4 reading ability. In a longitudinal study by Bryant et al. (1989), which investigated 64 British children's knowledge of nursery rhymes, rhyme detection ability, phoneme awareness, reading, and spelling skills, children's knowledge at age 3 was related to rhyme awareness at age 4;7, and strongly correlated to phoneme awareness at age 5. Furthermore, rhyme detection was stronger in predicting reading and spelling skills at age 6;3. In addition, the longitudinal study by Juel (1988) of 54 American children from first grader through fourth grade found that those students who had problems with PA tasks (blending, segmentation and manipulation tasks) continued to experience reading difficulties as they grew older and progressed to fourth grade. Moreover, the review study by Melby-Lervåg et al. (2012) on children's reading ability supported the three measures: phonemic awareness, rime awareness and verbal short-term memory, as longitudinal predictors of individual differences in reading development. The review concluded that these measures differ in terms of their predictive relationships with word reading skills. According to Melby-Lervåg et al. (2012), phonemic awareness was significantly concurrently associated with word reading skills, while a weak correlation was found for rime awareness and verbal short-term memory.

In conclusion, the evidence from the studies of Hulme et al. (2002), Muter et al. (2004) and Muter et al. (1998a-b) supports the argument that measures of phonemic awareness are more powerful concurrent and longitudinal predictors of children's reading skills than rime awareness in the early stages of reading. All those concurrent and longitudinal correlational analyses were consistent with the view that phonological analysis skills contribute significantly more to children's decoding abilities. Thus, PA is often used to assess children's reading skills in many ways (see **section 2.4.1**).

2.3.1.3 Relations between PA and decoding in opaque and transparent orthographies

The studies reviewed above found that PA is a powerful predictor in the assessment of children's reading skills. However, as a complex skill, which requires both cognitive and linguistic abilities, learning how to read also differs somewhat according to the language in which individuals are learning (Ziegler and Goswami, 2005). Learning to read, particularly in Alphabetic orthographies, requires the explicit ability to analyse, segment and map speech at the level of the phoneme (Ramirez et al., 2010; Sawyer and Fox, 2012). Since alphabetic orthographies map phonemes of spoken language, phonemic awareness is a primary requirement for word decoding in these languages (Saiegh-Haddad and Geva, 2008). Research shows that decoding in transparent alphabetic orthographies is faster in comparison to English (Seymour et al., 2003), while the predictors of reading skill remain the same (Caravolas et al., 2005).

Extant literature is dominated by studies on the English language. However, these studies provide evidence that the ‘transparency’ of a writing system (orthography) plays an important role in reading and specifically the role of PA. The level of difficulty in children’s learning to read heavily depends upon the question of how the letters of the language can be mapped to speech sounds (phonology). As stated by Seymour (2003), a word in those languages with a transparent orthography is pronounced based on its spelling. These languages have been demonstrated to pose fewer challenges for those learning to read.

According to Katz and Frost (1992), the role of PA varies in languages of different orthographic depth. Their orthographic depth hypothesis argues that transparent orthographies tend to support the word recognition process, since they involve the alphabetic process, whereas opaque orthographies involve the formation of both alphabetic and logographic foundations (Katz and Frost, 1992). There is some evidence that children learning to read transparent languages demonstrate more rapid development in reading skills. This evidence can be found in Patel’s (2004) comparison of RRS in children speaking English (an opaque orthography) with Dutch (a language with a transparent orthography).

In many cross-linguistic studies, English was compared with languages with more regular orthography, such as Turkish and Finnish, because they differ markedly in their phonological as well as orthographic structure. In one of the earliest studies, Öney and Goldman (1984) compared pseudo-word reading skills of 94 Turkish children (mean age= 5.9- 6.9) with 44 American children (mean age 6.3-7.2). The study was conducted twice; first, when the children were in first grade, and then again when they were in third grade. Participants were given letter tasks, decoding and PA tasks, including syllable and phoneme tapping, phoneme segmentation, initial and final phoneme deletion. The results of the study showed that the Turkish children were more accurate in decoding and also faster than the American children in the first grade with 94 percent accuracy. At the third grade, the Turkish children were still more fluent than American children in terms of reading ability (Öney and Goldman, 1984). Other studies comparing the differences between Turkish and English reading skills of children indicated similar results (Durgunoğlu and Öney, 1999).

In spite of PA's well-established role in reading, the contribution of PA to reading is found to be less important in transparent orthographies. For example, Landeral and Wimmer (2008) carried out a longitudinal study to investigate the development of word reading fluency and spelling in students acquiring the phonologically transparent German orthography. The students' (n=115) mean ages were 6.9 years at the beginning of Grade 1, 7.5 years at the end of Grade 1, 10.2 years in Grade 4, and 14.3 years in Grade 8 (*SD* = 0.5 years at all assessment points). Prediction measures (LK, phonological short-term memory, PA, RAN, and nonverbal IQ) were administered at the beginning of Grade 1; reading fluency was tested at the end of Grade 1 as well as in Grades 4 and 8. The findings showed that phonological measures contributed significantly to word reading fluency in Grade 1 only. Regarding RAN, it was a significant and consistent predictor of reading fluency in all grade levels. It has been concluded that the strongest specific predictor of reading fluency was RAN and not PA.

Moreover, in a cross linguistic study, Georgiou, Torppa, Manolitsis, Lyytinen, and Parrila (2012) examined the longitudinal predictors of nonword decoding and reading fluency in three languages that vary in orthographic depth: Finnish, Greek (both have consistent grapheme-to-phoneme mappings), and English. Eighty-two English-speaking, 70 Greek, and 88 Finnish children were followed from the age of 5.5 years old until Grade 2. Prior to any reading instruction, measures of PA, LK, and RAN were administered. In Grade 2, nonword decoding and text-reading fluency measures were administered. The results showed that the effect of PA on reading was limited to nonword decoding in English. These findings are in line with the findings of previous cross-linguistic studies (e.g., Georgiou, Parrila, and Papadopoulos, 2008; Mann and Wimmer, 2002) and provide support for the argument that reading development in transparent orthographies imposes fewer demands on PA than does reading development in opaque orthographies (e.g Mann and Wimmer, 2002).

Some studies have found PA's role be limited to the first year or two of schooling in transparent orthographies. For example, Furnes and Sameulsson (2011), in a longitudinal study tracking participants from kindergarten to Grade 2, explored the relationship between PA, LK and RAN and reading and spelling measures in cohorts of Scandinavian (n=280; mean age range =5.08-8.75) and American/Australian children (n= 1375; mean age range= 4.83- 8.33) learning to read in different alphabetic writing systems, namely Norwegian, Swedish (transparent) and English (less transparent). PA and RAN were measured in kindergarten and Grade 1. Measures of word recognition, phonological decoding, and spelling were taken starting in kindergarten through to Grade 2. The indings revealed that PA diminishes as a predictor of reading difficulties in transparent orthographies after the first years of schooling, but remained as a significant predictor in the English-speaking sample. Other studies of transparent orthographies (Dutch: van den Bos, 1998; Wesseling and Reitsma, 2000 (Finnish: Holopainen et al.; 2001 and German; Landerl and Wimmer, 2000) have reported that impairment in PA can explain reading difficulties in the first year in school but not in later schooling years. Moreover, it has been suggested that the consistency in transparent orthographies between grapheme and phoneme promotes the development of PA (Goswami et al., 2005; Seymour et al., 2003).

2.3.1.4 Relations between PA and Decoding in Arabic

As previously explained in Chapter One, Arabic is considered to be a transparent orthography. This means that it is relatively easy to derive the phonology from its orthography due to the near one-to-one association between letters and sounds (Mahfoudhi et al., 2011). The close phoneme-grapheme correspondence found in Arabic orthography can be expected to impact on the level of accuracy of single word decoding among Arabic speakers (Mahfoudhi et al., 2011).

Therefore, in relation to this study, Arabic is considered a language with transparent orthography, and the age range of the participants of this study is four to seven years. For this age range, all phonological information is provided and there is a consistent relation between letter and sound.

Evidence regarding the association between PA and reading in vowelized Arabic is reported in a number of studies, both concurrent and longitudinal. Almost all studies have found PA to be significantly associated with real word reading in the elementary grades.
Orthographic transparency is generally determined by the consistent and clear link between the corresponding sounds of written symbols. However, recent studies (for example, Asadi et al.'s (2017) cross-sectional study on 1305 Arabic-speaking children, mean age 6.8-11) observed that the vowelized version of Arabic did not behave as other transparent orthographies, such as Turkish and Finnish, to name a few. The features of Arabic orthography do not consistently enable correspondence between written language and spoken language. The use of diacritical marks for short vowels is visually complex and dense though providing full phonological information. Furthermore, the visual similarity between the diacritical marks makes the automatic perception as well as the orthographic pattern recognition a complex process. The results indicated that PA was significantly associated with vowelized reading script until Grade 6 (Asadi, et al, 2017). The findings are consistent with previous studies that showed the contribution of PA to reading in the Arabic context (see Al-Mannai and Everatt, 2005; Elbeheri and Everatt, 2007; Saiegh-Haddad and Geva, 2008; Taibah and Haynes, 2011). They are also consistent with other studies on alphabetic languages (See: Shatil and Share, 2003). The powerful effect of PA on vowelized Arabic was confirmed in a recent study. Tibi and Kirby's (2019) study on Grade 3 Arabic-speaking children (n=201, mean age= 6.5) examined whether the predictors of reading in English and other European orthographies function in the same way in Arabic. In this study, several measures were administered, including vocabulary, PA, RAN, orthographic processing, morphological awareness,

memory, and nonverbal ability. Furthermore, five reading outcome measures were used, in which words were presented in a vowelized script. All measures were administered within a two- to three-day period. The findings of the study showed that PA was the strongest predictor on almost all outcomes (Word Reading Accuracy, Pseudoword Reading Accuracy, Word Reading Fluency, and Text Reading Fluency), except Maze Comprehension. The largest final model effects were for word reading and pseudoword reading. Tibi and Kirby (2019) also underlined Caravolas et al.'s (2005) finding regarding the importance of PA in reading in many orthographies.

Mixed results on the relationship between PA and reading fluency were, however, found in Saiegh-Haddad's (2005) study of Arabic native speaking children (n= 42, mean age= 6.9 years old) which included phoneme discrimination and phoneme isolation and several other measures of reading-related processes. The results showed that phoneme isolation correlated with pseudoword reading fluency. However, no such correlation was found for phoneme discrimination.

In other studies, a longitudinal approach was taken, with PA measures administered at an early stage and reading assessed later. For example, Abu Ahmad, Ibrahim, and Share (2014) carried out a longitudinal study tracking native Arabic-speaking children from kindergarten (n=194; mean age: 5.9 years, SD: 3.6 months). In kindergarten, children were assessed on a variety of intra-lexical factors (phonemic awareness, phonological processing, visual-orthographic processing, pre-school print concepts and morphological awareness) and supra-lexical factors (general non-verbal ability, receptive vocabulary, syntactic awareness and working memory). At the beginning of Grade 2, children (n=177) were assessed again on word recognition and reading comprehension. The results revealed that the strongest predictors were phonemic awareness and phonological processing, followed by early print concepts, morphology and visual-orthographic processing. Furthermore, in a longitudinal study, Al-Sulaihim and Theo (2017) investigated the relationship between PA and literacy skills of

30 Kuwaiti children (mean age: 6.7) reading in MSA orthography. Data were collected twice during the school year with the aim of assessing the children's improvement regarding PA skills and reading abilities. The longitudinal results of Al-Sulaihim and Theo's study (2017) showed an improvement in PA skills once formal literacy training had been introduced. Similar to many studies reviewed here, these studies also indicated a reciprocal relationship between PA and reading ability.

Interestingly, however, studies report conflicting evidence as to whether PA is equally important to reading at all ages or school grades. In some studies, like Shatil and Share (2003) and Taibah and Haynes (2011), the researchers claim that the contribution of phonology to decoding in transparent orthographies is more crucial in lower grades and disappears in higher grades. Also, Al-Mannai and Everatt (2005) undertook a concurrent study in Bahrain to identify the best predictors of literacy skills in Arabic-speaking children in Grades 1-3, with mean age 6.8-8.4. Findings showed that non-word rhyming seemed to be more important in grade 1, whereas word rhyming was a reliable predictor of word reading in grades 2 and 3. Taibah and Haynes (2011) undertook a study in Saudi Arabia to examine children (n=237) from KG to third grade (mean age= 6.33 – 9.11). Findings revealed that PA predicted word recognition in all grades. However, they found that its predictive power varied across grades. Moreover, Taibah and Haynes investigated the ability of PA to predict reading using several measures including word reading, word decoding fluency text reading fluency, and comprehension fluency in Arabic. Findings revealed that PA predicts word recognition, accounting for 53% of the variance in kindergarten, 38% of variance in first grade, 32% in grade 2 and 46% of variance in 3rd grade. The predictiveness of PA for non-word reading fluency measures was (46%, 32%, 29% and 41% of the variances from kindergarten to 3rd grade respectively) started strong at kindergarten, but this ability decreased in Grades 1 and 2; however, it picked up again in Grade 3 as was true for word recognition. The results support the view that PA should best predict reading in the earliest grades, in transparent orthographies. However, Asadi et al. (2017) contradicted this claim, finding that PA contributed to decoding also in Grades 5 and 6. This finding might be explained by the transition period from reading vowelized to non-vowelized text (i.e. fourth grade). Children are forced to rely on other skills that help them to read non-vowelized text such as morphological knowledge (Saiegh-Haddad, 2013).Whilst the focus of the above studies was on vowelized orthography, others considered both the vowelized and non-vowelized forms.

In this study, non-vowelized script is not considered, since participants were still in the early stage of learning to read, and therefore exposed only to vowelized script. Nevertheless, it is of interest to comment briefly on other predictors that may be important for reading non-vowelized script (e.g. vocabulary, morphological awareness).

Asadi and Khateb (2017) undertook a study in Israel to examine Arabic speaking children (n=458) which included: 228 first graders (Mean age=6.99; SD= 9.4) and 230 second graders (mean age=7.96; SD=10.2). Asadi and Khateb investigated the role of vocabulary in the context of both non-vowelized and vowelized Arabic scripts. Findings suggested that the contribution of vocabulary increased with grades, and it was slightly higher in non-vowelized than in vowelized orthography. This was inconsistent with previous studies of Suggate et al. (2014) which suggested that children gradually rely less on vocabulary, if phonological information was presented for accurate reading. The sample size of the study was 133 and the participants were children and they recruited from two countries' first year of primary year i.e., Germany (age 7;2) and New Zealand (age 5;8) as well as from kindergarten in Germany aged 5;0 in order to provide age schooling matched samples. The increased contribution of vocabulary in the second grade in both vowelized and non-vowelized Arabic could be explained by diglossia of children. In the early stages of schooling children's oral language is poorly developed but after formal instruction in the literary language, their vocabulary would be richer. Asadi and Khateb also investigated the contribution of PA to word reading and found it was significant and similar for both vowelized and non-vowelized script, although this contribution declined with age. A similar trend was found by Asadi et al. (2017a) when they found morphology highly correlated with decoding and fluency. The contribution of morphology to decoding and fluency was found to be significant in the first and fourth grade, while in fifth grade it was significant to decoding only. The immediate exposure to morphological patterns might explain the contribution in the first grade. However, the morphological contribution in the fourth and fifth grade is consistent with previous results in Arabic (Abu- Rabia, 2007).

In contrast to these studies that affirm the contribution of morphology, Saiegh-Haddad and Geva (2008) argued that morphology does not contribute to decoding ability in Arabic as it does in English, beyond the PA contribution. According to their findings the reason behind that claim is the difference in the morphological transparency of the two languages. These findings supported the proposition that morphological contribution to reading is more significant in transparent morphologies (see Saiegh-Haddad and Geva, 2008).

The progression of reading acquisition in Arabic follows two stages learning to read a transparent orthography, then drawing on a variety of other tools and information to establish the word in use. A summary of studies investigating PA and other RRS can be found in **Table 2.1**, following:

AUTHOR/S	AIM	TASKS	AGE RANGE	KEY FINDINGS
Al-Mannai and Everatt (2005) *Bahrain*	To identify the best predictors of literacy skills in Arabic-speaking children	• Literacy measure (single-word reading and spelling) • Non-word reading (ability to decode letter strings) • PA • Short-term memory • Non-verbal abilities	Mean age: 6.8-8.4 Grades 1-3	• Measures of phonological skills (decoding and PA) predict reading and spelling ability
Saiegh-Haddad (2005) *North of Israel*	To investigate the correlates of reading fluency in Arabic particularly diglossic and orthographic factors	• Measures of basic reading processes including two cognitive processes (RAN and short-term working memory), and two phonological processes (phoneme discrimination and phoneme isolation), and one orthographic (letter recoding speed). • Phonological processing (phoneme isolation and discrimination for MSA/SAV) as independent measures.	Mean age= 6.9 Grade 1	• All predictor measures, except phoneme discrimination, correlated with pseudoword reading fluency. • Phonological processing (phoneme isolation and discrimination) for MSA phonemes was more challenging than that for SAV phonemes, phonological skills were not found to affect reading fluency directly. • The strongest predictor of reading fluency in vowelized Arabic was letter recoding speed. • Letter recoding speed

AUTHOR/S	AIM	TASKS	AGE RANGE	KEY FINDINGS
				was predicted by memory, RAN, and phoneme isolation.
Elbeheri and Everatt (2007) *Egypt*	To investigate the relationship between phonological processing and reading ability	• Rhyme awareness • Phoneme deletion • Retention/manipulation of digit name sequences • RAN • Decoding novel letter strings, distinguishing between similar words • Identifying words within a string of letters • Correctly spelling dictated text	Ages: 9-11 Grades 4-5	• Found relationship between literacy ability, decoding and phonological processing. • Showed that Arabic-speaking children with dyslexia had poor phonological processing skills.
Levin et al. (2008) *Palestine*	To investigate Arabic literacy acquisition as an intervention targeting Palestinian children with low SES	• PA • LK	Ages: 5-6 Kindergarten	• Early intervention with Palestinian children of low SES improved PA and LK. • All participants confused vocabulary in variants of Arabic, indicating issues specific to Arabic orthography.

AUTHOR/S	AIM	TASKS	AGE RANGE	KEY FINDINGS
Saiegh-Haddad and Geva (2008) *Canada*	To investigate the relationship between PA and morphological awareness (MA) in English/Arabic Bilingual children and their relationship with reading fluency	• PA • MA • Word recognition • Pseudoword decoding • Derive-word reading fluency • Oral language proficiency	Ages: 8-11 Grades 3- 6	• Correlation between PA in English and Arabic • No correlation with MA • PA predicted reading cross-linguistically • MA only predictive for English • PA was the only factor that predicted word decoding accuracy
Smythe, Everatt, Al Menaye, He, Capllini, Gyarmathy and Siegel (2008)	To investigate predictors of word-level literacy amongst in five diverse languages (Arabic, Chinese, English, Hungarian and Portuguese)	• PA (alliteration and rhyming) • Phonological decoding: non-word reading • Phonological access: rapid naming tasks • Phonological/working memory tasks • Sound discrimination • Word reading and spelling	Grade 3	Decoding and phonological-processing skills were good predictors of word reading and spelling among Arabic- and English-speaking children only.
Tibi (2010) *United Arab Emirates*	To determine the level of difficulty of four PA tasks in Modern Standard Arabic (as opposed to the dialectal variant)	• Initial sound identification • Rhyme oddity • Syllable deletion • Phoneme segmentation	Grades 1, 2 and 3	Identifying the initial sound in a word and rhyme oddity were much easier for children to perform than syllable deletion or phoneme segmentation.

AUTHOR/S	AIM	TASKS	AGE RANGE	KEY FINDINGS
Taibah and Haynes (2011) *Jeddah, Saudi Arabia*	To investigate the contribution of PA, RAN and phonological memory to basic decoding and reading fluency skills in Arabic	• Elision and blending • RAN (objects, colours, letters, numbers) • Non-word repetition • Digit span	Ages: 6-9 Kindergarten to Grade 3	• Significance of RAN increased with age • PA was more significant than RAN • Phonological memory had no relation to reading performance
Elbeheri. Everatt, Mahfoudhi, Abu ADiyar and Taibah (2011) *Kuwait*	To determine the influence of orthographic processing on comprehension within a group of Kuwaiti LD children	• Reading comprehension fluency • Orthographic discrimination • Phonological processing (decoding and awareness • Rapid naming • Memory	Grade 2-5	• Orthographic measure predicted variability in comprehension fluency. • It was significant in the older mainstream children (grades 4 and 5) when controlling for phonological processing. • The LD group showed little evidence of an influence of phonological processing but did of orthographic processing.

AUTHOR/S	AIM	TASKS	AGE RANGE	KEY FINDINGS
Elmonayer (2013) *Egypt*	To investigate the effect of dialogic reading (DR) on the promotion of Arabic phonological awareness skills of Egyptian kindergartners	• DR activities designed to improve PA skills, using an alphabet storybook • Syllable, rhyme and phoneme awareness	Ages: 5-6 Kindergarten	Children in the experimental group had higher PA levels in the post-tests than those in the control group.
Zayed et al. (2013) *Egypt*	To examine the relationship between PA, working memory and risk of dyslexia	• Rhyme awareness, • Phoneme isolation/blending	Ages: 5-6 Kindergarten	• Children at risk for dyslexia demonstrated deficits in PA in comparison to those not at risk • Strong relationship between PA and working memory • PA and working memory may have an effect on phonological sensitivity and reading skills (found a correlation between working memory deficit and PA skills)

AUTHOR/S	AIM	TASKS	AGE RANGE	KEY FINDINGS
Ammar and Ridha Ben Maad (2013) *Tunisia*	To investigate the effect of Arabic orthography on PA acquisition	• Phoneme counting, segmentation, and deletion	Ages: 5-8 Kindergarten – Grade 2	• Oral processing of Arabic is unlike English and acquisition of phoneme awareness is more determined by consonant position (which is more salient than vowel position in Arabic) • Manipulating syllables is easier than phoneme manipulation • Phoneme deletion is easier than phoneme segmentation.
Asaad and Eviatar (2014) *Israel*	To examine the influence of visual perceptual abilities, letter names and PA to text reading speed and accuracy	• RAN (letter naming task) • PA including (phoneme segmentation, blending phonemes, syllable deletion and sound deletion).	Ages: 7-10.9 Grades 1,3 and 5	• All levels of PA contribute significantly to reading accuracy, to the same degree. • Reading Speed RAN (letter naming) is crucial in first and fifth grade.

AUTHOR/S	AIM	TASKS	AGE RANGE	KEY FINDINGS
Abu Ahmad, Ibrahim and Share (2014) *Israel*	To compare the contribution of intra-lexical versus supra-lexical factors, and investigate differences in later word recognition and reading comprehension	• Visual orthographic processing • PA • Phonological memory (RAN-pseudo-word repetition) • Pre-school literacy • Word recognition and reading comprehension • Reading ability measures (word naming, pseudo-word naming) • Reading comprehension measures	Mean age: 5.9 K-2 grades	• Decoding in Arabic at the beginning of Grade 2 is relatively poor compared to English and Hebrew. • Word recognition skill was found to depend mainly on sub-lexical and lexical abilities in Grade 2. • The stronger predictors were phonemic awareness and phonological processing followed by early print concepts, morphology and visual-orthographic processing. • Reading comprehension skill was found to rely heavily on decoding skill but also on higher-order linguistic and cognitive abilities.

AUTHOR/S	AIM	TASKS	AGE RANGE	KEY FINDINGS
Layes, Lalonde and Rebai (2015) *Algeria*	To assess the effectiveness of PA-based intervention to improve reading of words and pseudo-words with 4th and 5th grade dyslexic and non-dyslexic children	• Phonological training of two metalinguistic RRS: PA and RAN • Training program focused on phoneme/syllable identification, phoneme matching/blending, and word segmentation.	Age: 10 Grades 4-5	• Dyslexic group performed significantly better in all post-training measurements, increasing reading, phonological processing, and metalinguistic-related skills, indicating a strong relationship between these variables. • The non-dyslexic group only improved in PA.
Makhoul (2017) *Northern Israel*	To examine the link between PA skills and the development of reading skills among 1st grade children with Arabic as L1	• Auditory word detection • Syllable/phoneme detection • Rhyme generation/detection • Initial/last sound isolation • Reading assessment measures Pseudo-word reading • Context-free oral word reading • Text reading	Grade 1	• Better performance in PA measures noted in higher-scoring group (HG) • After training, significant improvement was noted in PA in both HG and lower-scoring group (LG) • Moderate positive correlation was found between PA and reading performance in

AUTHOR/S	AIM	TASKS	AGE RANGE	KEY FINDINGS
				HG. Strong positive correlation was found in LG. • LG showed lower reading performance when compared to HG.
Al-Sulaihim and Theo (2017) *Kuwait*	To investigate the relationship between PA and literacy skills in Arabic-speaking Kuwaiti children learning MSA orthography	• PA tasks (syllable awareness, rhyme awareness, sound matching (initial/final position), phonemic isolation, elision) • LK task • Single word reading task	Grade 1	Longitudinal results indicated a general improvement in PA skills following introduction of formal literacy instruction
Asadi and Khateb (2017) *Israel*	To investigate the contribution of PA, RAN to reading (vowelized and un-vowelized).	• PA (phoneme deletion and phonemic segmentation) • RAN (letters and objects)	Mean age: 6.99-8 Grade 1-2	• PA contribution was similar for both grades in vowelized and un-vowelized orthographies. • RAN was weak and similar in both versions.
Asadi, Khate, Ibrahim and Taha (2017)	To examine the contribution of linguistic and cognitive variables	• PA (phoneme deletion and phonemic segmentation) • RAN (letters and digits)	Mean age: 6.8-11.7 Grade 1-6	• PA contributed to decoding • RAN contributed to fluency.

AUTHOR/S	AIM	TASKS	AGE RANGE	KEY FINDINGS
Israel	to reading processes			
Schiff and Saiegh-Haddad (2018) *Palestine*	To investigate the contribution of PA in spoken and standard Arabic word reading	• PA (phoneme and Syllabic structures)	Mean age: 7.7-15.6 Grade 2, 4,6,8 and 10	Differences between students reduced toward the end of elementary school in both spoken and standard Arabic language varieties.

Table 2.1: Studies on RRS in Arabic-speaking contexts

2.3.1.5 Causal relations between PA and decoding

Considering the literature on PA, Wagner and Torgesen (1987) proposed three different perspectives on the connection between PA and reading. The first related to the influence of PA on reading development, while the second focused on the consequences of learning to read. The issue here is whether PA causes or is a consequence of reading skills acquisition. The third perspective focuses on the bidirectional association between reading development and PA.

2.3.1.5.1 Effects of PA on Reading Acquisition

In a longitudinal study, Wagner et al. (1997) revealed that word-level reading performance is influenced by children's performance in PA awareness. Wagner et al. tested 216 children's PA and reading performance from kindergarten to Grade 4. PA was evaluated by means of phoneme deletion, oddity, blending, and segmentation, while reading performance was assessed using pseudo-words and single words in the study population. The study concluded that the relationship is causal, but it is pointed out that other variables such as home literacy environment, cognitive and SES background may also be primary factors that are responsible for performance in reading.

Lundberg et al. (1988) endorsed the training or experimental approach to addressing the causal relationship. Different studies have found that early readers, following extensive PA development training, attained significantly higher reading skills compared to control children who lacked training. For example, Ball and Blachman (1991) carried out a study of American kindergarten children (n=90) to evaluate a PA training program effect on reading. Results indicated that PA instruction contributed to improvement in children's early reading skills. Maclean, Bryant, and Bradley (1987) also found that PA awareness has a causal relationship with reading acquisition. They showed that the training of incipient PA skills increased the ability to acquire reading. Faber (1992) suggested that this training can be done from two to three years of age and their results indicated that the development of PA skills may require formal or informal instruction on printed words and knowledge of the alphabet. The researcher also suggested that levels of PA and reading development do not depend on the general level of language development. Blachman, Ball, Black, and Tangel (1994) found that the benefit for reading that children received as a result of specific training in PA development persisted six years later.

A pioneering investigation by Bradley and Bryant (1983) revealed a causal relationship between phonological processing and reading acquisition, because early intervention to support phonological processes in the kindergarten years facilitated reading acquisition. A follow-up from three years of age indicated that there is a sequence of progressive complexity that goes from the ability to learn rhymes, and culminates in the successful learning of reading. This relationship remained significant when the socioeconomic level of the children was controlled. These same researchers applied intensive phonological training with children aged five and six, and the children who followed it had better reading and writing performance than the control group.

Studies conducted on dyslexia are viewed as supplementary evidence for the relationship between phonological skills and success in reading. Information regarding children who or dyslexic or poor readers has reflected phonological function deficits leading to poor reading. For example, a study by Bradley and Bryant (1978) revealed poor PA sensitivity in dyslexic children. The results of the study are based on comparison with younger children who were matched in terms of reading level.

A number of studies have also highlighted the significance of letter knowledge (LK) for children's recognition of the association between word patterns and their pronunciation (Muter, Hulme, Snowling, and Taylor, 1998; Bradley and Bryant, 1983). These authors further argued that LK and PA should be aligned with each other for the attainment of best possible progress in the area of learning to read.

2.3.1.5.2 Effect of Reading on PA Development

This section considers studies of the effects of reading skills on PA in adult participants. Firstly, a study by Read et al. (1986) compared two groups of Chinese readers, non-alphabetic vs. alphabetic. The non-alphabetic group were adults with literacy skills only in Chinese characters, and the other, adults with literacy skills in Pinyin. All participants in this study were assessed on phoneme segmentation (add or delete a single consonant at the beginning of a spoken syllable). The results revealed that participants with the ability to read logographic script (non-alphabetic) could not delete or add individual consonants in the spoken Chinese words. Conversely, participants with the alphabetic reading ability performed the task of phoneme segmentation accurately. Based on this, Read et al. (1986) concluded that the development of segmentation skills during the process of learning to read, alongside the ability to write alphabetically, is independent of cognitive maturation development or the inclusion of non-alphabetic literacy.

Other researchers have conducted studies to investigate whether segmentation ability can develop over time without literacy skills, that is, whether it may arise from cognitive development and experience with spoken language alone. For example, an earlier study by Morais et al. (1979) investigated Portuguese adults with poor literacy skills and others with good literacy skills. All subjects (n=60) were examined in two tasks: deleting sound from words and non-words and adding sounds to words and non-words. Morais et al. (1979) found that phonemic skills were weaker in the poor literacy group (20% correct responses) than in good literacy group (80% correct responses). Morais et al. concluded that awareness of speech as a sequence of phones is not attained spontaneously in the course of cognitive growth but demands specific training for reading in the alphabetic system. Thus, learning to read, be it an adult or childhood, allows an individual to develop the ability to identify specific speech units. This does not exclude a possible role of cognitive growth in phonetic awareness development, and specific training is not effective before several critical stages of development.

<u>2.3.1.5.3 Reciprocal Relations between PA and Reading</u>

The above research has shown PA to be a predictor but also a product of learning to read. Findings from other research studies have indicated that the association between the ability to read and PA is reciprocal (Ehri and Wilce, 1980; Stanovich, 1986; Burgess and Lonigan, 1998; Goswami and Bryant, 1990; Wagner, Torgesen, and Rashotte, 1994). Perfetti et al. (1987) in a longitudinal study including 82 children from first grade tested the children at four different points in one year focusing on their reading ability and PA tasks such as tapping, blending, and deletion. The findings showed that the deletion task tapped phonemic knowledge that had a reciprocal relation to reading. In contrast, blending tasks tapped knowledge of phonemes in a non-reciprocal association with reading gains.

Adebate concerning the methodology of investigating the causal role of phonological awareness in reading acquisition is the failure to consider that some phonological skills emerge before reading and may have a causal influence on reading development (Hulme, Snowling, Caravolas, and Carroll, 2005; Castles and Coltheart, 2004). Goswami and Bryant (1990) claimed that there are different levels of analysis for PA (see section). The researcher suggestion were that syllables and onset/rime awareness develops prior to reading and related to reading acquisition, while phonemes develops later and is possibly a consequence of learning to read. Findings was supported by Bradley and Bryant, (1978, 1983), their results showed that rime awareness ability develop early (4 or 5 years), prior to reading instruction, and is highly predict later reading. While the ability to segment words by phonemes, appears late (after 5 or 6 years), possibly as a consequence of learning to read (Liberman et al., 1974).

Numerous studies have shown that PA can predict future reading. Despite extensive research in this area, there remains much debate about the nature of the relationship between phonetic awareness and reading. Some studies support the hypothesis that phonological sensitivity has different influences in reading acquisition. For example, in a longitudinal study carried out by Cronin and Carver (1998), with subjects (n= 95) from two grade levels (primary and first grade), children were assessed on two phonological tasks (initial consonant discrimination task and a rhyme matching test). Findings revealed that the phonological tests predicted unique variance in reading attainment as measured at the end of the second year of the study.

In a longitudinal study, Hulme et al. 2002 assessed 5-6 years old on three different task (deletion, oddity, and detection) tapping awareness of four phonological units (initial phoneme, final phoneme, onset, and rime). The findings indicated that rime and onset skill did not make an additional predictive contribution after considering phoneme skills. The results of such studies have led to debates about the role of various phonological units for predicting reading skills in children.

Some researchers suggest that the two-way relationship between PA and reading appears relatively early in the development of literacy skills, before formal reading instruction begins. Burgess and Lonigan (1998) carried out a longitudinal study of school-aged children for a year (n = 115, age = 4-5 years). All children were checked for PA tasks (tasks performed for rhyme oddity detection, blending, alliteration oddity detection, and deletion), and children's LK was used as the key to their reading ability. The research showed that the LK of children at phase one (T1) predicted PA at phase two (T2) and vice versa. Accordingly, LK may better predict PA skills, particularly smaller units in the form of phonemes.

2.3.2 Letter Knowledge and reading Skills

LK refers to knowing the names of letters and the sounds they represent. The ability to identify and name letters is considered to be an important predictor for the development of reading, both in transparent and opaque languages (Puranik and Apel, 2010). Several authors (e.g. Stuart and Coltheart, 1988; Johnston, Anderson, and Holligan, 1996) have suggested that LK provides an important first step in acquiring phonemic awareness. They argue that the ability to segment words into syllabic and sub-syllabic units (implicit phonological awareness) develops into phonemic awareness when children learn the sounds associated with letters. LK may thus provide the medium by which pre-reading analytical skills develop into the reading-based analytical skills evident in phonemic awareness.

Over several decades, studies have consistently found LK to be a powerful predictor of reading ability in pre-school children. For the most part, these findings are based on considerable positive correlations over time between children’s letter-naming ability and their reading proficiency in Grade 1. There is evidence that the level of LK that a child has on school entry can be a powerful predictor of their performance in reading skills by the end of the first grade (Walsh, Price and Gillingham, 1988). Share et al. (1984) tested the efficacy of some 39 different variables as predictors of reading achievement among kindergartners on entry, including IQ, level of vocabulary, and socio-economic status. LK emerged as one of the best individual predictors and was also found to be second best predictor of children’s reading achievement in Grade 1. The top-ranking predictor at that specific stage was phoneme segmentation.

It should be noted, however, this area of research is inconclusive about the actual nature of the relationship between LK and reading. In the early 1970s researchers maintained that training children to name letters did not initially improve their word-reading skills, due to the decontextualized nature of the training they were given (Jenkins et al., 1972; Silberberg et al., 1972). It was later argued that the relationship between LK and reading skills is causal and that children are able to draw on letter-name knowledge to read words that are unfamiliar to them (Treiman and Rodriguez, 1999). More recently, Burgess et al. (2002) claimed that LK predicts reading ability because it serves as an indicator of the richness of the home literacy environment, a factor known to influence development of early reading skills. In their previous study on this subject, Burgess et al. (1998) argued that young children are more likely to know the names of letters rather than their sounds since parents encourage them to sing alphabet songs, based on letter names.

Research findings point to a distinction between LK i.e. letter-*naming* (LN) and letter-*sound* (LS) knowledge and suggest that each has a specific relationship with reading achievement throughout literacy development (Caravolas et al., 2001; McBride-Chang, 1999; Wagner, Torgesen, and Rashotte, 1994). Duncan and Seymour (2000) argued that combining scores for LN knowledge and LS knowledge into a single measure can provide fuller information about children's LK since it has been found that most children are only able to identify a letter on the basis of either its sound or its name. Other researchers have found evidence that good prior LN knowledge positively influences children's knowledge of LS (Burgess et al., 1998; McGuiness, 2004). The findings of Stahl and Murray (1994) indicate that letter-sound knowledge appears to be necessary for the child's ability to analyse spoken words into onsets and rimes. They found that the majority of the participants in their sample of kindergartners and first graders who were above chance on an onset-rime task also had good LS; only one child who knew fewer than 45 letters managed to pass chance level. McBride-Chang (1999) examined the development of LS and letter-naming ability (LN) and their association with other RRS, tracking their development in children (N=91) from starting kindergarten to halfway through Grade 1. Measurement of LS included phoneme awareness using segmentation and identification tasks. Participants' previous LN was found to predict subsequent levels of both LN and LS. However, LS only predicted subsequent LS. Only modest correlations between LN and LS were found. Both proved to be substantially correlated with alphabet ordering throughout the study, leading McBride-Chang (1999) to suggest that letters found toward the end of the alphabet should feature more frequently in tests because children knew the opening letters better. Both LN and LS contributed unique variance in predicting subsequent RRS. Participants' LS was dependent on the grapheme-phoneme relation providing support for Treiman et al.'s (1994) letter-name hypothesis. Results suggest that alphabet knowledge consists of both LN and LS and that these follow different developmental patterns.

There is also evidence that some letters of the alphabet may be easier to learn than others. Huang et al. (2014) explored factors that may affect LS knowledge with kindergartners from disadvantaged backgrounds (N= 1197, 5;0 years old). Their prediction was that numerous factors, including letter-name structure, letter-sound ambiguity, letter–letter-name knowledge, own-name advantage and PA, would affect the child's knowledge of a particular LS.

Table 2.2 shows factors affecting LS learning and the studies in which these effects were found:

Factor	Studies finding an effect on LS learning
Child's exposure to the letter: children are most likely to know the name of the first letter of their own name.	Treiman and Kessler, 2003
Letter's position in the alphabet: letter sounds occurring at the beginning of the alphabet are more likely to be known.	McBride-Chang, 1999
Frequency of letter occurrence: letter sounds that occur more frequently are more likely to be known.	Treiman, Kessler and Pollo, 2006
Letter name structure: letter sounds are more likely to be known if the name of the letter contains its sound especially if it is the first sound e.g. /d/ - D.	Kim, Petscher, Foorman and Zhou, 2010 McBride-Chang, 1999
The level of letter-name knowledge: This may help children to learn letter sounds as they consider it to be a stable referent for an abstract concept.	McBride-Chang, 1999 Share, 2004
The interaction between PA and the letter-name structure: This may contribute to letter-sound knowledge by making the letter sound more salient and easier to remember and associate with the letter form.	Foy and Mann, 2006
Letter ambiguity: When letters share sounds with other letters e.g. C and K for /k/ this causes more difficulty, especially in a non-transparent orthography.	Scanlon, Anderson and Sweeney, 2010

Table 2.2: Factors affecting LS Knowledge in Young Children

Furthermore, scholars such as Kim, Petscher Foorman, and Zhou (2010) argued that both letter knowledge and PA are essential elements of developing reading abilities in alphabetic orthographies. Likewise, letter-name knowledge has an immediate association with word reading. Thus, knowledge of letter-names gives children an impetus to connect letters and print. In this context, Kim et al. (2010) conducted a study where the primary objective was to look at letter sound knowledge as an element of letter-name knowledge, PA, and letter characteristics. They analysed whether young children would detect letter sounds as a component of realizing letter names and PA. In this examination, they explored the ability of 653 English-talking kindergarten children toward the start of the school year. Participants were evaluated on PA abilities by utilizing phoneme blending and onset-rime tasks. Letter knowledge was evaluated by utilizing letter-name and letter-sound practices. Results showed that letter-name learning has an undeniable effect on letter-sound learning. Likewise, PA was found to have an effect on letter-sound knowledge if the children knew the names of the letters. Treiman and Kessler (2003) argued that LN provide cues about LS (phonemes) which are essential for decoding. Further, LN knowledge has a direct relationship with word reading as letter-names provide a link between letters and print which supports children in learning to read.

Moreover, the role of LK in PA and reading acquisition was evidenced in several longitudinal studies. For example, Blaiklock (2004) carried out a longitudinal study on English-speaking children (n=36; mean age= 5;1 years). Blaiklock examined the relationship between reading and PA while controlling for extraneous factors, such as LK and pre-existing reading skills. All participants were tested over a period of two years, on several measures (vocabulary, LS knowledge, LN knowledge, phoneme awareness, rhyme awareness, and word reading). Findings showed that children were able to complete rhyme awareness tasks before they started to read and were unable to complete a phoneme deletion task until after they had developed word reading skills. Also, controlling for LK reduced most correlations between PA scores and later reading to non-significant levels.

It is worth noting that, to date, most of the studies finding a predictive relationship between LK and reading have been carried out in English-speaking countries with school children who have English as their mother tongue. It is, therefore, arguable that this predictive relationship may have been influenced by specific features of literacy acquisition related to English. Alternatively, it may be linked to a particular focus on letter-naming skills in the curriculum. This increases the interest in investigating whether this predictive linkage is found in other alphabetic languages. In this regard, researchers have reviewed some evidence of LK in Arabic. In a recent longitudinal study by Alsulaihim and Theo (2017) (n=30, mean age=6.7), the outcome of the PA assessment which was carried out twice in one year on Arabic-speaking children revealed that there was a significant association between LK and reading at T2 only. It is arguable that this outcome was affected by experience after children received literacy instruction.

The diglossic nature of the Arabic language (see **section 1.4.1**) was explored in an intervention study by Levin et al. (2008). Within the framework of the intervention study, implemented by teachers, Arabic literacy acquisition was studied among Israeli Palestinian low socioeconomic status kindergartners. Findings of this study indicated that, in the preliminary stages, letter naming, alphabetic awareness, and phonological awareness were very low. Whereas the comparison group recorded little or no progress for the whole year, the intervention group progressed noticeably on all three skills. Following the intervention, children reflected the diglossic nature of the Arabic letter name system in their preference for standard over colloquial names. There was a slight tendency to confuse adjacent letters that are not necessarily similar.

The results of the reviewed studies in this section indicate that both alphabet knowledge, as well as PA, are critical for children's reading acquisition with alphabetic orthographies (e.g. Blaiklock, 2004; Kim, Petscher Foorman, and

Zhou, 2010). In addition, phoneme awareness seems to be linked more to LK than the other two phonological levels (as shown by the results of the reviewed studies of the three levels of PA). These results were a motivation to include LK in the Arabic test battery of the current study.

2.3.3 Rapid automatized naming (RAN) and Reading Skills

Rapid Automatized Naming (RAN) or Rapid Naming (RN) is an essential cognitive skill which refers to the time required for a child to rapidly and accurately name an array of well-known visual stimuli (letters, digits, objects, or colours). The speed at which an individual can name/recognise items is believed to provide insights into his/her ability to retrieve associations between visual or graphological symbols and phonological codes (Taub and Szente, 2012). This ability is pivotal to reading as it involves a complex cognitive process. Inadequacies and deficits in both PA and RAN are some of the main reasons for reading disability in children. Saccades are used in RAN in the understanding and prediction of current and future reading difficulties in children. Saccades are rapid eye movements that can be generated on command (Termsarasab, Thammongkolchai, Rucker, and Frucht, (2015). Usually, in reading, children should be able to generate saccades rapidly as the eye moves. This means that children can disengage and rapidly move their eyes from one stimulus to another. Therefore, these tasks require a rapid, accurate shooting and fixation of the eyes on an object, as well as subsequent naming of the objects (Felser, Phillips and Wagers, 2017). Signs suggestive of a neurological disorder in children that could affect their reading ability include delay, slow movement, dysconjugate movement of the two eyes, and inaccuracy in refixation, either undershooting or overshooting of the eye on an object. Such eye movement disorders have been reported (Hutzler & Wimmer, 2004).

RAN has been connected to reading development. For example, a study was carried out by Georgiou, Papadopoulos, Fella, and Parrila (2012) on Greek students (n=68; mean age = 7.9 years, SD = 3.36) from Grade 2 to Grade 6.

Students were assessed three times on RAN (Digits and Objects), PA, orthographic processing, speed of processing, and reading fluency. RAN was strongly related to reading fluency and accounted for unique variance over and above the contribution of PA.

Evidence shows that the RAN-reading relationship is affected by the consistency of orthography, accounting for more variance in transparent orthographies than in opaque orthographies. Georgiou, Parrila, and Liao (2008a) carried out a study in which English-speaking children (n=40), Greek-speaking children (n=40), and Chinese-speaking children (n=40) were examined on RAN, reading accuracy, and reading fluency tasks in grade 4. The results revealed diffrences in the strength of the RAN-reading relationship across languages, suggested that different RAN components (articulation and pause time) might be responsible for the RAN-reading relationship in different languages. Importantly, the coefficients of determination between RAN and reading fluency in Greek were at least four times as large as the coefficients of determination between RAN and reading accuracy in English. This finding provides an explanation as to why some researchers have argued that RAN is more strongly related to reading in consistent orthographies, such as German, Italian, or Greek, in comparison to English (e.g., Di Filippo et al., 2005; Landerl and Wimmer, 2000; Mayringer et al., 1998; Nikolopoulos et al., 2006; Wimmer et al., 1999).

Landeral and Wimmer (2008), as discussed above, carried out longitudinal study to investigate the development of word reading fluency and spelling in students acquiring the phonologically transparent German orthography. The students' (n=115) mean ages were 6.9 years at the beginning of Grade 1, 7.5 years at the end of grade 1, 10.2 years in grade 4, and 14.3 years in grade 8 (SD=0.5 years, at all assessment points). Prediction measures (LK, short-term phonological memory, PA, RAN, and nonverbal IQ) were assessed at the beginning of grade 1; reading fluency was tested at the end of grade 1 as well as in grades 4 and 8. The finding showed that RAN was a significant and consistent predictor of reading fluency in all grade levels. It was concluded that RAN rather than PA, was the strongest specific predictor of reading fluency.

Moll, Fussenegger, Willburger, and Landerl (2009) conducted a study on German children from different sample (n = 1248; mean age = 10.52, 9.38, 9.35) to investigate the correlations between RAN, PA, phonological decoding (nonword reading fluency), and orthographic processing (word reading fluency and spelling). Results of a series of hierarchical regression analyses indicated that RAN explained more variance in word and non-word reading fluency than PA.

Ziegler et al. (2010) investigated the role of PA, memory, vocabulary, RAN, and nonverbal intelligence in reading performance across five languages lying at differing positions along a transparency continuum (Finnish, Hungarian, Dutch, Portuguese, and French). Results from a sample of 1,265 children showed that whereas the impact of PA was more important in inconsistent scripts, RAN appeared not to be significantly modulated by script transparency. This finding stands in contrast to previous findings that RAN was more important than PA in predicting reading in transparent orthographies, such as Dutch and German. It was explained that this might be due to assessing object RAN rather than alphanumerical RAN, since object RAN tends to have lower correlations with reading performance than alphanumerical RAN does (Vaessen et al., 2009). Second, it is probably misleading to think of RAN as an independent non-phonological component (for a review, see Vaessen et al., 2009). Also, many studies have shown that RAN is essential in fluent reading. Examples of these studies have been discussed previously (e.g. Georgiou, Parrila, and Liao, 2008; Landeral and Wimmer, 2008).

Torgesen et al. (1997) carried out a study on English speaking children (n=285) to test the hypothesis that individual differences in RAN made a unique contribution to explaining the growth of orthographic reading skills in 2 overlapping periods of development: second to fourth grade, and third to fifth grade. Findings showed that the contribution of RAN to later reading skill appears to diminish between kindergarten and Grade 3.

Some studies have also found that PA and RAN can be intertwined and related to each other. For example, Taub and Szente (2012) conducted a study to investigate the relationship between PA and RAN and reading fluency. Their sample consisted of 86 English-speaking participants attending Grades 1-4 in an inner-city school located in an urban environment with high rates of poverty. They found that both PA and RAN were valid and reliable predictors of participants' reading fluency but RAN accounted for more variance in predicting reading fluency than PA alone. They also found that RAN had a large direct effect on PA, which was unexpected, suggesting that RAN shares a large portion of variance with PA. Their results provide support for the use of assessments of both RAN and PA to identify students at-risk of developing reading difficulties as together they successfully predicted reading fluency.

Furthermore, RAN has been shown to predict reading above and beyond PA. With regards to age, Kirby et al. (2003) found that the relationship between RAN and reading ability was stronger in older age groups, whilst the study by Swanson et al. (2003) suggested the opposite, as they argued that the significance of RAN actually decreases with age. Kirby et al. (2003) found that PA and RAN measured in kindergartners were able to predict reading development to Grade Five, but their predictive abilities varied. PA was most strongly correlated with reading development in Grades One and Two. In the case of RAN, although this initially displayed a weaker predictive relationship with reading development, this increased year on year. The study also found that participants performing poorly in both PA and RAN tasks were most likely to develop reading difficulties by Grade Five, followed by those performing poorly in the RAN tasks alone. This may suggest that RAN plays a role in orthographic processing (Wolf and Bowers, 1999a) which in turn facilitates reading development (Ehri, 1997). This explains why kindergartners' RAN scores had greater predictive power for reading performance in the higher grades (Kirby et al., 2003). There is an agreement that RAN is the most important predictor of reading ability in children in all orthographies: in opaque orthographies, for example, Parrila et al. (2004) examined English-speaking children (n=161; mean age= 5.56 years), to find how measures of articulation rate, verbal short-term memory, RAN, and PA tasks administered in kindergarten and again in Grade1 jointly and uniquely predict word reading and passage comprehension variance at Grades 1, 2, and 3. Results revealed that in Grade 3 RAN still accounted for a significant unique variance.

Comparing transparent and opaque orthographies, Mann and Wimmer (2002) examined the predictors of reading in English and German languages. Children aged between 5-8 years were given two tests of phonological awareness (Phoneme Identity Judgment and Phoneme Elision), RAN-Colours, letter identification, and short tests of word and nonword reading accuracy and speed. They found that phonological awareness was the only significant predictor of both reading accuracy and speed in English, but in German orthography, RAN was the only significant predictor of reading speed.

2.3.3.1 RAN in Arabic

Arabic is a semitic language with orthography normally representing long vowels and consonants. It is also characterised by allography or linear and non-linear morphology, and diglossia (Daniel, 1992). As noted previously, vowelized Arabic orthography, in which written words are presented with short vowels indicated, is considered transparent, wherease non-vowelized script, in which only consonants and long vowels are shown, is considered a deep orthography (Asadi and Khateb, 2017). Researchers are of the opinion that there is a relationship between the predictors of reading including RAN and all the characteristics of Arabic orthography (Saiegh-Haddad, 2005). (See **Table 2.1** in **section 2.3.1** for a summary of previous studies on Arabic).

Many studies showed that RAN was a significant predictor of real-word reading. For example, Layes et al. (2017) carried out a study in three groups of Arabic speaking children: a group with dyslexia (n=20) from grade 6 compared to a group of normal readers matched in age (n=20) and a younger reader group from grade 4 (n=18). All children were assessed on word reading, reading comprehension, morphology awareness, and RAN in addition to a nonverbal mental ability test. Layes et al. found that RAN was an important predictor of word reading and reading comprehension.

Tibi and Kirby (2018) investigated PA and RAN as predictors of reading in Arabic. They carried out a study on third-grade Arabic-speaking children (n=201; Mean age=8.08 years, SD = 5.4). All children were examined on general cognitive ability, vocabulary, PA, RAN, word reading, and reading comprehension. Findings showed that RAN made unique contributions to each outcome, with RAN having its largest contributions to the two fluency measures (word and text reading fluency).

Similarly, Tibi and Kirby's (2019) study on Grade 3 Arabic-speaking children (n=201, mean age= 6.5) examined whether the predictors of reading in English and other European orthographies function in the same way in Arabic. In this study, several measures were applied to vocabulary, PA, RAN, orthographic processing, morphological awareness, memory, and nonverbal ability. Furthermore, five reading outcome measures were used, in which words were presented in a vowelized script. The findings of the study showed that RAN played a significant role, having significant effects on all outcomes in the final models. Its effects were particularly strong for Word Reading Fluency and Text Reading Fluency, which concurs with prior evidence on the role of RAN in fluency measures (Tibi and Kirby, 2018; Georgiou, Parrila, and Papadopoulos, 2008; Taibah and Haynes, 2011).

Taibah and Haynes (2011) provided evidence that RAN's effect was stronger with older children, based on a study undertaken in Saudi Arabia to examine children (n=237; mean age= 6;33 -9;11, KG–third grade). Taibah and Haynes investigated the ability of PA, RAN and phonological memory to predict reading using several measures including word reading, word decoding fluency, text reading fluency, and comprehension fluency in Arabic. Findings showed that RAN's predictive power was lowest in the first grade, it gradually increased and was strongest at Grade 3, when basic decoding skills are more automatic.

In contrast, another Arabic study showed that RAN has a stronger effect for first and fifth graders than third graders. The study was carried out by Asaad and Eviatar (2014) Arabic speaking first graders (n=31; mean age 7.02 years), third graders (n= 30; mean age 8.94) and fifth graders (n=35; mean age=10.88 years). Findings revealed that RAN had a stronger effect for first and fifth graders than third graders, Asaad and Eviater argued that the first-grade effect was related to letter learning whereas the fifth-grade effect was related to orthographic pattern learning.

2.3.3.2 Double deficit hypothesis

Dyslexia in children may be caused due to phonological impairments, rapid naming deficits or both. According to the double deficit hypothesis (DDH), proposed by Wolf and Bower in the year 1999, children who have both of these deficits have the most severe reading difficulties, while readers with phonological impairment have moderate reading difficulty and reading impairment is least in those with naming deficit only (Wolf and Bowers, 1999b). The existing evidence about DDH is, however, unclear and it has not been consistently documented (e.g., Kirby et al., 2003; Manis et al., 2000; Schatchneider et al., 2002). This may be attributed to three reasons: Firstly, there is a huge methodological variation in the previous work related to the participants' characteristics. Some of the studies included typically developing children and examined if they could be assigned to the three hypothesised deficit groups (e.g., Manis et al., 2000; Powell, Stainthorp, Stuart, Garwood, and Quinlan, 2007; Sunseth and Bowers, 2002), whereas other studies involved children with reading disabilities (e.g., Lovett et al., 2000). Manis and his associates included second-grade children (n=85) and grouped them to children with no deficits (n = 50), children with phonological deficits only (n = 13), those with naming-speed deficits only (n = 14), and those with a double deficit (n = 8). The double-deficit group had word identification scores ranging from 25th percentile to the 48th percentile, and thus not all the children in this group had reading problems (Manis et al., 2000). Similarly, Lovett et al. (2000) included 166 children of age between 7 and 13 years who were diagnosed with reading disabilities. They found that the double-deficit group showed the most severe deficits on measures of reading, spelling, followed by the phonological deficit group and the naming deficit group. Secondly, the inconsistency may be because of the use of different cut-off scores to identify reading difficulties in children. For example, a study which included typically developing children considered the 30th percentile as the cut-off score (Sunseth and Bowers 2002) while it was the 25th percentile in Manis et al.'s (2000) study. On the other hand, studies involving children with reading disabilities also used the 25th percentile as a cut-off score (Lovett et al. 2000). Therefore, the lack of a consistent cut-off score may have resulted in children being identified as having a double deficit in the absence of noticeable reading problems. Some other concerns were raised by Schatschneider and his colleagues (2002) about the use of arbitrary cut-off values, especially regarding children with double-deficit. They argued that greater severity of reading impairment in children with double deficit might be partly due to a statistical artefact while grouping children based on their deficits in the cognitive skills (Schatschneider et al., 2002). In contrast, Kirby et al. (2003) found evidence to support the double-deficit hypothesis. They conducted a longitudinal study to compare word identification, word attack, and passage comprehension in children with and without reading difficulties. They included children from kindergarten to Grade 5 (ages between 4-10 years) and divided them into three groups with reading difficulties, and a fourth no-deficit group. They showed that the children with no deficits performed consistently well, while the participants in the double-deficit group performed consistently poorly. Participants with single phonological deficits performed poorly at the beginning but performed similarly to no deficit children over time, and the naming-speed-deficit group showed poorer performance throughout, almost similar to the double-deficit group. Their study results showed that children with a double-deficit lagged behind the no deficit group by almost two years of achievement and lacked signs of improvement or catching up.

The validity of the double deficit theory has been tested not only in English as an opaque orthography, but also in transparent orthographies such as Spanish, Italian, and Dutch (Wimmer, Landerl, and Frith, 1999; Zoccolotti et al., 2005). In these orthographies, the most common reading difficulty for children was reading speed rather than decoding (Wimmer, Landerl, and Frith, 1999; Zoccolotti et al., 2005). A study by Wimmer et al. (2000) found that children with dyslexia showed accuracy for text and word reading, and nonword reading, but there was a difference in reading rate between children with and without reading deficits. They also observed that the phonological deficit group had a reading rate deficit only for text, but not for nonword reading. Unlike phonological deficit and no-deficit groups, both the naming deficit and the double-deficit groups differed significantly and exhibited reading rate impairments for text, words, and nonwords. Conflicting findings were observed from the previous literature concerning the double-deficit hypothesis related to consistent orthographies. Wimmer and his colleagues (2000) reported that children in the double-deficit group did not have any problems related to nonword reading accuracy. In contrast, a study done by Escribano (2007), in Spanish children with dyslexia, found that the double-deficit group and the phonological deficit group had low pseudoword reading accuracy compared to the no-deficit group and phonological deficit and the double-deficit groups had almost similar results. However, Escribano's results should be interpreted with some caution due to small group sizes and the power of detecting significant differences.

The most important implication of the double-deficit hypothesis is that RAN measures should be incorporated into kindergarten and first-grade level screening batteries, to aid in the early identification of difficulties. Researchers argue that phonological awareness, rapid naming and letter knowledge are the best predictors of reading acquisition (e.g.Bishop 2003; Catts1996; League and

Bishop 2004) and recommend that all should be part of any good assessment and remediation battery.

2.3.4 Section Summary

As shown in **Table 2.5** (Mapping of Research Aims, Objectives, Questions, Hypotheses and Statistical Tests Used), the primary aim of the study is to develop a test battery to assess RRS of Saudi children between ages 4 and 7. **Sections 2.2 and 2.3** served to provide evidence from the existing literature to support the choice of of PA, LK and RAN as a predictors of decoding, to be included in the RRS test battery.The review showed that studies in English and other language contexts, including Arabic, provided extensive evidence for the choice of these predictors of reading ability of pre-school and early school-age children. The reviewed emprical studies also confirmed that these predictors are significant for assessing reading skills, particularly in languages with transparent orthographies like Arabic. As the study aims to develop an RRS testing battery in Arabic, the evidence that was provided throughout the literature review above is particularly supportive regarding the claims of the present study. In the following section, the issue of how RRS can be teseed is considered.

2.4 Methods of Testing RRS

2.4.1 PA Skills Testing

Due to the complexity of PA, tasks aiming to evaluate this ability must include a range of different activities, designed to assess the different aspects of PA (Wium et al., 2011). The implication of findings on PA and literacy development for this study is that, to be developmentally sensitive, the PA assessment must take into consideration both dimensions of the PA construct – the size of the linguistic unit and the explicitness of the task (Stackhouse and Wells 1997) (see **section 2.3.1.1**). When designing PA tasks, it is therefore important to include varying levels of complexity of the linguistic units used (syllable, onset-rime, and phoneme) and to consider the cognitive difficulty relating to task demands (e.g. ranging from identification to manipulation). A wide variety of tasks has been used in the literature to measure various aspects and sub-sets of PA skills involving different degrees of cognitive difficulty and linguistic complexity. **Table 2.3** provides a brief explanation of these tasks, in ascending order of complexity. Thus, for example, similarity judgment which involves recognition of segments, is easier than isolation, which requires manipulation of a segment.

TASKS	EXPLANATION
Similarity Judgment	Requires child to judge if two stimuli are similar in some respect e.g. if two stimuli rhyme, giving a ‘yes’/'no' response (Treiman and Zukowski, 1996; Deacon and Kirby, 2004). Example: “<car>, <kit> – do these words rhyme?" Expected response: no
Odd One Out	Requires child to identify an item that differs from other items due to a phonemic property (Torgesen and Bryant, 1994). Example: "Which word begins with a different sound: “<car>, <kit>, <cup>, <bus>?" Expected response: <bus>
Completion	Requires child to complete a word by providing the final sound/syllable (Muter et al., 1997a). Example: "Here is a picture of a bat. Listen I'll say the first sound of the word /ba/. Can you finish the word *ba*____?" Expected response: /t/
Segmentation	Requires child to divide sentences into words, words into syllables or words into phonemes or sounds. An oral response indicating the individual syllable or phoneme is required (Dodd et al., 1996). This depends on the precise instructions given by the assessor, which could include providing a number or making a particular sound, e.g. one tap per syllable. Example: “How many parts [syllables] are there in the

TASKS	EXPLANATION
	word <coffee>?" Expected response: Two Example: "How many sounds does the word <cat>have?" Expected response: Three
Isolation	Requires child to identify a phoneme in a particular position in a word (initial, final or medial). An oral response indicating the phoneme is required (Stahl and Murray, 1994). Example: "Tell me the first sound of the word <car>" Expected response: /k/
Substitution/ manipulation	Requires child to isolate a sound in a word and change it to another sound, forming a new word which must be named as a response (Rosner, 1999). Example:"Say the word <cat>. Now say it again, but instead of /k/ say /b/" Expected response:/bǣt/
Blending	Requires child to blend units of sound to form words/non-words. Child must pronounce the word after blending the sounds together (Wagner et al., 1999). Example: "I'll say some sounds /k/ /æ/ /t/. Which word do these sounds make?" Expected response: <cat>.

TASKS	EXPLANATION
Deletion	Requires child to say a word and then repeat it, deleting one element. This element may be a root word in a compound or a syllable in a word. In a phoneme deletion task, the child says a word and then repeats it, deleting one of its phonemes. All require a verbal response of the phonological form that remains after the deletion (Rosner, 1999). Example: “Can you say <spaceship> without <ship>?” Expected response: <space> “Can you say <car> without /k/?” Expected response: /ar/
Generation/ Production	Requires child to generate a word with particular sound properties, e.g. that rhymes with a given word (Muter et al., 1997a). Example: "Can you tell me a word that rhymes with <cat>"? Expected response: <hat>, <bat>, <rat>....

Table 22.3: Types of tasks used for testing PA

As noted above, research has shown that PA plays a central role in the development of reading ability, which points to the importance of assessing this in young children, in order to establish whether or not they are developing normally. The four most widely used standardized tests for evaluating PA development in English language contexts (Dodd et al., 2000; Wagner et al., 1999) are outlined in **Table 2.4**.The tasks are divided into five categories designed to reflect the different stages of PA development:

1. Word level tasks (including simple word lists);
2. Syllable level tasks (including syllable segmentation, completion,

identification and blending);

3. Onset-rime level tasks (including rhyme awareness and production);

4. Phoneme level tasks (including phoneme awareness, isolation, blending, segmentation, completion and deletion);

5. Literacy tasks (including non-word spelling and reading).

The four standardised tests cover a range of age groups (from 3 to 24 years) but are typically applied to assess PA in pre-schoolers and young school-age children, and dyslexia in older participants. Since these tests have been validated in a range of studies, the tasks used in the test battery were modelled on this reliable foundation.

TEST		PIPA (Dodd et al., 2000)	PAT (Muter et al., 1997a)	CTOPP (Wagner et al., 1999)	QUIL (Dodd et al., 1996)	PhAB Frederickson et al. (1997)
Population Ages		3;0-6;11	4;0-7;11	5;0-24;11	6;0-12;0	6;0-14;11
Normative Population		Australian and British	British	American	Australian	English/bilingual speakers
Areas of Measurement		PA	PA, Speech Rate, LK	PA, Phonological memory, RAN	PA, Non-word Spelling & Reading	PA, RAN, Decoding
PA TASKS	**Word Level**		Speech Rate			
	Syllable Level	Syllable Segmentation of unfamiliar word	Syllable Completion	Syllable Blending (5-6 years)	Syllable Identification/ Segmentation	
	Onset-Rime Level	Rhyme Awareness (odd one out)	Rhyme Awareness (detection/production)		Rhyme Awareness (recognition), Spoonerism	Rhyme Awareness Spoonerism
	Phoneme Level	Alliteration awareness Isolation/Segmentation LK	Phoneme Awareness (Completion and deletion), LK	Phoneme Blending/Elision (7-24 years)	Phoneme Deletion/ Segmentation/Detection	Alliteration
	Literacy				Non-word Spelling Non-word Reading	Non-Word Test

PIPA=Preschool Inventory of PA, CTOPP=Comprehensive Test of Phonological Processing PAT=Phonological Abilities Test, QUIL=Queensland University Inventory of Literacy PhAb= Phonological Assessment Battery

Table2.4: Standardized Tests of PA in English

2.4.2 LK skills testing

LK tasks require children to identify the phonemes that correspond to graphemes, usually by presenting these on flashcards in a randomized order. LK is usually measured using a test administered during kindergarten which involves naming upper case letters. As previously explained, since the Arabic alphabet does not have upper and lower case letter forms, LK is usually measured using tasks that involve recognition of letters in the isolate form.

2.4.3 RAN skills testing

A RAN test can be performed in a continuous or a discrete format, but there is debate on different aspects of RAN testing. Some scientists believe that the continuous format is a better predictor compared to the discrete format, where individual stimuli are presented; the latter is often believed to be time consuming (de Jong, 2011). RAN is a valid and reliable tool that makes a consistent contribution in predicting reading and children with dyslexia can be grouped into those with either RAN deficits or phonological deficits or both.

Task analysis of RAN involves various component processes including perception, attention, visual processing and discrimination, pattern identification, integration of characters with stored orthographic and phonological representations, inferential thinking, semantic information activation and integration (Wolf and Bowers, 1999).

In addition, Bowey and his colleagues (2005) believed that in the early stages of reading development, RAN is mediated by both language knowledge (LK), and phonological processing ability (Bowey et al., 2005). However, in later stages of development, the phonological processing ability is the most important predictor. Reading performance may be best assessed using PA, while time taken for decoding of words and non-words can be best predicted by RAN (Cirino et al., 2005). In recent studies, articulation time and pause time in RAN have been major concerns.

2.4.4 Section Summary

Following on from the evidence of the associations between PA, LK, RAN and reading ability, in **section 2.3** justifying their selection for inclusion in a test battery, this section has provided information on testing approaches. The information above can be used to inform the development of tests as components of battery to be used with Arabic speaking children. In **section 2.5** studies which have investigated possible links between selected demographic factors -age, gender and socioeconomic status-, RRS and reading development are reviewed. A particular emphasis is given to the relatively limited amount of research in the Arab world concerning the influence of environmental factors on RRS and reading development. Within this area, the research on one such particular factor, Qur’anic recitation, is examined.

2.5 Demographic Factors, RRS and Reading Development

In addition to investigation of the predictive ability of RRS a great deal of attention has also been paid to exploring the possible links between PA as the key RRS, demographic factors, and reading development, with most studies focusing on the possible effects of gender, socioeconomic status and age. Evidence has emerged that these demographic factors can have significant effects on PA and on the development of reading skills although these may vary across cultures and languages. This section surveys the literature relating to the impact of these demographic factors on PA and their relationship with the development of reading ability, in order to provide the necessary context for this investigation.

2.5.1 Gender

Many factors may have a positive impact on development of literacy skills in young children. Gender is considered not only an important factor for general cognitive ability but also a significant determinant of academic achievement (Matthews, Ponitz, and Morrison, 2009). Previous literature showed that females often have better verbal and linguistic functions compared to males (Maccoby and Jacklin, 1974; Halpern, 1986). Gender differences in PA have not been established as a robust phenomenon, the evidence on this point is inconsistent. In a study by Burt et al. (1999) (*n*=57; 28 boys and 29 girls), typically developing children of mean age 4.4 years (3.4) were assessed and it was found that there was no gender difference in PA and processing skills. They assessed children on eight tasks: consistency of word production, phonological variability according to speech production task, non-word imitation, syllable segmentation, rhyme awareness, alliteration awareness, phoneme isolation and phoneme segmentation. Similarly, another study by Schaefer et al. (2009) that aimed to develop a comprehensive PA test battery for German-speaking preschool children of age 4 to 6 years (*n*=55; 30 boys and 25 girls) also revealed no gender effect.

Contrary to the previous studies' results, Burman et al. (2008) found that female students (age range= 9-15 years) performed better than their male on RAN task, and furthermore acquired higher scores on PA. In addition, a gender difference in PA favouring young girls was observed by Lundberg, Larsman and Strid (2012). Similarly, researchers explored gender factor in phonemic tasks and found that female participants performed better than males in most of the phonemic tasks such as middle and final phoneme detection tasks (Moura et al. 2009), in PA and reading ability (Below et al., 2010), PA and overall reading ability (Chipere et al.,2014).

2.5.2 Age

Age can also have a considerable effect on literacy development and has been covered widely in the literature on PA and reading ability. As reviewed in **section 2.3** each one of the RRS that has been studied in the literature indicated the factors of age and grade as among determinants of reading ability. Previous studies suggest that the relationship between PA skills and levels of speech-sound accuracy strengthens with age and that PA becomes more predictive (McDowell et al., 2007; Mohamed et al., 2019)

According to Lonigan et al. (1998) this is due to the fact that as children grow older, their PA skills become more stable. Moreover, as children develop, they are more likely to be exposed to more systematic instructional or educational activities related to literacy at home or pre-school (McDowell et al., 2007). In this sense, younger children may have a specific level of PA development, but they do not achieve this as a result of having insufficient exposure to activities that trigger this potential (McDowell et al., 2007). Thus, lower PA skills in young children may reflect a lack of capacity or exposure, or both. In older children, however, lower PA skills are more likely to reflect personal capacity, signifying that they are at risk of developing reading problems (McDowell et al., 2007). McDowell et al.'s research corroborated findings from the earlier study by Burt et al. (1999), in which age was found to be significantly correlated with performance on tasks targeting alliteration, non-word imitation, phonological variability, phoneme isolation and segmentation, with older children demonstrating higher levels of PA than their younger counterparts. Another study by Foy and Mann (2001) reported that age was significantly related not only to PA (phonemic awareness and rhyme awareness), but also RAN, vocabulary, LK, and articulation.

Age effect on PA has been observed in different orthographies. For example, in Schaefer et al. (2009) examined German-speaking preschool children (*n*=55; age range 4;0-6;9) on several PA tasks and found that older children significantly outperformed the younger ones on Rhyme-identification-input, Rhyme-production-output, Sound- identification-beginning-input Sound-identification- beginning-output and Letter-knowledge (Burt et al., 1999). Mohamed, et al. (2019) used PA tasks with Arabic speaking children, and their results using multi-variate analysis of variance (MANOVA) showed a significant age effect. Evidence about the developmental sequence of PA among Arabic-speaking children is scarce. Therefore, this research focuses on children aged four to seven years, covering the crucial transition period from less formal learning activities in the pre-school period to more systematic literacy development instruction at primary school, with the aim of investigating the developmental sequence in Arabic-speaking children.

2.5.3 Socioeconomic status (SES)

According to Menkes et al. (2006: 1132) "In general, socioeconomic circumstances play the most crucial role in success at learning to read in children". Snow et al. (1998) previously argued that the process of learning to reading relies upon early childhood experiences of both spoken and written text, as these help to nurture curiosity in young children and expose them to literacy. Studies show that most schoolchildren with learning difficulties in literacy come from vulnerable families where there are high rates of unemployment, casual employment, domestic abuse, restricted social support networks, housing problems/overcrowding, poor quality of home literacy environment (availability of books and hours spent reading) and low levels of educational achievement. These factors influence children's cognitive and linguistic development (Ison, 2004) For this reason, the impact of SES on the development of PA and reading ability has been the subject of many studies which have demonstrated a relationship between SES and PA, confirming that children from low SES backgrounds had delayed PA and literacy development in comparison with those from higher SES backgrounds (Lonigan et al., 1998; Nancollis et al., 2005; Noble et al., 2006; McDowell et al., 2007; McIntosh et al., 2007; Schiff and Lotem, 2010; Lundberg et al., 2012).

Longitudinal studies, such as that of Hecht et al. (2000), attempted to quantify the impact of differences in SES on decoding and reading comprehension skills among children in five American public schools, examining whether differences in social class explained levels of reading-related abilities in kindergarten and subsequent development of reading skills until fourth grade. The reading-related abilities they investigated were PA, rate of access to phonological information in long-term memory, and print knowledge. Hecht et al. (2000) found that the impact of social class on the development of reading skills varied over time. When levels of print knowledge in kindergarten were taken into account, SES was found to have the greatest influence on the development of reading skills (Hecht et al., 2000). Dickinson and Tabors (2001) concluded that in the home literacy environment of children from low-income families, activities such as reading stories at home had immediate and lasting effects on acquisition of reading skills. Factors such as the quantity and quality of books in the home, the number of stories and poems read to children and the number of people who read these all have a significant relationship with children's PA level recorded on school entry (Porta, 2008).

Noble et al. (2006) conducted a study with first graders (n=150) from different SES backgrounds and with varying levels of PA. They found a multiplicative relationship between SES and PA when assessing decoding skills, which suggested that decreased access to resources may increase cognitive risk factors for poor decoding; conversely, greater access to resources may help to buffer reading skills among children with weaker PA. Noble et al. (2006) concluded that attempts to identify the cognitive and experiential factors driving development of reading skills needed to acknowledge the complex, synergizing relations between these factors. In addition, Schiff and Lotem (2010) observed that children with low SES had slower reading ability. When phonological and morphological awareness were controlled, a small difference was observed in reading ability with grade between low and high SES groups. This shows the indirect effect of SES on reading ability. The study results suggested that preschool children from low SES background may have weaker PA skills than their peers of higher SES and consequently poor reading ability.

To date, few large-scale studies have specifically addressed the issue of the influence of SES on levels of PA among pre-school children but Lundberg et al. (2012) tested PA in a sample of over 2,000 six-year-olds, comparing two cohorts on two occasions during the pre-school year. They found a clear SES effect, indicating the influence of early language stimulation in the home environment and reported that those children who had already grasped the alphabetic code and had LK at the beginning of the pre-school year had the highest initial scores on the test.

Krishna Priya et al. (2018) studied the effect of SES on PA in different age groups in Malayalam (an Indian language) speaking children (n=280; age range= 3-7). Participants of this study were divided into four groups, which were divided into sub groups (mid and high) based on their SES. Children from high SES groups scored highest in syllable discrimination, but children from mid SES backgrounds scored lowest in the word segmentation task and highest on the syllable counting task, Overall, children from a mid SES background had lower scores on most reading tasks than children from high SES.

McDowell et al. (2007) in a cross- sectional study examined participants (n=700) of mean age range = 2.59-5.36 years. Participants were identified as from low and high SES based on preschool funding source. All were assessed on two measures of vocabulary, eight measures of phonological awareness, and two measures of speech sound accuracy. Results indicated the existence of a correlation between SES and letter names. Also, age was found to moderate the relation between SES and PA (i.e., relations between SES and PA were amplified with increases in age).

Several studies have indicated that SES and PA are related (e.g. Noble et al., 2006, McDowell, Goldstein, 2007), but not many studies in Saudi Arabia have examined the possibility of an effect of SES on PA. Therefore, the present study attempts to study the PA in Saudi Arabic- speaking children of different socio-economic status.

2.5.4. Socio-demographic factors in Arabic

Very little existing research conducted in Arabic-speaking contexts has focused specifically on the possible relationship between extrinsic factors such as demographic variables and RRS and their influence on the development of children's reading ability. In some cases, there are culturally-related reasons for these gaps in the literature. Much of the focus on the influence of gender and on socioeconomic status in educational studies in many industrialized Western nations since the 1960s has been prompted by broader social and political concerns regarding inequality. In many Arab nations, these debates have only begun to surface as theoretical concerns within academia relatively recently, and the area of gender and inequality remains a particularly contentious issue in many of the more conservative Islamic societies. In addition, whilst socioeconomic status or class has long been viewed as of crucial importance in the Western world as a key factor in determining an individual's life chances, even today in many Arab societies, tribal and clan-based identities are considered more important in determining someone's life chances (Farsoun, 2006).[5]

However, some relevant data could be gathered from existing research on literacy development. In the case of gender, since most schools, even for the younger age groups, are usually single sex in the Arab context, results in studies are often presented by gender, even though comparative analysis may not have formed the focus of the research. The same is true in the case of age-

[5] The collected essays in the book *Arab Society: Class, Gender, Power and Development* edited by Nicholas S. Hopkins and Saad Eddin Ibrahim (2006) Cairo; New York: American University in Cairo Press provide a good overview of ongoing debates in the Arab world and their impact on the Saudi context.

related statistics, since literacy-related studies usually target particular age groups. However, data regarding socioeconomic status of the kind which would now be routinely collected by Western educationalists including parental occupations and educational attainment, family income or area of residence, rarely features in Arab studies.

A study done by Arafat et al. (2017) assessed the role of age and family SES in predicting children's literacy skills in kindergarten and literacy achievements (PA, word writing and text reading) in first grade (n= 109; mean age = 5.72 years, SD = 4.98). They included Arabic speaking Israeli children from a low SES background and found direct effect of SES on PA and word reading, but an indirect effect was observed in text reading. The results not only revealed that early literacy skills of children in kindergarten predict later achievements in first grade, but also highlighted the contribution of age and SES to children's literacy skills in transparent Arabic orthography.

Given that studies conducted outside the Arab world appear to indicate the crucial importance of state versus private sector educational provision and the influence of the home environment in developing literacy skills, the decision was taken to include analysis of this, along with gender and age, in this study.

2.5.5 Socio-cultural factors in Arabic

According to Gillon (2012), PA can be influenced by a wide range of extrinsic factors including socio-cultural variables together with the educational setting and the teaching methods used. As noted in the introduction to this thesis, for several reasons it has been decided to include SES, since no studies on PA have been conducted in the Saudi context.

Very little literature has investigated the impact of socio-cultural factors in determining reading ability in Arabic-speaking children. One key factor is the practice of Qur'anic recitation (*tajwid*) in Islamic cultures.In Saudi schools, children are required to learn to recite the Qur'an from a very young age. Although the purpose of teaching them Qur'anic recitation is, first and foremost, related to spiritual rather than linguistic development, given its central importance in Islamic education, several studies have examined the possible impact of this traditional form of education on RRS and literacy in Arabic. There is some evidence to suggest that this may have an impact on decoding in reading behaviour. In this research, the focus is on those children who are exposed to additional tuition in this area.

Wagner (1994) concluded from his Literacy Project based in Morocco that memorization of the Qur'an helps reading acquisition in Arabic. Boyle (2004) also noted that being able to recite passages before actual decoding can help children with reading, especially those who are at risk of not being able to read fluently. This suggests that this method of religious instruction may also have a significant impact on the way in which young children develop PA. According to Haeri (2011), exposure to this type of Qur'anic teaching improves memory, articulation, fluency, and audio discrimination. Zaid (2011) claims that Qur'anic recitation can encourage reading skills and grammar acquisition, as well as developing creative writing skills.

Two other studies focused on the impact of Qur'anic recitation on bilingual Muslim children in the UK context, where this is still considered an important element of Islamic practice. The importance of the development of decoding and comprehension skills was emphasized by Burgoyne et al. (2009). They argued that the cultural practice of learning to recite the Qur'an promotes a concept of reading that tends to be more focused on decoding skills than comprehension and word meaning (Burgoyne et al., 2009). In their study of Year Three children (aged 7;0-8;0) in the UK who were learning English as their second language, Burgoyne et al. (2009) reported that children who had additional tuition in reading and reciting the Qur'an were observed to demonstrate high levels of decoding skills and low reading comprehension.

Rosowsky (2010) focused on the reading ability of Muslim bilingual secondary school children (aged 11;0-12;0) in the UK, who had learned to read in English as a second language. A difference was observed between reading fluency and reading comprehension among the participants, and Rosowsky also suggests that this shows decoding is a "cultural practice" (2010: 60). He argues that learning to read the Qur'an, a culturally significant practice in which fluency of articulation is prioritized over complete understanding of meaning, accounts for a high level of decoding ability coupled with a low level of comprehension (Rosowsky, 2010). For example, in a test where children were required to substitute a selected word within a sentence with one from a list of options, they frequently chose the word with the highest level of graphophonic similarity to the original word, even when this was incorrect. The bilingual children exhibited a very high level of decoding skills, and mobilized these skills at the expense of semantic and syntactic cues when tackling tests of reading comprehension (Rosowsky, 2010).

However, there are no studies that investigate this phenomenon further among younger children or those who have learnt to read in Arabic as their first language. This study will address this issue, by taking a measure to assess the influence of extra exposure to Qur'an recitation classes on children's RRS.

2.6 Significance of the study

As a practising speech therapist, I am only too aware that despite the vast sums of money spent by the Saudi Ministry of Education on educational provision in the Kingdom[6] and the concerted efforts being made to produce children who can read independently and effectively, many pupils in primary education will never reach the level of proficiency required to ensure success in their future studies. A key problem in the Arab world generally and in Saudi Arabia specifically is that there is currently no systematic approach taken to identifying those children who fail to progress with their reading skills. Consequently, the remedial programmes which would help to address problems with literacy have not yet been put into place.

This creates frustration for teachers, who see some pupils struggling to keep up with their classmates but do not have the tools to diagnose the nature of the problem and devise individualized interventions. It also creates immense worries for parents, who cannot understand why their child is falling further and further behind in their studies. Some parents choose to ignore this problem in the hope that it will eventually disappear. It was my experience of these situations that initially inspired me to focus on this area of RRS research and in particular, to wish to make a contribution to the future development of the much-needed comprehensive standardized reading-skills assessment battery.

[6] In 2014, the amount spent on the education sector accounted for 25% of the total budget in the Kingdom (http://www.us-sabc.org/custom/news/details.cfm?id=1541).

The personal observations motivating this study are supported by Tibi and Kirby (2019)'s observation that 'The nature of reading instruction in Arabic classrooms needs to be studied, and instructional studies designed to improve the teaching of PA and decoding are needed'. The development of appropriate instruction and remediation programmes depend on evidence of skills that support literacy development, and standardized screening and assessment tests to children who have difficulties with these.

As demonstrated in **section 2.3** a plethora of studies have been conducted on reading skill acquisition by children whose first language is English, and there is a growing body of research on other languages including Arabic. Research evidence on English is complemented by standardized assessments of reading and reading-related skills (see **section 2.4**). While research on English and other language contexts is beneficial for understanding generally how children learn to read, every language is unique in terms of its features and also with regards to its socio-cultural setting. The research conducted with 4-7-year-old Saudi children in this study will add to the existing literature on the Arabic language, largely involving children aged 5 years and above for whom literacy acquisition should be well underway. This will also provide new evidence on the development of PA, LK and RAN in Saudi children and will contribute to a greater understanding of the relationship between these RRS and reading skills according to teacher report.

Studies of this type not only add to the existing knowledge in this important area, but also help to advise policy-makers and curriculum planners, and to inform teachers and other educational practitioners who are responsible for developing literacy skills. A key objective of this research is to develop and trial a psychometrically robust assessment battery of RRS for Saudi children aged 4;0-7;0 years old, producing a tool that can help to identify children with reading-related difficulties, who are at risk of later reading difficulties, and lay the foundations for a fully standardized measure. As Tibi and Kirby (2019) argue, 'because of the absence of standardized measures in Arabic, many of the measures used do not have demonstrated psychometric properties of use in a monolingual Arabic-speaking population' (p.1009). Having studied children in Grade 3 (mean age approximately 8 years), they also believed that 'inclusion of earlier grades could provide a picture of how the predictors [of reading skills] develop' (p.1009). The RRS battery developed in this study will go some way to addressing both these gaps in the research evidence, and will allow timely interventions and targeted remedial help to be made available to children at risk.

2.7 Research Aims, Objectives, Questions and Hypotheses

The broad aim of this research was to develop and trial a comprehensive battery of tests to assess reading-related skills (RRS) in Arabic-speaking 4;0-7;0 year-olds that is reliable and informative and could serve as the foundation for the future development of a comprehensive standardized RRS assessment battery designed for use in the Saudi context.

Study One aimed to develop a RRS test battery and use this to obtain normative data and explore the influence of biological factors (gender and age) and environmental factors (socioeconomic status and exposure to the study of Qur'anic recitation) on the development of RRS (Frith, 1995) in a sample of typically developing 4;0-7;0 year-old Arabic-speaking children in Saudi Arabia.

Study Two aimed to evaluate the reliability and validity of the RRS test battery developed for Study One.

The research objectives were:

1. To create a test of RRS, including phonological awareness (PA), letter knowledge (LK), and rapid automatized naming (RAN), drawing on previous research in other languages (particularly English) and taking into account the features of Arabic.
2. To provide normative data for clinical assessment and as a foundation for future standardisation of the tests
3. To provide evidence of the effects of biological and environmental factors on children's performance, and demonstrate that the battery is sensitive to age and therefore informative about development of RRS.
4. To provide evidence of test-retest reliability for all RRS tasks.

5. To provide preliminary evidence of the validity of the RRS measures in relation to teacher ratings of children's reading abilities, with potential to identify at risk of developing reading difficulties.

The following research questions were derived from these objectives:

- **RQ1:** To what extent does gender affect performance in RRS as measured by the RRS assessment battery?
- **RQ2**: To what extent does age affect performance in RRS as measured by the RRS assessment battery?
- **RQ3:** To what extent do socioeconomic variables (school type, parental education, and a composite category combining property type and family income) predict performance in RRS as measured by the RRS assessment battery?
- **RQ4:** To what extent does the amount of exposure to the study of Qur'anic recitation affect the performance of children in RRS as measured by the RRS test battery?
- **RQ5:** To what extent is the RRS test battery a reliable and valid instrument for measuring RRS (PA, LK and RAN) in Arabic-speaking children and identifying those at risk of developing reading-related difficulties?

The hypotheses evaluated in the two studies are rooted in the research findings reviewed in this chapter, summarised below:

2.7.1. Gender

Gender has been widely shown in the literature to have an impact on literacy development (Nancollis et al., 2005). The preponderance of evidence indicates that girls generally perform better than boys in terms of verbal and linguistic functions (Maccoby and Jacklin, 1974; Halpern, 1986).

Some studies reported gender differences in PA (e.g. Burman et al, 2008; Lundberg et al., 2012, Chipere et al. 2014, Mohamed et al., 2019) and rapid naming (e.g. Burman et al, 2008). As a result, it is expected that in this study, girls will perform better than boys in PA, LK and RAN tasks. Thus it is hypothesized that:

H1: Gender will affect children's performance in RRS, with girls gaining significantly higher scores in PA, LK and RAN tasks.

2.7.2 Age

As discussed previously, age has been found significantly associated with the literacy development of children. The literature indicates that age is correlated with performance in PA tasks (Burt et al., 1999; Anthony et al., 2002; McDowell et al., 2007 and Duranovic et al., 2012). There is evidence that the relationship between PA and speech sound accuracy skills becomes stronger with age (McDowell et al., 2007). As children get older, PA skills begin to stabilize (Lonigan et al., 1998). Levels of exposure to more formal instructional or educational activities related to literacy, either at home or pre-school, also influence development of PA. Age has been shown in Arabic speaking children (Al-Sulaihim and Theo, 2017; Mohamed et al., 2019). Hence, it has been assumed that children within the higher age bands will perform better than those in the lower age bands in PA, RAN and LK tasks. Thus it is hypothesized that:

H2: Age will affect children's performance in RRS, with those in higher age bands gaining significantly higher scores in PA, LK and RAN tasks.

2.7.3 Socioeconomic Status

Although many studies have examined the impact of SES on reading ability, relatively few have investigated the influence of SES on reading ability via its impact on PA, LK, RAN, especially among pre-school children (Lundberg et al., 2012). The literature in this area has shown a clear relationship between SES and PA (Nancollis et al., 2005; McDowell et al., 2007; Lundberg et al., 2012, Arafat, et al., 2017). Several of these studies have indicated that children from low socioeconomic backgrounds had delayed PA and literacy and RRS development in comparison with their counterparts from higher socioeconomic backgrounds (Nancollis et al., 2005; Noble et al., 2006; McIntosh et al., 2007; McDowell et al., 2007). As a result, it is hypothesized that:

H3: Socioeconomic status will predict children's performance in RRS, with those from higher socioeconomic backgrounds achieving significantly higher scores in PA and LK tasks than those from lower socioeconomic background.

2.7.4 Exposure to Qur'anic recitation

As in some other Islamic nations, Qur'anic recitation is an obligatory element of the curriculum in Saudi Arabia for children from a very young age. A limited range of studies have suggested that the way in which Qur'anic recitation is taught and the type of skills acquired when performing this practice may impact positively on particular RRS, including decoding and PA, which may in turn help to enhance reading ability and overcome reading difficulties (Wagner, 1994; Rosowsky, 2001; Robertson, 2002; Boyle, 2004; Burgoyne et al., 2009; Haeri, 2011; Zaid, 2011).

As a result, it has been assumed in this study that those Saudi children who have extra exposure to Qur'anic recitation will demonstrate better reading ability overall. Thus is hypothesized that:

H4: Children who have extra exposure to study of Qur'anic recitation will perform better in RRS (PA, LK and RAN), when controlling for age.

2.7.5 RRS (PA, LK and RAN) as predictors of reading ability

2.7.5.1 Phonological Awareness

The relationship between PA and early reading has been recognized in the literature (Stahl and Murray, 1994, Caravolas et al., 2005; Ziegler et al., 2010). PA has been shown to play a fundamental role in enhancing the alphabetic reading ability of children (Adams, 1990; Goswami and Bryant, 1990; Stanovich, 1994; Lyon, 1995; Gottardo et al., 1996; Muter and Snowling, 1998; Foy and Mann, 2003). The relationship between PA and reading ability has been found in many orthographies (e.g. Seymor, 2003; Patel, 2004; Melby-Lervåg, et al., 2012). This linkage has also been demonstrated in studies focusing on the relatively transparent vowelized Arabic script used with young learners. PA has been found to be related to a variety of reading measures in Arabic in kindergarten to Grade 3, especially word reading accuracy (Al-Mannai and Everatt, 2005; Asaad and Eviatar, 2014; Layes, Lalonde, and Rebai, 2015; Taibah and Haynes, 2011). In a recent study by Tibi and Kirby (2019), PA was a significant and unique predictor of every reading outcome, after controlling for age, gender, cognitive ability, and vocabulary.

PA is the predictor that has been studied most often in Arabic. Almost all studies have found it to be significantly associated with real word reading in the elementary school grades (Abu Ahmad, Ibrahim, and Share, 2014: Grades K–2; Al-Mannai and Everatt, 2005: Grades 1–3; Asaad and Eviatar, 2014: Grades 1, 3, and 5; Asadi and Khateb, 2017: Grades 1 and 2; Asadi et al., 2017: Grades 1–6; Elbeheri and Everatt, 2007: Grades 4 and 5; Elbeheri, Everatt, Mahfoudhi, Abu Al-Diyar, and Taibah, 2011: Grades 2–5; Smythe et al., 2008: Grade 3; Taibah and Haynes, 2011: Grades K–3). Children's performance in PA tasks has been found to correlate with reading skills development (e. g., Georgiou et al., 2008; Pennington et al. 2001). Deficits in phoneme awareness cause difficulties in learning to read (Warmington and Hulme 2012). As a result, it has been assumed that children achieving high scores in PA tasks will demonstrate better reading ability (see H5).

2.7.5.2 Letter Knowledge

Although this has been the subject of fewer studies than PA (Foulin, 2005), LK is also regarded as a significant skill in early literacy development. Past studies on literacy acquisition have demonstrated that LK is a powerful pre-school predictor of learning to read (Foulin, 2005), sometimes even the best single predictor (McBride-Chang, 1999, Share et al., 1984).

In addition LK has also been found to predict reading ability concurrently and longitudinally (Hogan, Catts and Little 2005; Johnston et al. 1996; Muter et al. 1998, 2004).

The predictive relationship between pre-school LK skills and school reading skills has been verified in Arab studies. For example, in a study done by Al-Sulaihim and Theo (2017) a significant association between LK and reading was observed over time and the authors believed that LK is an important factor for acquisition of early reading skills.

As a result, it is assumed that children achieving higher scores in the LK task will demonstrate better reading ability (see H5).

2.7.5.3 Rapid Automatized Naming

The literature review has suggested that in addition to PA, RAN is also an important element in the development of word decoding (Taub and Szente, 2012) and it has been demonstrated to have a significant association with reading ability (Denckla and Rudel, 1976; McBride-Chang and Manis, 1996; Kirby et al. 2003; Bowey et al., 2005; Leopla et al., 2005; Taub and Szente, 2012; Warmington and Hulme, 2012).

In Arabic studies, Naming speed has been found to predict word reading speed or fluency (Asaad and Eviatar, 2014; Saiegh-Haddad, 2005), as well as also word reading accuracy (Al-Mannai and Everatt, 2005), in studies covering kindergarten to Grade 5. Taibah and Haynes (2011) found that both RAN predicted a variety of reading skills in Arabic speaking children (K-3 Grade).

In most studies, RAN was a significant predictor of real word reading (Taibah and Haynes, 2011), reading fluency (Tibi and Kirby, 2018; Taibah and Haynes, 2011). In a recent study, Tibi and Kirby (2019) found that RAN was a significant and unique predictor of every reading outcome, after controlling the effects of age, gender, cognitive ability, and vocabulary.

Since RAN is considered to be an important predictor of different aspects of reading skills in children, it is assumed that children who achieve high scores in RAN tasks will demonstrate better reading ability. In view of the above research evidence in the predictive power of PA, LK and RAN generally and in Arabic (vowelized orthography) specifically, the following hypothesis is developed:

H5: Children with higher teacher ratings of reading abilities will perform significantly better on all RRS tasks than children with lower teacher ratings.

The aims, objectives, questions, hypotheses, statistical tests of this study are all aligned visually Table 2.5:

Table 2.5: MAPPING OF RESEARCH AIMS, OBJECTIVES, QUESTIONS, HYPOTHESES AND STATISTICAL TESTS USED

Study One aimed to develop a RRS test battery and use this to obtain normative data and explore the influence of biological (gender and age) and environmental factors (socioeconomic status and exposure to the study of Qur'anic recitation on the development of reading-related skills in a sample of typically developing 4-7-year-old Arabic-speaking children in Saudi Arabia.	• To create **a test of RRS, including PA, LK, and RAN,** drawing on previous research in other languages (particularly English) and taking into account the features of Arabic.			
	• To provide evidence of the effects of **biological (age-gender) and environmental factors (Socioeconomic status-extra exposure to study of Qur'anic recitation)** on children's performance, and demonstrate that the battery is sensitive to age and therefore informative about development of RRS.	• **RQ1:** To what extent does **gender** affect performance in RRS as measured by the RRS assessment battery?	**H1: Gender** will affect children's performance in RRS, with girls gaining significantly higher scores in PA, LK and RAN tasks.	Unrelated T-test: to determine if **gender** significantly affected scores for each of the RRS tests.
		• **RQ2:** To what extent does **age** affect performance in RRS as measured by the RRS assessment battery?	**H2**: **Age** will affect children's performance in RRS, with those in higher age bands gaining significantly higher scores in PA, LK and RAN tasks.	One-way ANOVA with **age** group as the independent variable (3 groups: 4;0-4;11, 5;0-5;11, 6;0-7;0) and scores for each of the RRS tests as the dependent variable.
		• **RQ3:** To what extent do socioeconomic variables (school type, parental education, and a composite category combining property type and family income) predict performance in RRS as measured by the RRS assessment battery?	**H3: Socioeconomic status (SES)** will predict children's performance in RRS, with those from higher socioeconomic backgrounds achieving significantly higher scores in PA and LK tasks than those from lower socioeconomic backgrounds.	Correlation between three key **SES** factors (school type, parental education, and property type + family income) and children's performance in PA, Blending and LK. Regression Analysis to identify which **SES** measures (school type, parental education, and family income +property type) can predict performance on the outcome measures (composite PA measure, Blending and LK.

		• **RQ4:** To what extent does the amount of exposure to the study of Qur'anic recitation affect the performance of children in RRS as measured by the RRS test battery	**H4:** Children who have **extra exposure to study of Qur'anic recitation** will perform better in RRS (PA, LK and RAN), when controlling for age.	Two- way ANOVA using levels of exposure to **study of Qur'anic recitation** as the independent variables and scores for each of the RRS tests as the dependent variable)
Study Two **aimed to evaluate the reliability and validity of the RRS test battery developed for Study One.**	• To provide **evidence of test-retest reliability for all RRS tasks.**			Intra-class Correlation Coefficient (Cronbach's alpha values above 0.7 for each task).
	• To provide preliminary evidence of **validity of the RRS measures in relation to teacher ratings of children's reading abilities**, with potential to identify those at risk of developing reading difficulties.	• **RQ5** To what extent is the RRS test battery a reliable and valid instrument for measuring RRS (PA, LK and RAN) in Arabic-speaking children and identifying those at risk of developing reading-related difficulties?	**H5**: Children with higher teacher ratings of reading abilities will perform significantly better on all RRS tasks than children with lower teacher ratings.	Two-way ANOVA/ANCOVA with ability group as the independent variable (two groups: highest-ranking children vs lowest-ranking children, according to teacher-rated reading ability, with scores for each task as the dependent variable).

2.8 Conclusion

The literature reviewed in this chapter indicates that, in keeping with theoretical models of the skills entailed in decoding written language, there is a broad and long-standing tradition of research investigating the development of RRS such as PA, RAN and LK, and their relationship to reading. This research has shown that PA, RAN and LK are related to concurrent decoding skills in many languages, including several varieties of Arabic, and are predictive of later decoding skills. The causal relations between RRS and decoding have been shown to be complex and reciprocal. Most notably, the acquisition of literacy skills including letter knowledge supports the more complex and explicit levels of awareness measured in tasks tapping phoneme-level awareness, which is crucial for accurate decoding of unfamiliar words. Whether RRS support or are supported by decoding skills, these research findings demonstrate the importance of assessing RRS to identify children at risk of reading difficulties and guide intervention to support their RRS and reading development. Recent research confirms the need for a standardised assessment of RRS in Arabic to compare children's skills with those expected at their age. In line with Frith's comprehensive model of reading which underpins the English Phonological Assessment Battery (PhAB) (Frederickson et al., 1997; see **section 2.2.5**), it is important to establish extrinsic factors that influence children's skills in order to take these into account when comparing individual children with norms. The research reported in this thesis goes some way to addressing this need for an evidence-based assessment in Arabic by developing a comprehensive battery of RRS, collecting normative data in children aged 4-7 years, evaluating test reliability, and investigating a range of extrinsic factors.

The questions and hypotheses addressed in this research are linked to the adapted version of Frith's (1995) comprehensive framework of reading skill development presented in **section 2.2.5**. Questions and hypotheses 1 and 2 relate to the impact of biological factors of gender and age on behavioural measures of RRS (PA, LK and RAN), while 3 and 4 relate to the impact of environmental factors, namely SES and exposure to Qur'anic recitation. Finally, question and hypothesis 5 focus on the reliability and validity of the behavioural measures in the RRS battery as indicators of children's skills.

The following chapters present the methods and results of the two studies addressing these questions and hypotheses.

Chapter Three introduces the overall methodological approach adopted in this investigation and then provides a detailed description of the methodology for Study One, together with the results of this study. It charts the process of developing the RRS test battery designed to investigate performance levels and PA, LK, and RAN profiles in typically developing Arabic-speaking children from different socioeconomic backgrounds in Riyadh, Saudi Arabia. The pilot study that was used to trial the test battery is also described in this chapter, together with the participants, materials, and testing procedures employed in Study One. The chapter also charts the development of the questionnaire devised to capture demographic and socioeconomic data, and the methods employed to codify and analyse all the data collected. The final element of Study One examines the possible influence that the amount of exposure to Qur'anic recitation may have on the performance in RRS of children who receive additional tuition in this area. The results of Study One are also presented. Descriptive analyses of the statistical tests conducted are followed by analyses of the inferential statistical tests. It also provides the results of the regression analysis used to identify key factors affecting RRS development in Saudi children.

Chapter Four presents a detailed description of the methodology for Study Two, which was designed to validate the RRS test battery and assess its predictive abilities. The results of Study Two are also presented, including both descriptive and inferential statistical analyses.

Chapter Five presents an analytical reflection on the data which were presented in the previous chapter, considering in detail the five hypotheses H1-H5 which the two studies conducted in this research were designed to test. It determines the extent to which these can be said to be supported or not on the basis of the findings. At the same time, it compares and contrasts findings from this research conducted in an Arabic-speaking Saudi context with evidence from other cultural and linguistic settings, attempting to account for any results which appear to contradict previous studies in this area.

In **Chapter Six**, conclusions drawn from the research findings are discussed in relation to the implications for use of the battery as an assessment tool of RRS. Limitations of the study are also considered, and directions for future studies are suggested.

3 CHAPTER THREE: STUDY ONE

3.1 Introduction

This chapter begins by presenting the overall methodological approach adopted in this investigation which aims to investigate performance levels and the profiles of PA, LK, and RAN in typically developing Arabic-speaking children in Riyadh, Saudi Arabia. The first element of Study One explored the potential impact of gender, age, and SES on participants' levels of RRS while the second element examined if exposure to additional tuition in Qur'anic recitation affected these levels. First, the research approach is presented (**section 3.2**) followed by a detailed account of the design of Study One (**sections 3.3 to 3.7**) and development of materials and testing procedures employed (**sections 3.8-3.11**). The methods used to codify and analyse the data collected are also explained (**section 3.12**) together with the ethical dimensions which need to be considered in research of this type (**section 3.13**). The results of the statistical analysis for both elements of Study One are presented in **section 3.14**.

3.2 Research Approach

3.2.1 Inductive vs. deductive

Academic research may be divided into two distinct strands: inductive and deductive. Inductive research tends to develop hypotheses as the investigation is in progress, and often entails a more open-ended, exploratory form of study (Ritchie and Lewis, 2010). It is characterized by investigative approaches that seek to identify patterns in the data and create working hypotheses and theories from this process (Ritchie and Lewis, 2010). It also typically focuses on data relating to the subjective worldview of the participants (Ritchie and Lewis, 2010). In contrast, deductive reasoning is employed when the investigation begins from fixed hypotheses that can be tested (Ritchie and Lewis, 2010). A deductive research approach, often described as a 'top-down' approach, is particularly useful for testing hypotheses and theories derived from an existing body of research. The available literature on PA and other RRS has advanced various hypotheses and theories which have been incorporated into the research questions addressed by this research, suggesting that this investigation was well-suited to a deductive research approach. These hypotheses are stated in the Literature Review.

3.2.2 Qualitative vs. quantitative

A further key division between research approaches concerns the use of quantitative versus qualitative research strategies and methods. A qualitative strategy “usually emphasises words rather than quantification in the collection and analysis of data” (Bryman cited in Harmmersley, 2013: 1) and is most appropriate for studying attitudes and examining people’s views, beliefs and interests. Qualitative data usually depend on observation, interview, and spoken or written accounts such as diaries (Landy and Conte, 2010: 61).

One of the strengths of qualitative research is that it examines the experience as lived by participants in real and natural contexts (Klenke, 2008) and qualitative methods usually require researchers to observe a phenomenon in its normal context, taking a naturalistic approach, without manipulation or interference. For example, when investigating classroom behaviour, researchers would do this *in situ*, observing behaviours as they occur in their normal context. The socially constructed nature of what is studied, the relationship between the researcher and what is studied, and the situational constraints that affect the inquiry are all important in qualitative research (Houser, 2014). Consequently, it has been claimed that one of the advantages of qualitative methods over quantitative ones is that the former provide “deeper insights into the user experience, answering more explanatory questions” (Gramatikov and Barendrecht, 2010: 48).

This has encouraged the claim that quantitative research methods are superficial (Rubin and Babbie, 2009) and are used to conduct studies in a “contrived context” (Bryman, 2012: 408). Despite this limitation, it was decided that quantitative methods were the most appropriate for these studies. Quantitative methods centre on the collection and analysis of numerical data (Creswell, 2003) gathering these from tests, rating scales and questionnaires. A quantitative methodology is positivist in nature and attempts to create standardized data by observing phenomena independent of their context (Creswell, 2003). Whilst qualitative researchers attempt to study the complexity of the natural setting, quantitative researchers make efforts to “control the extraneous variables in the natural environment” (Houser, 2014: 79), their aim being to create reliable, generalizable outcomes. Quantitative research is compatible with statistical analysis, which facilitates efficient and quick results (Creswell, 2003). By following well-established procedures to ensure reliability and validity, research can be replicated and then analysed and compared with similar studies across categories and over time (Houser, 2014). Quantitative researchers aim to avoid personal bias by keeping a distance between themselves and participating subjects by using participants who are unknown to them (Creswell, 2003).

3.3 Research Design

This research design may be described as cross-sectional. Unlike longitudinal research design, cross-sectional research involves looking at a population at a given moment in time, rather than taking multiple measures over a long period. In essence, it offers a snapshot of the population under investigation, rather than attempting to test the same group repeatedly at intervals (Bowling, 2009). To answer the research questions, this study aimed to take a snapshot of a population of Saudi children from a range of backgrounds and compare the results among different subgroups of gender, age, SES and levels of exposure to tuition for Qur’anic recitation within that population. An outline of Study One is presented in **Table 3.1**.

	DATE	EVENT	PURPOSE
Pilot study: Battery	February 2010	Pilot study	Test construction and initial testing (34 children)
	March 2010		Reviewing and updating test
Pilot study: Questionnaire	April 2010	Pilot questionnaire	Feedback sought from colleague, sociologists, and parents
Recruitment Phase	September 2010	Initial contact with administrative centres for Riyadh (NSEW)	Distribution of introductory letters about research, Parental Information and Consent Form

	DATE	EVENT	PURPOSE
	October 2010	Second contact with administrative centres	Confirmation of schools willing to participate
	October/ November 2010	Contact with administrators for participant schools	Explanation of research aims and provision of parent questionnaire. Participants meeting criteria identified by relevant party at school
Data Collection	November 2010-June 2011	Data collection	Testing participants, maintaining contact with schools and asking parents to complete questionnaire

Table 3.1: An outline of the research design

3.4 Assessing Factors Affecting RRS

Study one involved testing boys and girls from a sample of private and public kindergarten and schools in the Saudi capital, Riyadh, whilst their parents were asked to provide information about their socioeconomic status. This section details how Study One participants were sourced, approached and selected. An additional component was added at a later stage to this study which focused on whether the level of exposure to tuition in Qur'anic recitation might be another factor which would impact on performance in RRS. This component is covered below (**section 3.11**) and required teacher input to the study.

3.5 Consent

Before beginning this investigation, consent was sought from those identified as key stakeholders. Initially, the Saudi Ministry of Education was contacted requesting permission to conduct the study and permission was granted (see Appendix A). Schools in Riyadh were sampled according to the procedure detailed below and administrators at the selected schools were invited to participate in the study by letter. This invitation included a brief summary of the study design and rationale, and schools were asked to give their consent to participate in the investigation (see Appendix B).

Once ministerial consent had been received, school administrators were asked to identify potential participants who met the inclusion criteria. Consent for their child to be assessed was then sought from parents of potential participants

(Appendix C). The researcher had no personal contact with parents or children until they had been successfully recruited by the school official.

Before conducting the battery of tests, the researcher obtained consent from children using a specially designed information sheet which used visuals to explain in a simplified form the nature of the study and what participants would be required to do (Appendix D). The researcher read through the form with each child, ensuring they still wished to consent to participate. Informed consent was thus obtained from all key stakeholders in the research process.

3.6 Sampling method

Riyadh was selected for this study because of its central importance as the capital of Saudi Arabia, and its demographic diversity, allowing a broad cross-section of Saudi society to be represented in the study. The target population was children attending kindergarten and elementary schools (first grade) in the capital. A representative sample of participants was selected using an adapted cluster sampling method in which the population is divided into sub-populations (Bowling, 2009). Units of interest are grouped together in clusters and then these are typically sampled using simple or systematic random sampling (Bowling, 2009) as illustrated in **Figure 3.1**.

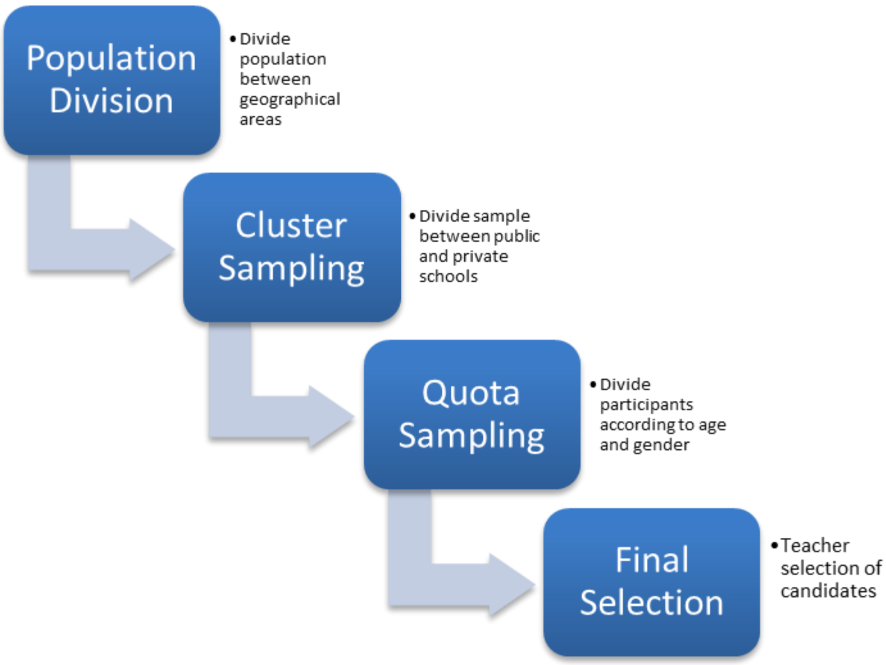


Figure 3.1: Sampling Procedure

In terms of educational administration, Riyadh is divided into four areas: North, South, East and West. The researcher contacted the administrative centre for each of these geographical areas to outline the aims of the research project, answer any questions about it and confirm that it had received ministerial approval. Each centre was then provided with copies of the letter for schools and parental consent forms and asked to invite all the schools in the area to participate, thus avoiding any bias in school selection. Each centre was given a two-week timeframe to obtain consent from schools wishing to participate to ensure that the study kept to the envisaged timeframe. After the two-week deadline, from a possible total of 153 kindergarten and elementary schools in Riyadh, some 47 agreed to participate. Their geographical distribution is shown in **Table 3.2**.

Riyadh Area	No. of schools	Potential participants
North	79	19[7]
South	16	8
East	24	10
West	34	10
Total	**153**	**47**

Table 3.2: Population by area and number of schools consenting to participate

Some potential participant schools were excluded by the researcher because they did not have suitable facilities where tests could be run or had insufficient numbers of pupils meeting the selection criteria. Participating schools were selected from this short-list to reach the total target number of participants required. A non-random cluster-sampling method was applied to divide the sample between public and private schools and the final number of schools selected from each area was distributed as shown in **Table 3.3**. An extra public school had to be found in the north to ensure that all the quotas for gender and age range would be met for a representative sample.

[7] The relatively low level of responses from North area schools can be explained by the fact that many failed to respond within the time limit set although further responses were received after that had elapsed.

Riyadh Area	Public	Private
North	5	4
South	4	4
East	4	4
West	4	4
Total	**17**	**16**

Table 3.3: Public/private school distribution in sample

A non-random quota sampling method was used to determine the subjects required for the study using available data (numbers of males and females, and different age bands), in order to produce a representative sample of these groups, according to their distribution in the population as a whole (Bowling, 2009).

Once the list of participating schools had been finalized, potential candidates for the study were identified according to the selection criteria:

1. The child was aged 4;0-7;0;
2. Both of the child's parents were of Saudi nationality;
3. Arabic was the first language of the child and both his/her parents;
4. Arabic was the main language spoken in the family home;
5. No concerns existed regarding the child's speech and language development.

A further check was made to verify these criteria had been correctly and consistently applied.

Once the children eligible to participate in the study had been identified, the researcher used a quota sampling method to divide up participants representatively according to area, type of school (public or private), gender and age. This produced a representative sample of some 384 children in total.

All parents of these study participants were then surveyed by telephone by the researcher to ascertain information about the family's background and socioeconomic status (SES) and their child's linguistic and educational development.

3.7 Participants

The key characteristics of the final representative sample of participants in Study One are summarized in **Table 3.4** with a total of 384 students being identified as being eligible for testing.

<table>
<tr><th colspan="4">PARTICIPANTS BY SECTOR AND AREA</th><th colspan="2">GENDER</th><th colspan="3">AGE</th></tr>
<tr><th>N</th><th>S</th><th>E</th><th>W</th><th>Male</th><th>Female</th><th>4;0-4;11</th><th>5;0-5;11</th><th>6;0-7;0</th></tr>
<tr><td colspan="4">PUBLIC SECTOR</td><td rowspan="2">96</td><td rowspan="2">96</td><td rowspan="4">123</td><td rowspan="4">128</td><td rowspan="4">133</td></tr>
<tr><td>48</td><td>48</td><td>48</td><td>48</td></tr>
<tr><td colspan="4">PRIVATE SECTOR</td><td rowspan="2">96</td><td rowspan="2">96</td></tr>
<tr><td>48</td><td>48</td><td>48</td><td>48</td></tr>
<tr><td colspan="4">TOTAL = 384</td><td colspan="2">TOTAL = 384</td><td colspan="3">TOTAL = 384</td></tr>
</table>

Table 3.4: Characteristics of the Study One sample

	NORTH		SOUTH		EAST		WEST	
AGE	**M**	**F**	**M**	**F**	**M**	**F**	**M**	**F**
4;0-4;11	**15**	**15**	**15**	**16**	**16**	**16**	**14**	**16**
5;0-5;11	**15**	**17**	**15**	**15**	**16**	**16**	**18**	**16**
6;0-7;0	**18**	**16**	**17**	**17**	**16**	**16**	**16**	**17**
TOTAL	**48**	**48**	**47**	**48**	**48**	**48**	**48**	**49**

Table 3.5: Breakdown of gender and age by region

Equal numbers of participants were recruited from each of the four areas of Riyadh (i.e. North, South, East and West) and numbers of participants from each of these metropolitan areas were equally split between children attending private sector schools (*n*=48) and those attending public sector schools (*n*=48). It is worth noting that regardless of sector, all participants are taught the Saudi National Curriculum through the medium of Arabic but supplementary English classes are offered in the private schools. The gender and age breakdown by region is shown in Table 3.5. The age bands for the sample were set at four to four years eleven months (4;0-4;11 = Kindergarten second year), five to five years eleven months (5;0-5;11 = Kindergarten third year) and finally six to seven years (6;0-7;0) (elementary school first grade).

3.8 Design and Development of the RRS Test Battery

3.8.1 The need for the test battery

This section describes the processes and methods employed to create the RRS (PA, LK and RAN) test battery in Arabic that formed the primary research tool for Study One. The process of questionnaire development which was also a key component of Study One is examined later (**section 3.10**).

In creating an appropriate test battery for this investigation, the intention was to replicate the approaches, conventions, and methodologies used to assess PA, LK and RAN that have been developed in the UK, the US, New Zealand and other English-speaking environments (Stahl and Murray, 1994; Yopp, 1988; Gillon, 2012). While some RRS procedures and methodologies have been developed in Arabic-speaking contexts, in comparison with those developed for English-language speakers, most have lacked a sufficient degree of rigour, depth, quality, and comprehensiveness (Taibah and Haynes, 2006). There is a demand from many sources to develop standardized assessment tools in Arabic that would be appropriate for wider application.

The tests currently available in the Arabic-speaking context may be categorized as either formal or informal. Formal, albeit non-standardized, tests are usually based on English-testing materials that have been translated into Arabic (Taibah, 2006). These include the Receptive-Expressive Emergent Language Scale-2 (REEL-2) test developed by King Faisal Specialist Hospital speech therapists, together with the PA and RAN tasks developed by Taibah (2006), an Arabic adaptation of the Comprehensive Test of Phonological Processing (CTOPP), originally developed for English-language speakers by Wagner et al. (1999). The process of developing test batteries for Arabic-speaking contexts has not been properly documented and few have been subjected to appropriate validation.

Informal tests, by contrast, have usually been developed by clinicians based on their personal experience, one example being the Arabic Articulation Check List developed and used in-house at the King Faisal Specialist Hospital and Research Centre. Using these informal tests can result in a lack of coordination and consistency across clinical centres, which work largely in isolation from one another.

Both formal and informal tests suffer from weaknesses and limitations. Although the former, in particular REEL-2, can be considered more advanced than other tools in the Arabic-speaking context, the fact that these tests have been directly translated from English is a major drawback. Given that the structures of English differ considerably from those in Arabic, simply translating English-based tools into Arabic is ineffective. As Geisinger (1994) observes, tests of this kind require careful adaptation, rather than direct translation, not least in order to avoid unintentional cultural bias when dealing with another population operating in an entirely different cultural and linguistic context.[8]

Informal testing is, by its very nature, highly subjective and inconsistent, leading to a high risk of unreliable data. Since these assessments have not been standardized it is difficult, if not impossible, to objectively identify performances below the normal range, compare performances across age groups or track progress in individuals.

This dearth of standardized assessments in Saudi Arabia means that there is an obvious and important gap in knowledge about the acquisition of RRS in Arabic-speaking children. Consequently, this means Saudi children are disadvantaged because deficits in PA, LK or RAN linked to dyslexia and other reading and writing difficulties cannot be professionally detected and may be left untreated. This lack of practical tools to guide intervention can have an enduring and profoundly negative impact on children's development, education and social life.

3.8.2 Standardisation of the Saudi test battery

Having identified the need for the test battery, this study aims to provide a basis for formal, standardized, valid, and reliable RRS assessment tools, specifically designed and shaped to meet the needs of Saudi children and developed in line with existing best practice. In creating the assessment tools, this study followed the criteria prescribed by Gillon (2012) for producing valid and reliable

[8] See guidelines of the Multilingual Affairs Committee of the International Association of Logopedics and Phoniatrics (IALP) concerning test translation. Available online at: www.http://ialpasoc.info

standardized tests that are applicable in multiple contexts. These criteria include the following:

- A description of the sample population is needed which should include details such as SES and geographical residence (to be aware of possible linguistic variations).
- The sample size for standardized tests must be at least 100 individuals to allow for the degree of diversity and complexity needed for a representative perspective.
- A range of measures must be put in place to test the reliability and validity of the study.
- Information concerning variability in participants' performance is required.
- Administrative procedures and test scoring must be recorded and explained.

Although the test battery developed as part of this research follows Gillon's standardisation principles, it does not strictly conform to the more exacting definition of a standardized test framed by Fischer and Milfont (2010) because the time and finance required to satisfy this were not available for this study. They note that the aim of standardisation is to "reduce the influence of any extraneous variable on the test or experimental performance of participants" (Fischer and Milfont, 2010: 89) and that this should apply equally to procedure, interpretation and scoring. According to Fischer and Milfont (2010), scores in psychological tests are typically assessed against norms derived from a representative sample of the population, meaning that the researcher should usually define the characteristics of the population before administering the tests.

In the case of this investigation, the sample was drawn from only one region of Saudi Arabia and is therefore only representative of the Riyadh area. Thus, although the sample was produced in a representative way, it was not created on the basis of population distribution, resulting in a normative sampling process rather than true standardisation.

In addition, no initial audiology assessment was administered to participants, since this would have proved too costly in terms of resources. It is possible, therefore, that some participants had mild hearing problems, or were affected by some temporary illness during the course of the investigation which meant they were unable to hear the words being used sufficiently well. This may also have had an impact on the results. However, best practice as identified by Gillon (2012) was followed wherever feasible within the practical resource constraints imposed on the study.

3.8.3 Validity and reliability

When creating the test battery, every effort was made to ensure that the test was valid and as reliable as possible. Each task was developed in consultation with practitioners and with reference to previous studies. Although no clear precedent in Arabic existed for many of the tests described above, they were nevertheless considerably informed by learning from previous research (Gillon, 2012). In addition, the new tests were piloted to reveal any problems with the data collection, procedure and content. Each test began with practice items before the main assessment task in order to ensure each participant was fully aware of the process and demands of the task.

It was important to ensure that variations in pronunciation and non-significant mistakes did not invalidate the test itself. This was achieved by anticipating mistakes that might be made by the children, and categorising these as significant or non-significant to the assessment outcomes. In addition, it was important to ensure that the test was entertaining and engaging for participants. Care was taken to avoid making the assessment process too lengthy or tedious, in an effort to avoid test fatigue. Also, some flexibility was allowed during assessment sessions. Thus, if a child was reluctant or unwilling to cooperate on one test element, this was revisited later.

Particular care was taken to ensure that the vocabulary items used within each task were appropriate and familiar to participants (Gillon, 2012). Unfortunately, there is no database for Saudi Arabic providing acquisition norms for words which would have formed a reliable source of age-appropriate vocabulary so expertise on this was sought from a range of practitioners including speech therapists and pre-school and elementary teachers.

The difficulty of identifying appropriate vocabulary was further compounded by the fact that there are several spoken dialects of Arabic within Saudi Arabia which are often markedly different from Modern Standard Arabic (MSA) and it was thought possible that given the demographic diversity of the Saudi capital, some participants would use non-local variants. However, although participants may speak different dialects at home, they are all educated through the medium of MSA which is also the language used in storybooks. Thus, in the absence of an adequate existing vocabulary corpus, all of the items used in the tests were taken from age-appropriate storybooks written in MSA.

The words chosen from these storybooks were short and commonly used in everyday life in both MSA and dialects to ensure that test participants were as familiar as possible with the material. A concerted attempt was made to choose words which were concrete nouns and, with few exceptions, these were names of everyday objects, animals, vegetables or fruits. In most cases, the singular form of the noun was used.

These measures were intended to cater for the possible diversity of national backgrounds and dialects in the study sample. Furthermore, the process of vocabulary selection needed to take into account the knowledge and cultural background of participants, to ensure that material was familiar and did not detract from the test's purpose by causing confusion. Moreover, cultural sensibilities needed to be taken into account, in order to ensure that participants were not exposed to anything that might be considered age-inappropriate or offensive. These criteria were applied to the selection of materials for all of the tests, with the exception of the syllable segmentation and blending tasks, where unfamiliar words needed to be used to ensure that the test was not biased by prior knowledge.

Pictures are widely used when testing RRS, particularly in cases where the tasks require non-verbal responses (PIPA, Dodd et al., 2000; PAT, Muter et al., 1997a) so it was decided wherever possible to use words that could be represented pictorially, as this would be a useful means of both engaging the child's attention and reducing memory load. Testing the hypothesis that pictures helped to reduce cognitive load in PA tasks, Gibbs (2003) found that performance improved when visual stimuli were used in tests. Alloway et al. (2004) suggested that picture-based tasks exert fewer demands on memory load and provide a stimulating and enjoyable testing process for children, as well as encouraging shy or reticent participants to speak. In addition, picture use has been demonstrated to be effective in triggering children's semantic representation of words (Stackhouse and Wells, 1997). However, images were not used in the blending task, to allow this task to be completed using a lexical route (recognized regular/irregular words) or a non-lexical route (sounding out letters to read a word).

On the other hand, pictures may prove distracting for children, diverting their attention away from the sounds that were the focus of the test, so care was taken to ensure that the images used were clear, simple and easy to understand. Images were taken from two internet sources and selected for their clarity and attractiveness, and then edited to ensure they were similar in style and that any distracting background or additional elements were removed, preventing bias towards particular images. If an appropriate image could not be found, a professional artist was asked to provide one. In addition, the ethnic, educational and socioeconomic background of the children was taken into consideration when selecting images.

To ensure these images successfully represented the required word, a pilot test was undertaken with seven Arabic-speaking children who were asked to identify each depicted object. This created further confidence that the images and the concepts they represented would be easily recognizable by test participants, and that content was age appropriate for participants. An example of a test battery image is provided in Appendix O.

3.8.4 Overview of the test battery

In order to create the test battery, a number of tasks were assembled that were relevant to the assessment of the children's RRS as outlined in **Table 3.6**.

CONSTRUCT	TASK	SOURCE
Phonological Awareness (PA)	Rhyme awareness (8 test items)	Self-developed based on Dodd et al. (2000)
	Syllable segmentation (8 test items)	
	Alliteration awareness (8 test items)	
	Phoneme isolation (8 test items)	
	Blending (20 test items)	Taibah (2006)
Letter Knowledge (LK)	Identification of name or sound of all 28 letters of the Arabic alphabet presented in random order on flashcards (28 test items)	Self-developed based on Dodd et al. (2000)
Rapid Automatic Naming (RAN)	Four subtests, designed to measure the rate at which participants were able to correctly name randomized items consisting of (1) letters; (2) objects; (3) digits and (4) colours, (30 test items)	Taibah (2006)

Table 3.6: Overview of tasks and sources

With the exception of the blending and RAN tasks, which were taken from an Arabic source (Taibah, 2006), all the other test battery tasks were devised by the researcher and modelled on Dodd et al. (2000).

3.8.5 PA assessment tasks

Participants were assessed using five tasks which evaluated five different components of PA, namely rhyme awareness, syllable segmentation, alliteration awareness, phoneme isolation, and blending.

3.8.5.1 <u>Rhyme awareness task</u>

The rhyme awareness task was a "matching" task[9] which involves choosing the one of the three alternatives provided that rhymes with or sounds like the stimulus word. It was chosen as being most suitable for the targeted age group (4;0-7;0) while the visual component helped maintain participant attention.

3.8.5.1.1 <u>Materials</u>

All test items were simple, easily recognized, monosyllabic words with closed syllables (see **Table 3.7**). A and B were used as practice examples followed by eight trials each consisting of: (1) stimulus word; (2) target (rhyming) word and (3) two distractors. One distractor was phonologically related to the stimulus word, sharing either the same vowel or initial consonant. The other was completely unrelated to the stimulus word, sharing no segment. Thus, in trial

[9] This task could also be described as a 'rhyme detection' task, as participant select from several alternatives. However, since it requires participants to actively choose the item that matches the stimulus word, it is classified here as a matching task.

number two, stimulus word /*dam(m)*/ and related distractor /*dub(b)*/ share the same initial phoneme /d/. The unrelated distractor, /*riz(z)*/, does not share any phonemes with the stimulus word. In trial number three, the stimulus word /*t̠o:g*/ and related distractor /*Θo:b*/ share the same vowel /o:/.

With respect to the nature of the relationship between the stimulus and distractor, a limited number of one-syllable words in Arabic would have been accessible to the age range being tested. Thus, only two suitable items sharing the same consonant were found, the remaining eight items shared the same vowel. One of each of these types was used for practice. The order in which target word, related and unrelated distractor was presented in each trial varied to ensure participants did not respond simply on the basis of item positioning.

3.8.5.1.2 Procedure

Firstly, the child was asked to name the images in each trial to check s/he used the right term in each case. Trials A and B were presented for practice, allowing for corrective feedback from the examiner and giving the child time to become familiar with the testing method. This ensured that children fully understood the task before engaging in the eight test trials on which they would be scored.

For each trial, the examiner placed three images on the table corresponding to the target word and the two distractors. The child was shown the image of the stimulus word and told: "*Here is a picture of a* [stimulus word] *and there are two more pictures down here"*. All instructions and explanations were given in the children's mother tongue, Arabic. Then pointing to each of these in turn, the examiner said: "*Which of these* [says name of first image, then second, then third] *rhymes with* [stimulus word]?" The examiner continued through all eight trials without providing feedback.

3.8.5.1.3 Scoring

A response was scored correctly if the child pointed to the correct image, or repeated the target word. If the child pronounced the target word incorrectly, for example, [*θam*] for /*fam*/, this was also scored as correct. Correct responses received one point; no point was awarded for an incorrect response (see Appendix G).

		Stimulus	English	Rhyming word	English	Related distractor	English	Unrelated distractor	English	Stimulus/distractor similarity
Practice Items	A	fa:rr فـار	Mouse	na:r(r) نـــار	Fire	fe:l فيل	Elephant	ke:k(k) كـيك	Cake	/f/- initial consonant
	B	lo:z لـــوز	Peanut	mo:z مــوز	Banana	θo:r ثــور	Ox	xe:t̲ خيـط	Thread	/O:/- vowel
Test Items	1	fa:s فـاس	Axe	na:s نـاس	People	fu:l فـــول	Bean	ħu:t حوت	Whale	/f/- initial consonant
	2	dam(m) دم	Blood	fam(m) فـــم	Mouth	dub(b) دب	Bear	riz(z) رز	Rice	/d/- initial consonant
	3	t̲o:g طوق	Swim ring	fo:g فوق	Up	θo:b ثـــوب	Dress	θu:m ثوم	Garlic	/o:/- vowel
	4	xat̲ (t̲) خـــط	Line	bat̲(t̲) بـــط	Ducks	dʒad(d) جـد	Grandfather	zir(r) زر	Button	/a/- vowel
	5	fi: ʃ فيش	Plug	ri: ʃ ريـــش	Feather	di:k ديـك	Rooster	ʕe:n عيـن	Eye	/i:/- vowel
	6	xad(d) خــد	Cheek	Jad(d) يـــد	Hand	bar(r) بـــر	Desert	sin(n) سن	Tooth	/a/- same vowel
	7	ra:s راس	Head	ka:s كـــاس	Glass	ʃa:l شـال	Scarf	be:t بيت	House	/a:/-same vowel
	8	xe:l خيل	Horse	ðe:l ذيل	Tail	be:d̢ بيض	Egg	s̲u:f صوف	Wool	/e:/-same vowel

Table 3.7: Rhyme awareness stimuli[10]

[10] The duplicated sound in the pronunciation of Arabic words, represented by the *shadda* diacritic marking in Arabic morphology, is indicated here by the doubling of the letter repeated within brackets.

3.8.5.2 Syllable segmentation task

The aim of the syllable segmentation task was to test whether children are able to segment the syllables of an unfamiliar or relatively unknown word, and whether this differed for stimuli with initial syllable stress (for example, /'dun–ja:/) and those where the stress falls elsewhere (for example, /mi ʤ–ra–fah/).

3.8.5.2.1 Materials

Test items chosen for this task were unfamiliar words unlikely to be present in the child's everyday vocabulary (see **Table 3.8**) in order to test segmentation skills, rather than recall of familiar items. In addition, no visual stimuli were used, thus requiring participants to draw on their own memory.

	Item No.	No. of Syllables	Items in Arabic	Items in IPA with Stress	English Gloss
Practice items	A	2	صدى	sa–**'da:** 2nd syllable	Echo
	B	2	شمـــام	**'ʃam**–ma:m 1st syllable	Melon
Test items	1	2	دنيــــا	**'dun**–ja: 1st syllable	Life/world
	2	2	ابـــهام	'**ʔb**–ha: m 1st syllable	Thumb
	3	2	دلــــو	**'da**–lu: 1st syllable	Bucket
	4	3	شماعة	ʃam–**'ma:**–ʕah 2nd syllable	Clothes rack
	5	3	مجرفة	**'mi ʤ**–ra–fah 1st syllable	Shovel
	6	3	مروحة	mir–wa-**ħah**' 3rd syllable	Electric fan
	7	4	جاذبية	ʤa:–ð–**'bi**-jah 3rd syllable	Gravitation
	8	4	مدفعية	mad–fa–**'ʕi**-jah 3rd syllable	Artillery

Table 3.8: Items in the syllable segmentation task

The task included 10 items, two for practice and eight for testing. Items ranged in length from two to four syllables and included words with stress either on the initial or non-initial syllable. This means that children have to rely on their skill of segmenting words into syllables and not on their memory in spotting the position of the stress or obtaining cues from the examiner[11] (see **Table 3.8**).

Items became progressively longer and the stress position was varied randomly within each item, regardless of its length in syllables. The examiner was careful to ensure consistency in her pronunciation of syllabic stress patterns since there is evidence that lengthening the final syllable can enhance performance in a segmentation task (Saffran et al., 1996).

3.8.5.2.2 Procedure

The method used to administer this task was developed specifically for this investigation. Two practice items, A and B, were included to allow corrective feedback from the examiner and allow familiarisation with the testing method, ensuring the child fully understood the task before beginning. The examiner explained the nature of the task as follows: "*When we say each word, we say it with a tap on the table, like this: /sa–**'da:**/"* (practice item A). The examiner tapped the table twice, once for each syllable in the word */sa–**'da:**/* and then encouraged the child to copy this by repeating the word and marking each syllable with a tap: *"Now you do it."* During the practice trial, if the child responded correctly to item A, the examiner moved to item B. If the response was incorrect, a further demonstration was given and feedback offered. After the practice items, the examiner proceeded with the remaining eight test items, without giving further feedback or correcting the child's responses. If the child did not respond to a particular item, the examiner continued with the next item, returning later to any omitted items to ensure there were no non-responses.

3.8.5.2.3 Scoring

If a child responded with the appropriate number of taps but did not vocalise the word, the answer was accepted as correct. Correct responses received one point; no point was awarded for an incorrect response (see Appendix G).

[11] Some studies have suggested a link between stress and performance in syllable segmentation tasks. Sebastián-Gallés and Costa (1997) found evidence of this occurring in Spanish. In addition, Saffran et al. (1996) found that young adults use multiple segmentation cues (lexical, syntactic, and stress-pattern) in combination to break up continuous speech. Furthermore, research suggests that segmentation cues can be used flexibly so that remaining cues are relied upon more heavily when other information is missing.

3.8.5.3 Alliteration awareness

This task was intended to explore children's ability to identify items that differed according to their phonemic properties (Torgesen and Bryant, 1994), spotting the one from the group of four presented that did not have the same initial sound. This odd one out method was used in preference to a matching task because it requires the child to segment syllables and compare between words (Passenger et al., 2000).

3.8.5.3.1 Materials

Each trial consisted of four items: three sharing the same initial consonant and the odd one out (target item) having a different initial consonant (Dodd et al., 2000). **Table 3.9** shows the three foils followed by the target item in each case. In the actual testing session, the target item was randomly placed to prevent the child from responding purely on the basis of order of presentation. Words ranged in length from one to three syllables, and became progressively longer.

Initial consonants of both target and foil items also varied in phonological proximity in terms of their place and manner of articulation, the expectation being that the greater the numbers of differences between target and foils the easier it would be for the child to distinguish between these. Thus, in practice trial A, the three foils, i.e. /*ba:b(b)*/, /*bat(t)*/ and /*be:t(t)*/ begin with the voiced bilabial stop /b/ whilst the target item /*se:f(f)*/ begins with /s/, a voiceless dental fricative. Foils and target were thus characterized by maximum distance, differing with respect to voice, place and manner. In other trials, target and foils differed in just one or two features. These differences are mapped onto **Table 3.9.** Target words are bolded in red or blue. Target words coloured red are expected to prove harder for children to than blue ones due to the higher number of differences involved.

				Manner of articulation								Place of articulation									
		Items	No. of differences	Nasal	Stop	Affricate	Fricative	Approx	Trill	Voiced	Voiceless	Labial	Inter dental	Dental	Alveolar	Post- alveolar	Palatal	Velar	Uvular	Glottal	Pharyngeal
Practice Items	A	ba:b(b)			+					+		+									
		bat̲(t̲)			+					+		+									
		be:t(t)			+					+		+									
		se:f(f)	**3**				**+**				**+**			**+**							
	B	fam(m)					+				+	+									
		fa:r(r)					+				+	+									
		fi:l(l)					+				+	+									
		luwz(z)	**3**					**+**		**+**				**+**							
Test Items	1	di:k(k)			+					+				+							
		dub(b)			+					+				+							
		dam(m)			+					+				+							
		bejd̪	**1**		**+**					**+**		**+**									
	2	naħil		+						+				+							

	Items	No. of differences	Nasal	Manner of articulation							Place of articulation									
				Stop	Affricate	Fricative	Approx	Trill	Voiced	Voiceless	Labial	Inter dental	Dental	Alveolar	Post-alveolar	Palatal	Velar	Uvular	Glottal	Pharyngeal
	naxil		+						+				+							
	nisir		+						+				+							
	dʒaˈzar	**2**			**+**				**+**				**+**			**+**				
3	masdʒid		+						+		+									
3	maktab		+						+		+									
3	mifta:ħ		+						+		+									
3	**fusˈta:n**	**2**				**+**				**+**	**+**									
4	ʕasal						+		+											+
4	ʕalam						+		+											+
4	ʕinab						+		+											+
4	**samak**	**3**				**+**				**+**			**+**							
5	ʃuwkah					+				+					+					
5	ʃura:b					+				+					+					

	Items	No. of differences	Manner of articulation								Place of articulation									
			Nasal	Stop	Affricate	Fricative	Approx	Trill	Voiced	Voiceless	Labial	Inter dental	Dental	Alveolar	Post- alveolar	Palatal	Velar	Uvular	Glottal	Pharyngeal
	ʃaʕar					+				+					+					
	ħisˈa:n	**1**				**+**				**+**										**+**
6	dʒubun			+					+							+				
	dʒamal			+					+							+				
	dʒazar			+					+							+				
	daˈradʒ	**1**		**+**					**+**				**+**							
7	madʒallah		+						+		+									
	muθallaθ		+						+		+									
	maxaddah		+						+		+									
	ˈʃadʒarah	**3**				**+**				**+**					**+**					
8	mist̪arah		+						+		+									
	manʃafah		+						+		+									
	mit̪ragah		+						+		+									

					Manner of articulation							Place of articulation									
		Items	No. of differences	Nasal	Stop	Affricate	Fricative	Approx	Trill	Voiced	Voiceless	Labial	Inter dental	Dental	Alveolar	Post-alveolar	Palatal	Velar	Uvular	Glottal	Pharyngeal
		burtuˈqa:l	1		+					+		+									

Table 3.9: Phonological proximity of alliteration awareness task items

3.8.5.3.2 Procedure

Before each trial, the child was asked to name all the pictures to ensure the expected term was used. If a difference in terms of accent or dialect used was noted, the examiner modelled the desired pronunciation or naming convention, repeating this until the child was able to reproduce this. Then the examiner instructed the participant as follows: "*Your name starts with* [utters the initial sound of the child's name]. *I know other words that start with the same sound* [Examiner says two words starting with same phoneme]. *I am going to say those words again, but this time I'll add another word that starts with a different sound. Then I want you to tell me, which of those three words does not start with the same sound*?" Initial trials A and B were used for familiarisations purposes and to provide feedback in case further explanation was needed. The examiner then proceeded to the eight trials and no further feedback was provided whilst the assessment was being administered.

3.8.5.3.3 Scoring

A response was scored as correct if the child pointed to the correct image, or repeated the target word. If the word was pronounced incorrectly, for example, if the child said /*damak*/ instead of /*samak*/, this was also scored as correct. Correct responses received one point; no point was awarded for an incorrect response (see Appendix G).

3.8.5.4 **Phoneme isolation task**

This subtest revealed participants' ability to recognise word-initial consonants, separate them and pronounce them. This particular task was specifically selected to target the older children in the sample but was completed by all age groups.

3.8.5.4.1 Materials

The test comprises two one-syllable and six two-syllable words, with most of the latter having the stress on the first syllable which allowed for investigation of the effects of stress position on phoneme isolation. The test targets a range of initial consonants: stop, fricative and approximant in two-syllable words with initial and non-initial stress; onsets in the one-syllable words are clusters comprising fricative + approximant (with fricative target) and nasal + stop (with nasal target). **Table 3.10** lists the words used for the phoneme isolation task.

		Arabic Item	Items: IPA	No. of syllables	Placement of stress	Single or cluster	Stress Feature
Practice Items	A	خـــس	xas(s)	1	1	Single	Fricative
	B	شمـــس	ʃams	1	1	Single	Fricative
Test Items	1	فـراش	fra: ʃ	1	1	Cluster	Cluster
	2	مقـص	mgas(s)	1	1	Cluster	Cluster
	3	أرنب	'ʔar nab	2	1	Single	Stop
	4	شنـــطة	'ʃant̠ah	2	1	Single	Fricative
	5	لمـبـة	'lam bah	2	1	Single	Approximant
	6	كتـــاب	ki 'tta:b	2	2	Single	Stop
	7	عصفور	ʕas 'fu:r	2	2	Single	Fricative
	8	لـوحة	'lo: ħ ah	2	1	Single	Approximant

Table 3.10: Phoneme isolation task items

3.5.5.4.2 Procedure

At the start of the test, the child was asked to name all the images to ensure that he/she knew the required term and, additionally, the child was reminded of the names of the objects in all the test trials. The examiner introduced the task to the child as follows: "*Here is a picture of* /xas(s)/ *The first sound of* /xas(s)/ *is* /x/ *or* /xa/ [makes the sound of the initial phoneme]. *Let's try some others*." For each item, the examiner said the word and then waited for the child to say this.

As in the previous tasks, two trials A and B were presented as practice, allowing corrective feedback from the examiner and giving participants time to familiarise themselves with the method of testing before engaging in scored trials.

3.5.5.4.3 Scoring

If the child's response was either the first consonant (for example /x/) or the first consonant followed by /a/ (for example /xa/), this was considered correct. When naming consonants in spoken Arabic, the consonant is followed by /a/. However, if the child offered [ki] for /*kita:b*/, instead of [k] or [ka], this was scored as incorrect, since the child has given the first syllable of the word. This may indicate a limitation within the test design as in all but two of the words used in the test the consonant was actually followed by the vowel /a/ meaning that in some cases it was not possible to know if the child had isolated the phoneme or was pronouncing the first syllable since these would have sounded the same. However, if a child made consistent and systematic substitutions for the consonant, for example fronting initial velar stops due to an articulation problem, the production of [t] for initial /k/ was accepted. Correct responses received one point; no point was awarded for incorrect responses (see Appendix G).

3.8.5.5 Blending task

This task used Taibah's (2006) Arabic adaptation of Wagner et al.'s (1999) CTOPP blending task, originally an English language test.

3.8.5.5.1 Materials

This task contained 20 items. These varied in length, with three types of segments which the child was required to blend: the largest segments are syllables, for example /*zaj'_tun*/; intermediate segments are onset plus rime, for example /*s_ur*/; and the smallest are phonemes, for example /q/, /r/, /d/. The order of presentation progressed from larger to smaller segments as follows:

- The first three items presented two syllables making a two-syllable word, for example, */fus-ta:n/.*
- The next eight items presented onset plus rime making a one-syllable word, for example */n- a:r/.*
- The remaining nine items presented a sequence of phonemes making up words that vary randomly from one to five phonemes, for example a-l-w-a:-n.

3.8.5.5.2 Procedure

The examiner demonstrated the tasks as follows: "*Listen carefully. I'll say a word in small parts, one part at a time. I want you to put these parts together to form a whole word. Are you ready*?"

Participants were trained with six practice test items, for example, two syllables forming /*ha:shim*/ [a male name]; onset+rime forming one-syllable words; first phoneme isolated, for example, /*beit*/ ‘house’; and segments forming words of one syllable, such as /*fi:*/ ‘in’. These initial trials were presented to participants with examiner feedback, before proceeding to the twenty test items without feedback on which children were scored.

3.8.5.5.3 ***Scoring***

Responses were judged to be correct if the child produced the target word. If the response was influenced by accent or incorrect pronunciation but was clearly the target word, it was accepted. If the correct lexical item was produced, but the child added the extra phoneme marking the feminine rather than masculine ending in Arabic, this was also accepted. Correct responses scored one point; incorrect responses received no mark, as did a lack of response (see Appendix G).

3.8.6 Letter knowledge task

The purpose of this task was to determine whether the child could identify the phonemes corresponding to graphemes or could provide the letter name. Both of these responses were accepted when scoring the task.

3.8.6.1 Materials

The child was required to identify the sound or name of all 28 letters of the Arabic alphabet, which were individually presented using flashcards in random order (see Appendix G). As previously noted, Arabic script is cursive and letters can take four different forms so in the test all the letters were presented in their final unconnected form (i.e. the shape used when printing the letters of the alphabet).

3.8.6.2 Procedure

Each child was shown one flashcard at a time and was asked to give the name or sound of the letter displayed. If an articulation error was committed, the examiner said a word starting with the same letter and asked the child to pronounce this in order to confirm whether he/she had a misarticulation problem.

3.8.6.3 Scoring

If the name of the letter or its corresponding sound was produced, a score of one was noted; 0 was awarded for incorrect responses. If the child made consistent and systematic errors in pronouncing a letter, this was treated as a correct response as he/she knew the grapheme but could not pronounce it (Appendix G).

3.8.7 <u>RAN task</u>

Taibah’s (2006) RAN task from his Arabic adaptation of Wagner et al.’s (1999) original CTOPP was also used as part of this battery. This task measures the speed of phonological access to lexical information. This study used four RAN subtests, designed to measure rates at which participants can correctly name randomized items consisting of letters, objects, digits or colours presented to them as flashcard images.

<u>3.8.7.1</u> ***<u>Materials</u>***

A set of 30 flashcards was produced consisting of representations of:

- Six individual Arabic letters:

ن /no:n/ ل /la:m/ ي /ja: ʔ/ م /me:m/ ص /sa:d/ ك /ka:f/

- 12 everyday objects:

يـــد	**بــاب**	**أرنب /**	**سمك**	**قـلم**	**بيت**
/jad/	/ba:b/	/ ʔ rnab/	/samaka/	/qalam/	/be:jt/
Hand	door	rabbit	fish	pen	house
ورد	**كـورة**	**شجرة**	**كرسي**	**مفتاح**	**بطـــه**
/ward/	/ku:ra/	/ʃadʒarah/	/kursi/	/miftah/	/bata/
Flower	ball	tree	chair	key	duck

- Six digits (1, 4, 5, 6, 7 and 9) and
- Six colours (red, blue, yellow, green, brown and black).

These items were then incorporated into two cards, each one presenting an array of letters, objects, digits and colours randomly arranged in rows.

<u>3.8.7.2</u> ***Procedure***

Each child was first shown the whole set of 30 flashcards described above and was asked to name all these items to ensure familiarity with them. If the child was unfamiliar with any of these items, the examiner did not proceed with the test since this would not provide a valid measure of speed of retrieval of familiar words. The actual test required the participant to name all the items presented randomly on both tests cards, one after another, as quickly as possible.

The examiner placed the first test card in front of the child, concealed with paper, and the child was instructed in Arabic as follows: "*I want you to name these items* [images of letters/objects/digits/colours] *as fast as possible starting from the top row at the right* [examiner points], *as soon I remove the paper. Even if you make a mistake, just keep on going. Get ready: go!*"

As the child completed each row, the examiner pointed to the next row to guide the child. When the first test card had been completed, the examiner produced the second card and repeated the process.

3.8.7.3 Scoring

Correct responses received one point; no point was awarded for incorrect responses. In addition, the time taken to name all items on the test card was noted in seconds. A stopwatch was used to measure the time taken from naming the first item on the test card to the last. Any misnamed or omitted test items, or hesitations lasting more than two seconds were counted as errors. In Taibah's (2006) task, any child making more than four errors on the first test card did not proceed to the second. However, in this study, the researcher administered the second test card regardless of the number of errors on the first. The final score for this task was a combination of the mark for correct responses and the average time taken to complete each test card (Appendix G).

3.9 Pilot study: RRS test battery

Extensive pilot testing was carried out to ensure the validity and reliability of the RRS assessment results. This was particularly important in this study because the initial version of the test battery was entirely new and had not been subjected to prior testing. Following the piloting process, various adjustments and modifications were made to both test content and testing procedures, resulting in a higher degree of confidence for the main testing process.

The pilot study tested 34 pre-schoolers who met the inclusion criteria. **Table 3.11** presents the characteristics of the pilot study sample. Individual testing took place in February 2010 at schools in Riyadh.

Group Number	Type of school	Age group and gender					
		4;0-4;11		5;0-5;11		6;0-7;0	
1	Public	M = 1	F=1	M = 4	F= 4	M = 8	F=2
2	Private	M = 4	F=0	M = 0	F= 2	M =6	F= 2
TOTAL		6		10		18	

Table 3.11: Pilot study sample characteristics

The pilot study was generally successful. The children responded well to the tasks, seemed at ease with the testing environment created, and were not confused by instructions or explanations. The pilot also showed that the structure, order and organisation of the tasks was viable; the majority of the amendments were made to the number of items per task and stimuli used.

Originally, there were 10 items for the rhyme awareness task, 13 for alliteration awareness, and 15 for both syllable segmentation and phoneme isolation tasks. This was standardized to eight items per task (plus two practice items) to make the test less time-consuming and tiring for young participants, and to facilitate statistical analysis.

Some words or pictures originally selected were found to be problematic and were eliminated as part of the reduction of task items. In the alliteration awareness task, most children referred to the image of the penguin as simply ‘bird’, not having acquired this more specialized vocabulary item, possibly for cultural reasons, and it was eliminated. Two items, namely /ɣaza:l/ ‘deer’ and /dura:q/ ‘peach’, were removed from the phoneme isolation tasks because they proved difficult to represent unambiguously pictorially. In the rhyme awareness task, one whole trial was eliminated because most children could not name /qo:s/ ‘rainbow’, /ti:n/ ‘fig’, and /raf/ ‘shelf’ correctly. Finally, three of the original words selected (/riz(z)/ ‘rice’, /dʒad/ ‘grandfather’ and /fa:r/ ‘mouse’) were swapped between trials after it was noted that the distractors originally chosen contained two items rather than one which rhymed with the stimulus provided. These changes were vindicated in the full study, as testing demonstrated that the tasks and pictures used were appropriate for young children. Apart from item changes, careful attention was paid to randomisation when positioning correct response to prevent response bias in all trials. Less extensive piloting of the blending and RAN tasks was needed as both had been used previously in Taibah's (2006) study.

Testing procedures were also finalized during the pilot study. Children were tested individually in a quiet room provided by their respective schools. Researcher and child sat side by side throughout the testing session, meaning stimulus material was clearly visible to both. The test battery was completed within one session lasting 35 to 45 minutes for each child. Parents did not attend these sessions and if teachers were present during the testing, they were asked not to interfere or assist the child in any way.

As previously noted, as the Kingdom's capital city, Riyadh attracts incomers from all over Saudi Arabia and beyond, leading to a linguistically diverse setting which is reflected in the primary schools. Thus, it was decided that if a child produced variants in terms of vocabulary or pronunciation, the researcher would provide the preferred word and pronunciation to maintain uniformity of stimulus items for all participants. Children were also asked to repeat words in order to ensure retention of the preferred version. This was important because some tasks depended on accessing the given phonology and pronunciation. In addition, extra care was taken during the presentation of all practice and test items to maintain the same pitch, intonation pattern and volume and to avoid placing undue stress on or lengthening the differentiating sounds.

Participant responses were noted on a score sheet (see Appendix G) and every session was recorded using a high-fidelity, portable, digital voice recorder (VN-5500PC, Olympus) to ensure responses were available for later reference. Participants did not see the score sheet and results remained confidential. As usual when conducting research with young children, responses were praised regardless of their accuracy, and after several responses, a reward (in the form of a small toy or sticker) was given to participants as encouragement to continue. If a child did not give the correct response, no clues were given.

In order to check the reliability of the task scoring, an independent speech therapist was provided with test recordings, a scoring sheet (see Appendix G) and the criteria established by the researcher and was asked to independently score 20 of the recorded sessions. This allowed the researcher to check for inter-rater agreement (Bowling, 2009). When compared, the scoring profiles of the researcher and independent second marker proved almost identical, thus validating the reliability of the assessment procedure.

3.10 Questionnaire

3.10.1 Advantages and disadvantages of the questionnaire method

For Study One, the primary data relating to RRS was collected using the test battery, as described above. However, following criteria specified by Gillon (2012), supplementary information relating to each participant's age, gender and SES was also gathered, using a parental questionnaire. These data were also used when testing the hypotheses relating to factors that could potentially affect RRS development and, in turn, reading ability. This section describes the process of creating and administering the questionnaire used in this study.

Data collection via questionnaire has distinct advantages. In particular, when dealing with large sample sizes, surveys are relatively quick to administer and collect, providing an inexpensive and efficient method of gathering data relating to multiple research questions (Walonick, 2004). This data collection format is also familiar to most people, meaning that participants do not generally require lengthy procedural explanations. Furthermore, when closed questions are used, responses can be easily entered into software packages such as SPSS, thereby facilitating the analysis process.

Despite its advantages, there are also drawbacks to this method of data collection. There is a high risk of low response rates to questionnaires, although this can be avoided by giving careful thought to the questionnaire design process and how the questionnaire is administered. Walonick (2004) notes that survey response rate can range from 10% to 90% depending on the type used, illustrating the key role that design can play in determining the success of this research method. In this study, the questionnaire was piloted to ensure the optimal design (Appendix E provides the English translation of the original Arabic pilot) and was subsequently rewarded with a high response rate.

3.10.2 Considerations in developing questionnaire structure and content

The process of questionnaire development and validation followed the criteria developed by Walonick (2004) and followed the seven-step guidelines recommended by Churchill and Iacobucci (2002) as outlined in **Figure 3.2**. Parents of study participants were asked to complete a questionnaire designed to solicit information to build up a detailed picture of the child's socioeconomic and literacy environment, and could be correlated with their performance in the RRS tasks. The final version of the questionnaire consisted of two sections. Section A asked respondents to provide the child's name, gender, date of birth, contact telephone number, location of family home, type of school attended (private or public) and school district (North, South, East or West).

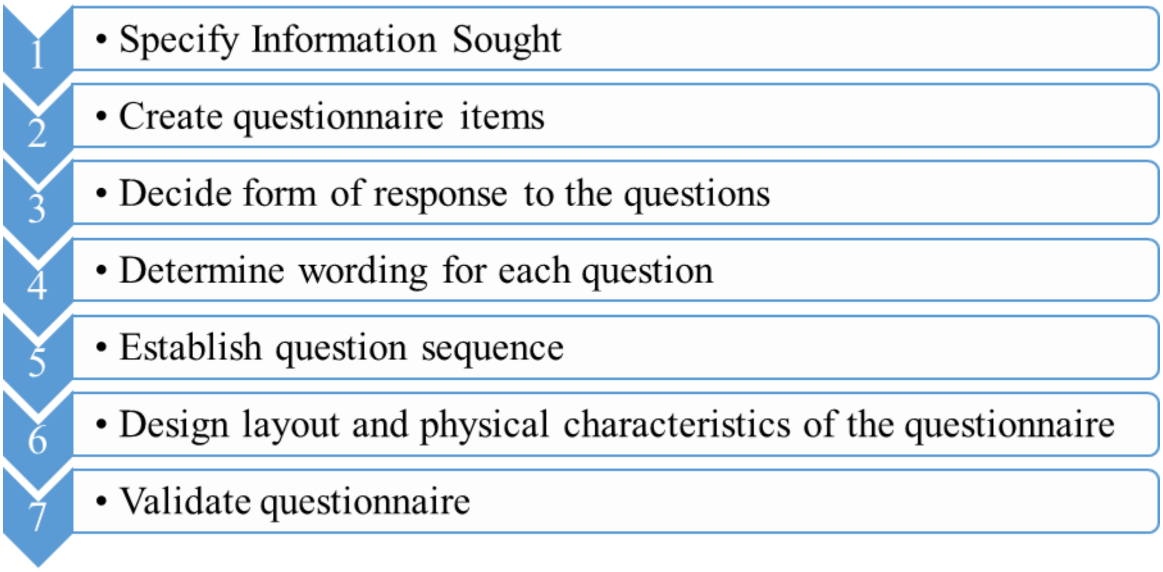


Figure 3.2: Questionnaire design process. *(Source: based on Churchill and Iacobucci, 2002).*

Section B consisted of 18 multiple-choice questions, with some also asking for further details. Items 1-9 related to the child's language background, parental level of education and occupation, monthly family income, and information about the family home. The remaining items covered aspects of the child's home environment including contact with a nanny, access to books, television and computers, and learning support. In addition, a final question was added to ascertain if the child received extra tuition in Qur'anic recitation so that this area could be explored as a supplementary strand to Study One.

Survey questions were closed rather than open ended as this quantitative approach enabled results to be transferred easily to an SPSS package that would subsequently perform the statistical analysis. This familiar format also made completing the survey much easier and less time-consuming for participants.

Clear and concise instructions were provided to respondents together with an accompanying explanation regarding why the information was being collected and its importance within the study. The number of questions was kept to a minimum and particular attention was paid to wording to ensure this was clear, unambiguous and as free of jargon and technical terms as possible. Care was taken to formulate simple questions, using neutral language throughout and avoiding leading questions. As recommended by Churchill and Iacobucci (2002) and Walonick (2004) the wording and content of the questionnaire was developed in a pilot-testing phase (see **section 3.10.4**) to remove any potentially confusing, ambiguous or redundant items and ensure options provided were appropriate for the target group.

3.10.3 Procedure

The questionnaire was administered to parents by the researcher herself, either by phone or in person. Although this proved time-consuming it had a number of advantages. It considerably increased the response rate in comparison to that of the pilot-testing phase when questionnaires were mailed to parents. It also improved the quality of responses, since pilot study returns were often incomplete and/or incorrectly filled in. Moreover, it considerably shortened the overall time required for data collection as it eliminated the time needed to allow participants to complete and return questionnaires. The telephone survey method also allowed the researcher to put respondents at ease, answering any queries and reminding them about the importance of the research and reassuring them about confidentiality. Parents were re-informed about why the information was needed, how it would be used and stored, and how it related to the study. It also facilitated explanation of any items which were not understood.

3.10.4 Questionnaire pilot

The questionnaire was extensively pilot tested before administering it to parents. Informal feedback on early versions was solicited from friends and colleagues to pinpoint any linguistic issues relating to wording, terminology or style, and to provide comments on its general user-friendliness. As a result, some language was simplified as it was thought too specialist in nature, the order of the questions was adjusted and the format simplified.

Questions were grouped more logically, with language-related items placed together at the start with SES-related items following these and arranged thematically: parental educational level, family income, accommodation and child's educational environment at home. The format was also adjusted to ensure consistency in presentation of multiple-choice answers and make questions and responses easier to read.

More formal consultation regarding the questionnaire took place with two sociologists, one from the Department of Sociology, King Saud University, Riyadh, and the other from the Social Research Unit at the Saudi Ministry of Labour. Feedback was sought on questionnaire format and content. These specialists recommended additions to the categories covering educational qualifications and accommodation. They also pointed out that obtaining SES information from parents could be difficult because they might not be aware of the relevance of research. In addition, as previously noted, Saudis can be sensitive about providing personal information, particularly relating to income. Specialists also cautioned against using email to distribute questionnaires as levels of computer literacy and internet access might be limited among the sample population.

Following further revisions, a notice was circulated to all schools asking parents for feedback on the questionnaire content and language. In addition, a group of mothers from diverse socioeconomic classes was used to provide feedback on the appropriateness of questions and response options. These responded to an open invitation posted in the Administration Department, King Saud University. Further small changes were implemented to the accommodation items.

The original breakdown of income ranges was based on one used by a well-known Saudi mobile phone company as part of their marketing research materials.[12] During the pilot stage, this was found to be insufficiently discriminating and consequently the number of categories for this answer was increased, allowing for a more precise categorisation of middle-income respondents. However, at the analysis stage, it proved necessary to once again use broader ranges to facilitate the analysis of results.

At the piloting stage, 34 questionnaires were distributed by post to parents. The response rate was relatively low (20 completed questionnaires returned), which led to the decision to administer the questionnaire by telephone or in person. The questionnaire underwent considerable changes from initial draft, pilot phase (Appendix E) and final version (Appendix F). The number of items was substantially reduced (from 25 to 18) since some either duplicated information already known from other sources or were not of direct relevance to the study's aims. One leading question was identified and eliminated as it was thought it might have influenced respondents to give the "right answer" to show they were "good" parents. Question order was also changed again to improve logical flow by grouping related questions. Finally, detailed information pertaining to research participants was removed as it was realized that this data had already been collected as part of the sampling process.

Given the disappointingly low return for the pilot study questionnaire, largely due to a general lack of awareness concerning academic research, time and effort was invested in maximizing the response rate for the main study. The information about the study and a paper copy of the questionnaire was originally distributed by the school manager so that parents would have some background information. The researcher then personally contacted every family involved by phone and went through the questionnaire with one of the child's parents. This meant that any unclear points could be instantly clarified, reassurance could be given about confidentiality and further explanation provided concerning the nature of the study and its purpose. The researcher

[12] At the time of writing, Mobily are considered the second biggest mobile company in Saudi Arabia. Their Market Research Department was happy to provide this information by e-mail when the researcher contacted them.

used this opportunity to explain to parents the importance of having data of this kind to develop diagnostic tests that would benefit Saudi children. Although this was very time consuming, it produced an excellent return rate.

3.11 Study One: Culture-Specific Component

3.11.1 Tuition in Qur'anic recitation

As discussed in the literature review, findings from previous studies (Rosowsky, 2001; Robertson, 2002; Burgoyne et al., 2009) have suggested that the study of Qur'anic recitation which forms a part of regular religious practice in Islamic societies from a very early age might have a generally positive effect on the reading ability of Arabic-speaking children. However, no studies have specifically examined the impact of Qur'anic recitation on development of RRS.

Since children in Saudi Arabia are taught to recite the Qur'an from memory and to decode (but not read) extracts of this text from a very young age, the decision was taken to investigate whether this culturally specific factor might play a role in children's development of RRS. It is possible to compare performances in the RRS tasks by children with different levels of exposure to tuition in Qur'anic recitation since this subject forms part of the core curriculum but some children also take extra classes. The final item on the questionnaire made it possible to identify a sample of those children who received additional lessons in Qur'anic study. This section describes the research methods and procedures employed in this culture-specific component of Study One which addresses whether levels of exposure to tuition in Qur'anic recitation affect children's performance in all the measures of RRS.

3.11.2 Consent

As with the RRS component of Study One, before beginning this culture-specific investigation, any further informed consent needed was sought as necessary following the process outlined in **section 3.5**.

3.11.3 Sampling method

Questionnaire returns from the Study One sample described in **section 3.10** were used to identify children who received additional tuition in Qur'anic recitation. A further check was made to verify that this was the case and a non-random quota sampling method was then used to select participants for the culture-specific component of Study One by type of school (public or private), gender and age. This produced a sample of some 55 children in total.

3.11.4 Participants

The key characteristics of the participant sample for the culture-specific component of Study One are summarized in **Table 3.12**. It was not possible to create a wholly representative sample but there was a spread across public (*n*=33) and private (*n*=22) schools and a relatively equal spread of participants across all three age groups within each school sector. Although both male and females are represented, boys predominate in the sample.

ARTICIPANTS EXPOSED TO EXTRA QUR'ANIC RECITATION TUITION											
Public schools (*n*=33)						Private schools (*n*=22)					
4;0-4;11		5;0-5;11		6;0-7;0		4;0-4;11		5;0-5;11		6;0-7;0	
M	F	M	F	M	F	M	F	M	F	M	F
8	4	7	6	5	3	3	3	4	4	4	4
12		13		8		6		8		8	
TOTAL = 55											

Table 3.12: Age, gender and school type for participants exposed to extra Qur'anic recitation tuition

Once the 55 participants from the main sample who had extra Qur'anic recitation tuition had been identified, their existing scores for the RRS tasks were used in calculations which compared their performance with that of the other children in the sample who had received less exposure to this culture-specific element of Saudi education.

3.12 Data Coding and Analysis

All data collected from test battery RRS tasks and SES questionnaires were anonymized and each participant was randomly allocated a number, which did not contain any information which could be used to identify an individual. These anonymized data were input into an SPSS (v.22) package which was used to generate all statistical data.

Once the data had been collated, appropriate statistical tests were chosen to conduct the necessary descriptive and inferential analyses. Means and standard deviations for the RRS task results were calculated and correlated with relevant information derived from the questionnaire. Analysis of Variance was used to compare test scores between groups. Normality tests were also performed using Kolmogorov-Smirnov and Shapiro-Wilk tests. To supplement this, and ensure that the ANOVA test was reliable, additional tests such as the Kruskal-Wallis and Mann-Whitney U tests were run to investigate the results of particular variables. To investigate correlation between performance in the test battery and other variables, the Pearson Correlation and Non-Parametric Spearman's Rho were used. Regression analysis was also conducted to ascertain the accuracy of the RRS test battery in predicting levels of competence in reading-related abilities. The results of these analyses are presented below (**section 3.14**) and discussed in Chapter Five.

3.13 Ethical Considerations

This research needed to take into account a wide variety of ethical considerations. Since the study involved very young participants, it was imperative to ensure that a high degree of transparency in research procedures and aims of the investigation was evident at all times. Proposals for all components of Study One were approved by the School of Community and Health Science Research Ethics Committee, City University London. Once this clearance had been received, a process of informed consent (see **section 3.5**) was undertaken for the study and all those identified as key stakeholders (school officials, parents and children) were given appropriate information before consenting to participate, ensuring that they understood the aims of the study and the implications of their involvement. If at any point participants did not feel comfortable with what was being asked of them, they were given the opportunity to voice their concerns. If these could not be resolved satisfactorily, they were free to end their involvement in the investigation.

Parents were assured that all personal data would be appropriately protected and informed about how data would be stored. All possible measures were taken to ensure information supplied by questionnaire respondents or test participants was anonymized using a coding system to link children's test results with parental questionnaire responses. Paper-based information was stored securely in a locked room at all times and all digital information was kept in password protected files. Only information relating directly to the aims of the investigation was collected and stored, and any additional personal information was disposed of safely.

In addition, considerable efforts were made to ensure that parents and children felt comfortable with the testing procedure at all times. The test was designed to be fun for participants and they were praised and children were rewarded for their involvement. Particular attention was paid to children's emotional state during the test and if they were seen to be uncomfortable, the test was paused and the child reassured. The *Guidelines for Research with Children and Young People* devised by Shaw et al. (2011) offered helpful advice to best practice in this area. During the telephone questionnaires with parents, the researcher offered reassurance if respondents sounded anxious or confused.

3.14 Results of Statistical Analysis

The results of Study One are presented in the following sections. Both descriptive and inferential statistics are presented here in a variety of forms. When indicated, further tables and figures can be found in the appendices.

3.14.1 Participant characteristics

Since the aim of the study was to create and pilot a test battery to assess the RRS of PA, LK and RAN, the research was interested in analysing an initial distribution of scores on these key measures for the sample of participants. This distribution was based on the performance of typically developing Arabic-speaking Saudi children aged 4;0-7;0. The initial sample consisted solely of typically developing children since for educators, clinicians and professionals who provide assessment of and intervention in RRS, an understanding of typical development is of crucial importance to understanding disordered development. It is possible that some children with reading difficulties were inadvertently included in the two studies but none of those included had been formally diagnosed as such. All participants were included in the regression analyses to ascertain the diagnostic power of the RRS assessment battery.

The number of schools selected from each area of Riyadh was distributed as evenly as possible in terms of sector (public/private), as shown in **Table 3.13**.

RIYADH AREA	SECTOR	
	Public	Private
North	5	4
South	4	4
East	4	4
West	4	4
Total	**17**	**16**

Table 3.13: School distribution in sample in terms of area and sector

The following **tables 3.14** and **3.15** present the independent variables used to analyse participant performance using the RRS test battery.

A total of 384 participants were tested: 192 (50%) attend public schools and 192 (50%) private schools. In terms of gender, equal numbers of male and female students were tested in each educational sector. By age group, participants were divided as follows: 123 (32%) in the youngest group (kindergarten stage 4;0-4;11); 128 (33%) in the intermediate group (kindergarten stage 5;0-5;11) and 133 (35%) in the oldest group (6;0-7;0). Children in this group attended first grade in elementary school (see **Table 3.14**).

<table>
<tr><th colspan="4">PARTICIPANTS BY SECTOR AND AREA</th><th colspan="2">GENDER</th><th colspan="3">AGE</th></tr>
<tr><th>N</th><th>S</th><th>E</th><th>W</th><th>Male</th><th>Female</th><th>4;0-4;11</th><th>5;0-5;11</th><th>6;0-7;0</th></tr>
<tr><td colspan="4">PUBLIC SECTOR</td><td rowspan="2">96</td><td rowspan="2">96</td><td rowspan="4">123</td><td rowspan="4">128</td><td rowspan="4">133</td></tr>
<tr><td>48</td><td>48</td><td>48</td><td>48</td></tr>
<tr><td colspan="4">PRIVATE SECTOR</td><td rowspan="2">96</td><td rowspan="2">96</td></tr>
<tr><td>48</td><td>48</td><td>48</td><td>48</td></tr>
<tr><td colspan="4">TOTAL = 384</td><td colspan="2">TOTAL = 384</td><td colspan="3">TOTAL = 384</td></tr>
</table>

Table 3.14: Number of participants by area, sector, gender and age

PUBLIC SECTOR						PRIVATE SECTOR					
4;0-4;11		5;0-5;11		6;0-7;0		4;0-4;11		5;0-5;11		6;0-7;0	
M	F	M	F	M	F	M	F	M	F	M	F
30	31	32	32	34	33	30	32	32	32	34	32
61		64		67		62		64		66	
Total = 192						Total = 192					

Table 3.15: Number of participants by sector broken down according to age and gender

Although many previous studies used six-month age bands, initial analysis suggested these were not sufficiently discriminating (see Appendix H), so for the main analyses the age bands were collapsed and broader one-year bands adopted.

3.14.2 Gender and RRS

Initially, a descriptive analysis was carried out on the girls' and boys' scores for measures of PA, LK and RAN to calculate the means and standard deviations. The independent t-tests for gender differences on each task, by age group, were not significant. Therefore, gender differences on the PA, LK and RAN tasks were investigated for the sample as a whole rather than by age group. **Table 3.16** shows the mean/standard deviation (SD) performance in the RRS test battery, divided according to gender.

Measures	Gender	Mean	SD	df (*n*=384)	Sig.(2-tailed) *p* value
PA: Rhyme Awareness	Male	5.68	1.51	382	.50
	Female	5.78	1.52		
PA: Syllable Segmentation	Male	5.35	1.60	382	.60
	Female	5.45	1.71		
PA: Alliteration Awareness	Male	3.20	1.59	382	.07
	Female	3.52	1.77		
PA: Phoneme Isolation	Male	2.95	2.73	382	**.03***
	Female	3.55	2.68		
PA: Blending	Male	6.59	2.68	382	.33
	Female	6.86	2.78		
Letter Knowledge	Male	9.21	9.83	382	.08
	Female	11.03	10.20		
RAN: Colour Task valid average time[1]	Male	40.76	6.75	114	.39
	Female	39.65	6.97		
RAN: Object Task valid average time[1]	Male	47.23	6.81	112	**<.01****
	Female	43.90	6.62		
RAN: Letter Task	Male	38.07	7.38	61	.09

valid average time[1]	Female	34.83	7.40		
RAN: Digit Task valid average time[1]	Male	39.82	8.05	71	.17
	Female	37.33	6.80		

Table 3.16: Participant performance on RRS tasks according to gender

[1] Valid average time to perform tasks is shown in seconds
** t-value was significant at 0.05 level (p < 0.05) ** t-value was significant at 0.001 level (p <0.01)*

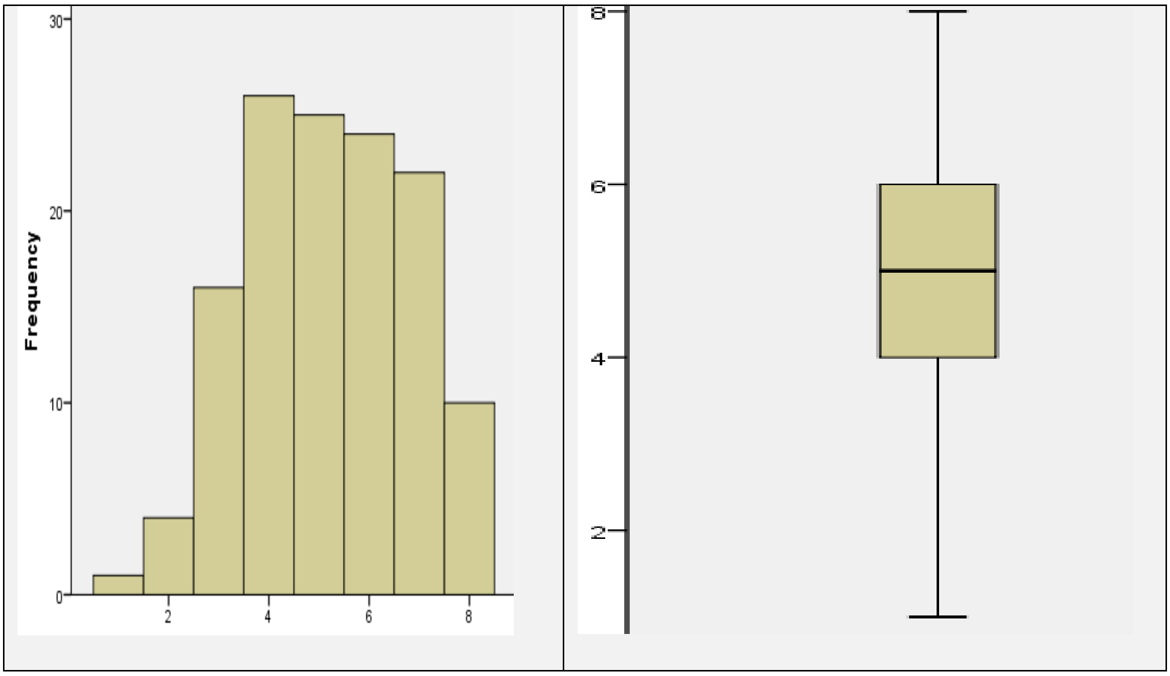
The only tasks that appeared to differentiate between males and females at a statistically significant level were PA: Phoneme Isolation ($p < 0.05$) and RAN: Object Task ($p \leq 0.01$) with mean scores showing that girls performed better on PA: Phoneme Isolation and achieved faster times than boys on the RAN: Object Task (see **Table 3.16**). However, neither of these results proved to be significant once a Bonferroni correction for multiple comparisons was applied as the significance value was reduced to .005.

Since gender did not appear to affect participant performance in any of the RRS tasks, the decision was made not to control for gender in any further multivariate analysis.

3.14.3 Age and RRS

3.14.3.1 Normality testing

Prior to running statistical analyses to assess the age sensitivity of the RRS measures, the data was assessed for normality according to age groups, since many statistical tests and procedures, such as parametric statistical tests, assume that data follows a normal distribution. This normality testing was carried out using both graphical and numerical means to help understand the distribution within the sample. Appendix I shows all the histograms produced for normality testing and an illustrative example of these is shown in **Figure 3.3**.



A. Histogram *B. Box plot with whiskers*

Figure 3.3: Representative graphical method for normality testing: Rhyme Awareness/Age Group 4;0-4;11

The means and medians for all the data sets were calculated and used to produce box plots for all the RRS measures. On a box plot (Figure 3.3 B) the length of the box is the interquartile range of the variable, containing 50% of the cases, the horizontal band inside the box representing the median value. The whiskers extending above and below the box represent the data range. Box plots were produced to obtain a graphical display of any outliers that had skewed the data. These usually appear in the box plot as tiny circles labelled with the ID number of the case. To be considered an ordinary outlier, a score should extend more than 1.5 box-lengths from the edge of the box (Pallant, 2010; Field, 2009). As **Figure 3.3** indicates, there are no ordinary or extreme outliers in the scores obtained for the Rhyming Awareness task by children in the 4;0-4;11 age group.

The histogram (**Figure 3.3 A**) shows that most participants scored between 4 and 7 (n=8) whilst the box plot shows that 50% of this age group obtained scores ranging from 4 to 6, the lowest score being 1 and the highest 8. This variation was confirmed by the numerical method. **Table 3.17** shows that the mean performance in the Rhyme Awareness task was 5.19 for the youngest group (SD 1.64).

AGE GROUP: 4;0-4;11	STATISTICS	STD. ERROR
PA: RHYME AWARENESS TASK		
Mean	5.19	
95% Confidence Interval for Mean		.145
Lower Bound	4.90	
Upper Bound	5.47	
5% Trimmed Mean	5.20	
Median	5.00	
Variance	2.673	
Std. Deviation	1.635	
Minimum	1	
Maximum	8	
Range	7	
Interquartile Range	3	
Skewness	-.088	.214
Kurtosis	-.748	.425

Table 3.17: Numerical method for normality testing: Rhyme Awareness/Age Group 4;0-4;11

Secondly, the skewness and kurtosis values were calculated for all PA and LK tasks according to age (Appendix J) to determine how the data are distributed and, in particular, if they are highly skewed. If this is found to be the case, median values are quoted because they give more information about the centre and spread of the data.

Skewness values must be taken into consideration when choosing parametric or non-parametric tests. Those tests with skewness values in the ±1.5 range were analysed using a parametric test, as in the case of Rhyme Awareness/age group 4;0-4;11. However, in the case of the LK task for the youngest age group, the skewness values were outside the ±1.5 range. Therefore, these data needed to be analysed using non-parametric tests. Skewness values were looked at more closely for all tasks (Appendix J).

As the distributions of the scores on the Histogram show (Appendix I), very few, if any, of these scores were normally distributed (although the distribution shown in **Figure 3.3** is not far off normality) and so it was decided to run extra statistical tests to test for normality. Both Kolmogorov-Smirnov[13] and Shapiro-Wilk tests were run and the results are shown in Appendix K. Kolmogorov-

[13] For a sample of this size, the Kolmogorov-Smirnov test was sufficient.

Smirnov and Shapiro-Wilk tests produced significance *p*-values of less than 0.05 in all measures for all age groups, indicating a non-Gaussian curve, meaning that on the PA and LK measures for the RRS test battery the data was not normally distributed.

Parametric tests require that assumptions regarding normal distribution are met. When data are distributed in a non-Gaussian manner, non-parametric tests are typically selected. However, sample size is of crucial importance. Parametric tests including both the *t*-test and the *F*-test (analysis of variance or ANOVA)[14] are still considered to be robust even if the population was non-Gaussian, as long as the samples are large enough due to the central limit theorem (Motulsky, 1995). Therefore, the decision was taken to run parametric analyses but to check the results of these analyses using their non-parametric equivalents, namely, Mann-Whitney U (for t-test comparisons) and Kruskal-Wallis (for ANOVA comparisons) tests, when investigating the effects of age on the different tasks (see Appendix L).

3.14.3.2 PA tasks: descriptive analysis

Firstly, descriptive statistics (mean, standard deviation, median, minimum and maximum scores) together with their associated box plots were extracted and produced for all five PA measures in the test battery (Rhyme Awareness, Syllable Segmentation, Alliteration Awareness, Phoneme Isolation and Blending) and the LK task relating to each of the six groups (three public sector, three private sector) using six-month age bands. This division is commonly used in RRS research on the basis that young children's language development changes rapidly. One-way ANOVAs were performed, taking age as a between-group factor. The results of this initial analysis are shown in Appendix H.

Overall, *F* values for PA and LK were found to be significant (large *F* values with $p < 0.001$). However, post-hoc univariate ANOVA analysis with a Bonferroni correction for multiple comparisons then showed no significant differences between immediately adjacent age bands for any measure, so the decision was

[14] One-way Analysis of Variance (ANOVA) is based on comparing the variance between the data in each group and was used in this study as a data analysis method to help to determine significant differences between the means of two or more groups. ANOVA can also be used to compare similarities or differences for three or more means. This is the same as using a t-test to compare the means of groups (Pallant, 2010; Field, 2009).

taken to collapse the age bands and adopt broader one-year bands for all the main analyses.

Results show that the intermediate and oldest groups performed better on the Rhyme Awareness task: 50% in both these groups had scores between 5 and 7. The youngest group not only performed less well but also exhibited the largest range of scores, with results spread across all values. The 5;0-5;11 and 6;0-7;0 groups produced comparable results with a minimum score of 1 for the oldest group and 2 for the intermediate group (**Table 3.18**).

Age Group	Mean	SD	Med	Min	Max (*n*=8)
4;0-4;11 (*n*=123)	5.17	1.65	5	1	8
5;0-5;11 (*n*=128)	5.93	1.49	6	2	8
6;0-7;0 (*n*=133)	6.05	1.26	6	1	8

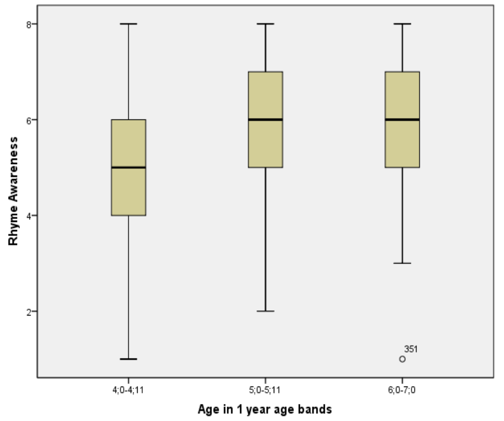


Table 3.18: Descriptive statistics and box plots: Rhyme Awareness
SD = Standard Deviation Med = Median Min = Minimum Max = Maximum

The scores for the Syllable Segmentation task improved with age. For the 4;0-4;11 and 6;0-7;0 age groups, 50% of scores were between 5 and 6, and 5 and 7, respectively, with standard deviations of 1.6 and 1.69, respectively. There were two ordinary outlier scores of 0 for the 4;0-4;11 age group, and 10 outlier values (including two extreme outliers) for the 5;0-5;11 group (**Table 3.19**).

Age Group	Mean	SD	Med	Min	Max (*n*=8)
4;0-4;11 *n*=123	4.80	1.61	5	0	8
5;0-5;11 *n*=128	5.44	1.48	5	1	8
6;0-7;0 *n*=133	5.92	1.69	6	1	8

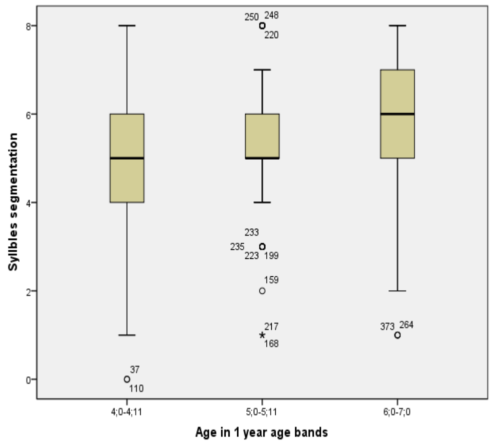


Table 3.19: Descriptive statistics and box plots: Syllable Segmentation
SD = Standard Deviation Med = Median Min = Minimum Max = Maximum

For Alliteration Awareness, mean differences indicate improving scores for the intermediate and oldest groups. The box plot for the youngest group shows 4 outlier values in the upper range, with one of those achieving the highest score of 6, and two outliers below the interquartile range. In this age group, 75% of children scored 3 or less, whereas the median for the oldest group was 4, with 50% of participants scoring 4 or above. The box plot distribution and standard deviation values also indicated that the performance spread increased with age

Age Group	Mean	SD	Med	Min	Max (*n*=8)
4;0-4;11 *n*=123	2.70	1.15	3	0	6
5;0-5;11 *n*=128	3.31	1.64	3	0	8
6;0-7;0 *n*= 133	4.02	1.91	4	0	8

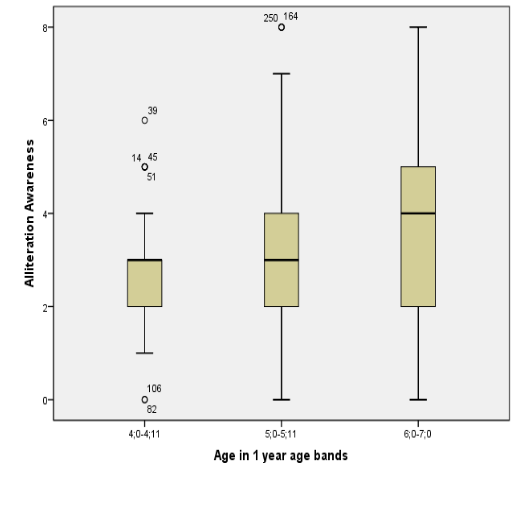


Table 3.20: Descriptive statistics and box plots: Alliteration Awareness
SD = Standard Deviation Med = Median Min = Minimum Max = Maximum

For Phoneme Isolation, mean and median both increase as the age of the children rises. Only 25% of children in the 4;0-4;11 age group achieved 3 or more, whereas 50% and 75%, respectively, of children in the two older age categories scored at least 3. The 5;0-5;11 age group showed the widest spread in scores, as reflected in the widest box plot and largest standard deviation.

Age Group	Mean	SD	Med	Min	Max (*n*=8)
4;0-4;11 *n*=123	1.74	1.89	1	0	7
5;0-5;11 *n*=128	3.18	2.84	3	0	8
6;0-7;0 *n*=133	4.71	2.47	5	0	8

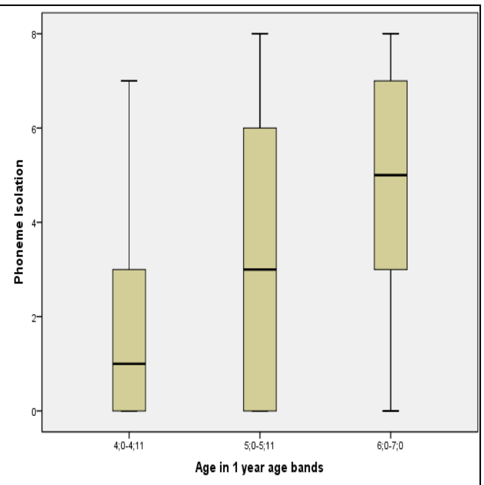


Table 3.21: Descriptive statistics and box plots: Phoneme Isolation
SD = Standard Deviation Med = Median Min = Minimum Max = Maximum

The differences in the means and medians for Blending showed that the overall results for the intermediate and oldest groups were higher, with the highest score of 16 being achieved in the 6;0-7;0 age group. For the 4;0-4;11, 5;0-5;11 and 6;0-7;0 age groups, 50% of scores were between 4 and 7, 5 and 8, and 6 and 10, respectively. There were three outliers in the upper range for the youngest age group, with one of those achieving a score of 14. None of the participants achieved the maximum score of 20.

Age Group	Mean	SD	Med	Min	Max (*n*=20)
4;0-4;11 *n*=123	5.71	2.29	5	1	14
5;0-5;11 *n*=128	6.23	2.33	6	0	11
6;0-7;0 *n*= 133	8.16	2.88	8	0	16

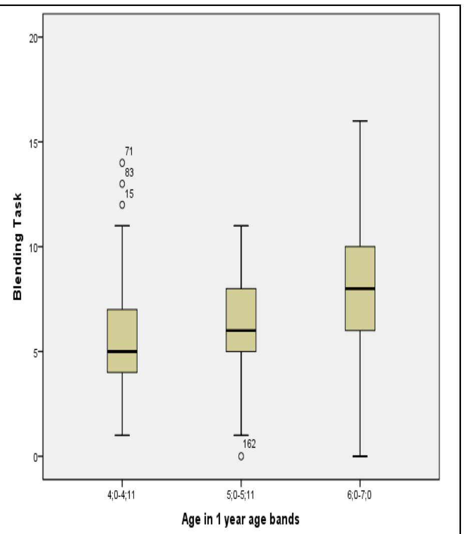


Table 3.22: Descriptive statistics and box plots: Blending task
SD = Standard Deviation Med = Median Min = Minimum Max = Maximum

3.14.3.2.1 PA tasks: inferential statistical analysis

A one-way ANOVA was carried out to investigate the effect of age on children's performance in the PA tasks, using one-year age bands (4;0-4;11, 5;0-5;11 and 6;0-7;0). Post-hoc comparisons across the three age groups were carried out, applying a post-hoc Bonferroni correction to investigate whether means differed significantly from each other.

Effect size (*d)* measures the extent to which two means differ, in terms of standard deviations. If there is a large overlap between two groups, the effect size will be relatively small and vice versa. Inter-group effect size was calculated for the three age groups (4;0-4;11 vs. 5;0-5;11, 4;0-4;11 vs. 6;0-7;0, and 5;0-5;11 vs. 6;0-7;0), using the formula:

$$\text{Cohen's } d = \frac{M_1 - M_2}{SD_{pooled}} \quad \text{where } SD_{pooled} = \sqrt{\frac{(SD_1^2 + SD_2^2)}{2}}$$

and *d* = effect size *M* = mean *SD* = standard deviation

Guidelines developed by Cohen were used to interpret the strength of the effect in the results (Pallant, 2010).

Cohen's *d*	**Squared size (SD units)**
Small	.2
Medium	.5
Large	.8

Table 3.23: Cohen's d guidelines for interpreting effect size strength

F-values for inter-group statistics showed an overall significant age effect (p<0.001) on performance in PA tasks (**Table 3.24**). A post-hoc Bonferroni correction test was carried out to determine the extent to which means for the age groups differed from each other.

Table 3.24: Means, SD, ANOVA, effect size and strength (Cohen's d) and p-values for pair-wise comparisons between age groups for PA tasks

		AGE GROUPS										
		4.0 4.11		5.0 5.11		6.0 7.0		Df	F	Effect size	Strength of effect size (Cohen's d)	*p* value for pair-wise comparison
	Measure	*M*	*SD*	*M*	*SD*	*M*	*SD*					
PA TASKS	***Rhyme Awareness (n=8)***	5.17	1.65	5.93	1.49	6.05	1.65	384	13.30***	a 0.48 b 0.60 c 0.09	Medium Medium Small	**<.001***** **<.001***** 1.00
	Syllable Segmentation (n=8)	4.80	1.61	5.44	1.48	5.92	1.69	384	15.78***	a 0.41 b 0.68 c 0.30	Small Medium Small	**.005***** **<.001***** **.042***
	Alliteration Awareness (n=8)	2.70	1.15	3.31	1.64	4.02	1.91	384	21.61***	a 0.44 b 0.86 c 0.40	Small Large Small	**.008**** **<.001***** **.001****
	Phoneme Isolation (n=8)	1.74	1.89	3.18	2.84	4.71	2.47	384	47.36***	a 0.61 b 1.36 c 0.58	Medium Large Medium	**.000***** **<.001***** **<.001*****
	Blending (n=20)	5.71	2.29	6.23	2.33	8.16	2.88	384	34.01***	a 0.23 b 0.95 c 0.74	Small Large Medium	.311 **<.001***** **<.001*****

Age groups: a=youngest vs. intermediate b=youngest vs. oldest c=intermediate vs. oldest
****F-value significant at 0.001 level (p< 0.001) ** F-value significant at 0.01 level (p< 0.01) *F-value significant at 0.05 level (p< 0.05)*

These findings confirmed the observations obtained from the confidence intervals (Appendix M). Comparing confidence levels showed that on the Rhyme Awareness task, the confidence intervals for the youngest group were 4.91-5.43 and did not overlap with the intermediate and oldest groups for which the confidence intervals overlapped substantially, being 5.67-6.19 and 5.80-6.30, respectively. This suggests no difference between the means for these older two age groups on this task. There was little overlap between the intermediate and oldest groups for the Syllable Segmentation task, with confidence intervals of 5.16-5.72 and 5.65-6.20, respectively. There was also some overlap between the confidence intervals of the youngest and intermediate age groups on the Blending task, with 5.26-6.15 and 5.79-6.67, respectively, suggesting a similar performance on this task by the two youngest age groups. For Alliteration Awareness and Phoneme Isolation the confidence intervals did not overlap for any of the age groups.

In **Table 3.24** post-hoc analysis of pair-wise comparisons showed that virtually all of the PA tasks produced a significant age effect between each age group, with only two exceptions. No significant difference was found for the Rhyme Awareness task between the intermediate and oldest groups ($p>0.05$), while the performance of the youngest group was significantly lower than that of the two older groups ($p<0.001$). In the Blending task no significant difference was noted between the youngest and the intermediate groups ($p>0.05$). These findings confirmed the observations inferred from the confidence intervals (Appendix M). This suggests that in the Rhyme Awareness task the difference between mean scores for the two older groups was very small. In the Blending task the performances of the two youngest groups were similar.

Table 3.24 also shows the results for effect size *d* (Cohen's *d*) indicating the effect that age has on the variance in test scores. The smallest effect size (0.09) was obtained in the Rhyme Awareness task for the intermediate and oldest groups, showing very similar performances between these participants and suggesting evidence of the ceiling effect**.**[15]

---

[15] This refers to the level at which an independent variable no longer has an effect on a dependent variable, or to the level above which variance in an independent variable is no longer measured or estimated.

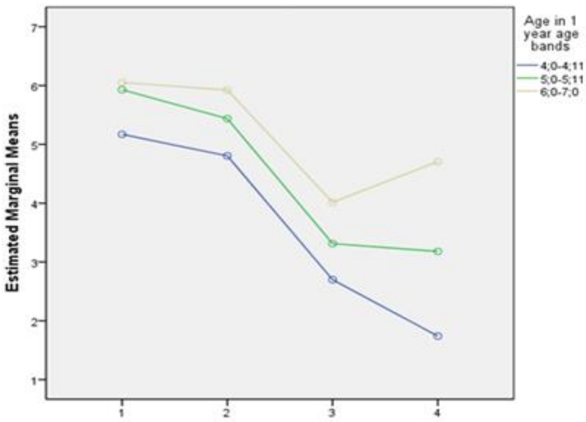
Syllable Segmentation and Alliteration Awareness tasks also produced the smallest effect sizes between the intermediate and oldest groups, whereas effect sizes were small for the Blending task between the youngest and the intermediate groups, indicating a move toward the floor effect.[16]

3.14.3.2.2 PA tasks: evaluating age sensitivity

Age was measured against task difficulty to find out which tasks proved easiest or most difficult according to the age groups, to determine the suitability of the four PA tasks devised by the researcher (i.e. Rhyme Awareness, Syllable Segmentation, Alliteration Awareness and Phoneme Isolation) for assessing children's skills by age. Mixed-design/split-plot ANOVA analysis[17] was used to test for differences using one 'between factor' (age with three levels) and one 'within factor' (PA with four levels).

As previously noted (**Table 3.24**), performance varied according to the difficulty of the task. A line graph was created to illustrate where the interaction effect was occurring (**Figure 3.4**).

Graphical displays of this kind can furnish a basic understanding of participants' performance. However, they do not provide information about the presence or absence of significant differences between groups.



PA Tasks

Figure 3.4: Comparison of performance in four PA tasks by age group

---

[16] A floor effect arises when a data-gathering instrument has a lower limit to the data values it can reliably specify.

[17] This is used to test for differences between two or more independent groups whilst subjecting participants to repeated measures. Thus, in a mixed-design ANOVA model, one factor (a fixed effects factor) is a between-subjects variable and the other (a random effects factor) is a within-subjects variable.

KEY: 1= Rhyme Awareness 2= Syllable segmentation 3= Alliteration Awareness 4= Phoneme Isolation

Error bars clearly show the distribution of these overlaps (see **Figure 3.5**). While there are many overlaps between the age groups for Rhyme Awareness, Syllable segmentation and Alliteration Awareness tasks; there are no overlaps between age groups for Phoneme Isolation. This suggests that although these tasks may not be sufficiently discriminating for clinical purposes they may identify children with very marked impairment.

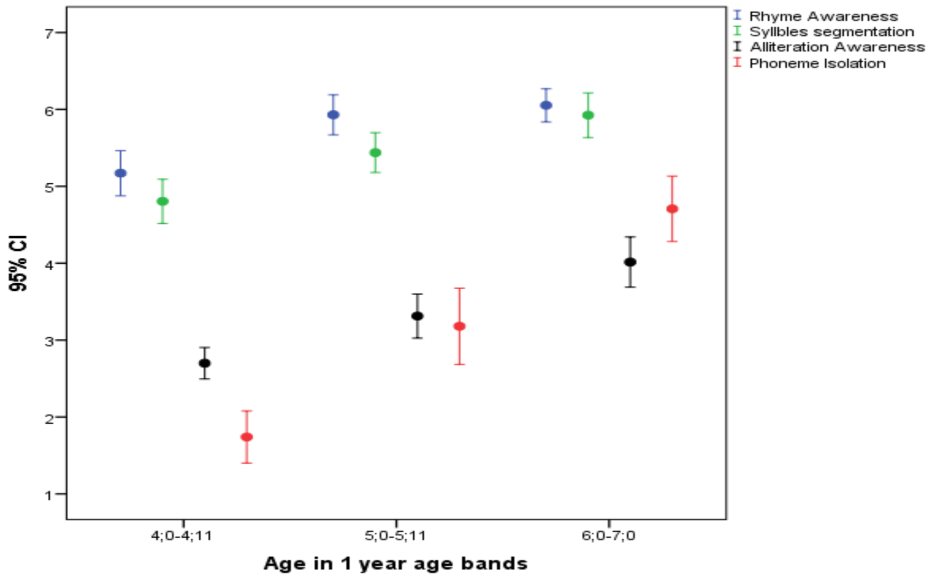


Figure 3.5: Comparison of performance in four PA tasks by age group using error bars

KEY: 1= Rhyme Awareness 2= Syllable segmentation 3= Alliteration Awareness 4= Phoneme Isolation

3.14.3.2.3 Composite PA measure: descriptive analysis

After establishing the age sensitivity for four PA measures in the RRS battery designed by the researcher, a single PA composite measure was created and used as a dependent variable in subsequent analysis. This composite measure was created across the four variables (Rhyme Awareness, Syllable Segmentation, Alliteration Awareness and Phoneme Isolation) by totalling the individual scores for the four tasks.

Performance improved with age. As the box plot shows, the youngest group showed the smallest spread (smallest interquartile range and standard deviation), with one ordinary outlier. The two older age groups exhibited similar spreads. All three groups exhibited non-symmetry. Comparing the values of the means for each group and box plot distributions, all groups show a positive skew. **Table 3.25** shows an improvement in performance in PA as age increases. The mean scores for children by age are youngest group, 14.41; intermediate group, 17.86; and oldest group, 20.70. Comparing the means for all three groups using the composite PA measures shows that children's achievement increases with age, meaning that this is an important factor in raising performance levels in PA.

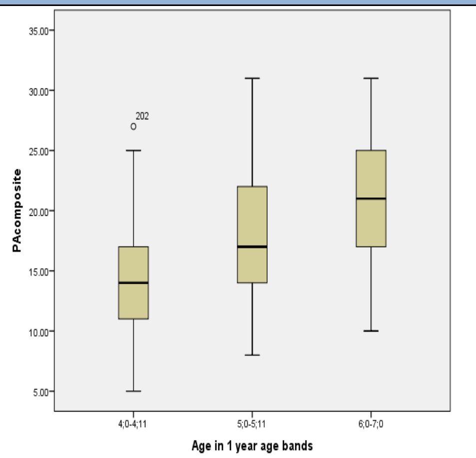
	COMPOSITE PA MEASURE*		
Age Group	**Mean**	**SD**	
4.00-4.11 (*n*=123)	14.41	4.39	
5.00-5.11 (*n*=128)	17.86	5.06	
6.00-7.00 (*n*=133)	20.70	5.07	

Table 3.25: Descriptive statistics and box plots: composite PA measure*.

**=* *Rhyme Awareness; Syllable Segmentation; Alliteration Awareness and Phoneme Isolation tasks*

3.14.3.2.4 Composite PA measure: inferential statistical analysis

As previously (see **section 3.14.3.2.1**), a univariate ANOVA was used to test for age group differences in scores and a post-hoc Bonferroni correction applied to determine whether age-group means differed significantly. Effect sizes were calculated to determine whether age was a significant factor in performance level and Cohen's guidelines (**Table 3.26**) applied to interpret effect sizes.

Table 3.26: Means, SD, ANOVA, effect size and strength (Cohen's d) and p-values for pair-wise comparisons between age groups for the composite *PA measure*

	AGE GROUPS										
	4;0-4;11		5;0-5;11		6;0-7;0						
MEASURE	M	SD	M	SD	M	SD	df	F	Effect size	Strength of effect size (Cohen's *d*)	*p* value for pair-wise comparisons
COMPOSITE PA	14.41	4.40	17.86	5.06	20.70	5.07	2	53.51*	a 0.73 b 1.32 c 0.56	Medium Very large Medium	**<.001***** **<.001***** **<.001*****

Age groups: a=youngest vs. intermediate b=youngest vs. oldest c=intermediate vs. oldest

**F-value significant at 0.05 level (p< 0.05) ** F-value significant at 0.01 level (p< 0.01) ***F-value significant at 0.001 level (p< 0.001)*

Table 3.26 shows that age has a significant effect on performance in the composite PA tasks as shown by the F value= 53.51* (*p*<.001). Pair-wise comparisons by age group show that performance in the composite PA tasks improved as age increased. The effect size of youngest vs. oldest (1.32) and intermediate vs. oldest (0.56) is significant with a very large and medium effect respectively.

3.14.3.3 <u>LK task</u>

<u>*3.14.3.3.1 Descriptive analysis*</u>

Performance for the LK task improves with age (**Table 3.27**). The mean values for performance on LK tasks almost doubled between successive ages (3.96 for the youngest group, 8.91 for the intermediate group, and 16.98 for the oldest group); the median values quadrupled.

The youngest group showed the smallest spread (smallest interquartile range and standard deviation), with four ordinary and nine extreme outliers. The two older groups exhibited a similar spread to each other. The oldest group had a very high median of 20, indicating that 50% of children recognized at least 20 letters. All three groups exhibited non-symmetry.

Comparing mean and median values for each group and box-plot distributions, the youngest and intermediate groups show a positive skew with most data values being low (mean>median), whereas the oldest group shows some negative skew (mean<median), suggesting that most children performed well on the LK task.

Age Group	Mean	SD	Med	Min	Max (*n*=28)
4;0-4;11 (*n*=123)	**3.96**	6.12	**1**	0	27
5;0-5;11 (*n*=128)	**8.91**	9.22	**5**	0	28
6;0-7;0 (*n*= 133)	**16.98**	9.59	**20**	0	28

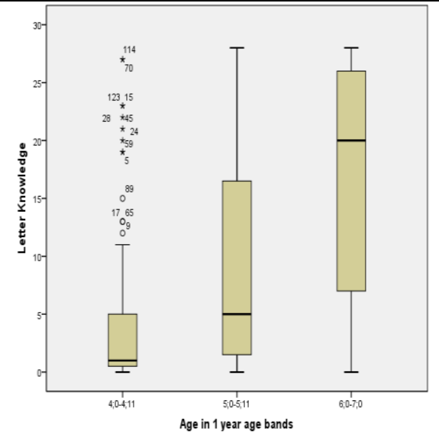


Table 3.27: Descriptive statistics and box plots: LK task

These results suggest that LK skill develops over time and that the children's ability to recognize letters increases as they get older.

3.14.3.3.2 Inferential statistical analysis

The same statistical analysis was carried out as previously for the PA tasks (see section **3.14.3.2.1**). *F*-values for between-groups statistics showed a significant age effect at $p>0.001$ on performance in the LK task, when carrying out pair-wise comparisons by age group (*Table 3.***Table 3.28**). Post-hoc analysis of these comparisons showed that age has a sig**n**ificant effect on the LK task with a medium to large effect size.

These findings confirmed the observations obtained from the confidence intervals (Appendix M) for LK, where the confidence intervals did not overlap at all for any of the age groups, with quite substantial differences between the values for successive age groups.

Table 3.28: Means, SD, ANOVA, effect size, effect strength (Cohen's d) and p-values for pair-wise comparisons for age groups on LK Task

	AGE GROUPS										
	4;0-4;11		5;0-5;11		6;0-7;0		Df	F	Effect size	Strength of effect size (Cohen's *d*)	*p* value for pair-wise comparisons
Measure	**M**	**SD**	**M**	**SD**	**M**	**SD**					
LK TASK	3.96	6.12	8.91	9.22	16.98	9.59	384	77.04***	a 0.65 b 1.66 c 0.86	Medium Very large Large	**<.001***** **<.001***** **<.001*****

Age groups: a=youngest vs. intermediate b=youngest vs. oldest c=intermediate vs. oldest

**F-value significant at 0.05 level (p< 0.05) ** F-value significant at 0.01 level (p< 0.01) ***F-value significant at 0.001 level (p< 0.001)*

3.14.3.4 RAN tasks

3.14.3.4.1 Attempted vs. valid RAN tasks

In the case of the RAN tasks, only valid cases were considered when performing statistical analysis. These were cases in which children had error scores of less than five in one or both forms of each RAN test. **Table 3.29** shows the number of attempted cases and those considered valid for all age groups across all four RAN tasks.

		ATTEMPTED CASES[1]			VALID CASES[2]		
	MEASURES	4;0-4;11	5;0-5;11	6.0-6.11	4;0-4;11	5;0-5;11	6.0-6.11
RAN TASKS	***Colour***	71	104	120	15	30	71
	Object	101	112	115	18	38	58
	Letter	2	26	72	-	11	52
	Digit	7	22	76	6	12	55
TOTAL CASES		**181**	**264**	**383**	**39**	**91**	**236**
		47.0%	**69.0%**	**99.0%**	**21.5%**	**34.5%**	**62.0%**

Table 3.29: Attempted and valid cases for all age groups across all RAN tasks. [1] *Attempted cases refers to all children in the pre-test.* [2] *Valid cases refers to all children who made less than 5 mistakes*

As **Table 3.29** shows, the number of valid cases increased markedly with age. The youngest group achieved the least valid cases across all RAN tasks and none of the children produced a valid performance for Letter Naming. The highest number of valid cases was always found to be in the oldest group, with the highest number of valid cases overall for the Colour Naming task. A clearer picture of the potential value/informativeness of the task for each age group can be gained from considering the percentages of attempted and valid cases for each group. It can be concluded that the RAN tasks were most useful for the oldest group.

3.14.3.4.2 Descriptive Analysis

In **Tables 3.30- 3.33** the figures refer to the time taken to complete each RAN task and are given in seconds. The values of the means and medians and the box plot distribution show that the older children performed faster on the Colour Naming task (**Table 3.30**) Some 50% of the intermediate group completed this task in less than 41 seconds; for the oldest group, less than 37 seconds. The youngest group recorded the slowest performance (with a median of 46 seconds) and data points tend to cluster around this time. There were ordinary outlier times of 60.20 seconds for the intermediate group and 55.10 seconds for the oldest

Age Group	Mean	SD	Med	Min	Max
4;0-4;11 (*n*=15)	46.68	7.39	46	35.50	58.50
5;0-5;11 (*n*=30)	41.60	6.52	40.60	31.25	60.20
6;0-7;0 (*n*= 71)	38.22	5.91	37.35	25.20	55.10

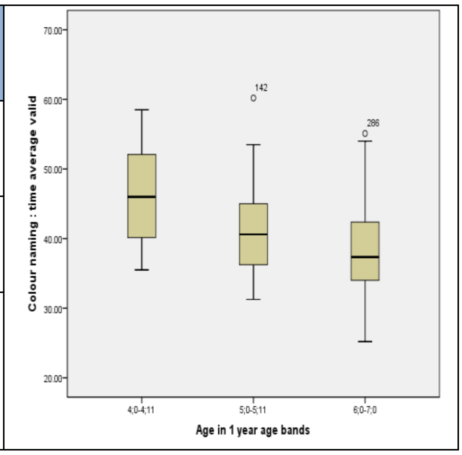


Table 3.30: Descriptive statistics and box plots of RAN: Colour Naming

For Object Naming (**Table 3.31**), the two youngest groups showed similar mean and median values (approximately 49 and 48 seconds respectively), with the intermediate group exhibiting a greater standard deviation and range. The oldest group performed fastest on the Object Naming task. The box plot for this group showed just one outlier value in the upper range with 50% of children in this group completing the test in 42 seconds or under.

Age Group	Mean	SD	Med	Min	Max
4;0-4;11 (*n*=18)	48.96	4.77	47.68	40.50	56.55
5;0-5;11 (*n*=38)	48.60	6.69	47.63	35.95	60.40
6;0-7;0 (*n*= 58)	42.75	6.36	42.00	31.00	59.00

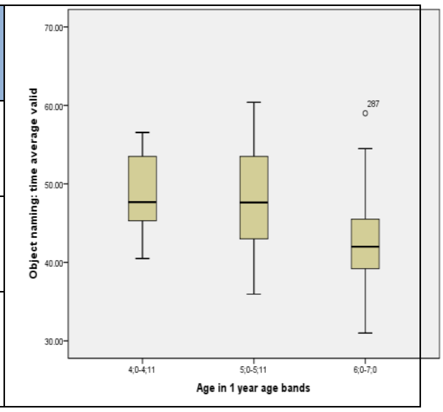


Table 3.31: Descriptive statistics and box plots of RAN: Object Naming

For Letter Naming (**Table 3.32**), there were no valid cases in the youngest group. For the other two groups, the mean and median scores for the oldest children were smaller than those of their intermediate counterparts, with the oldest group being fastest, completing the task in just over 24 seconds on average. The oldest group showed the broadest range, the intermediate group having just one ordinary outlier value of approximately 53 seconds.

Age Group	Mean	SD	Med	Min	Max
4;0-4;11 (*n*=0)	--	--	--	--	--
5;0-5;11 (*n*=11)	41.15	6.96	41.75	30.25	53.45
6;0-7;0 (*n*= 52)	35.23	7.27	34.33	24.05	53.00

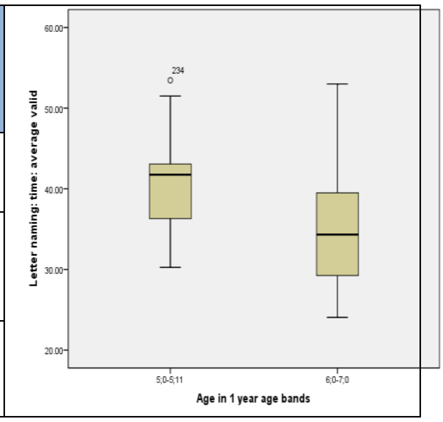


Table 3.32: Descriptive statistics and box plots of RAN: Letter Naming

In the Digit Naming task (**Table 3.33**), mean and median times both decreased with age, with the fastest times dropping from just over 40 seconds for the youngest group to 29 for intermediate, and just over 22 for the oldest. The intermediate age group showed the broadest spread of times taken to complete the Digit Naming task, as shown in the box plot and largest standard deviation.

Age Group	Mean	SD	Med	Min	Max
4;0-4;11 (*n*=6)	47.41	6.90	46.00	40.25	58.00
5;0-5;11 (*n*=12)	43.33	9.28	41.71	29.00	58.50
6;0-7;0 (*n*= 55)	36.82	6.19	37.00	22.20	47.75

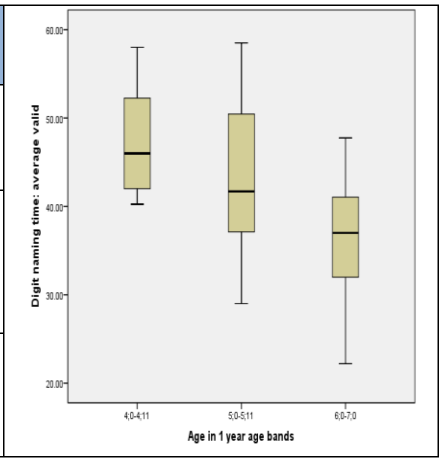


Table 3.33: Descriptive statistics and box plots of RAN: Digit Naming

<u>3.14.3.4.3 Inferential statistical analysis</u>

The same statistical analysis was carried out as previously for the PA tasks (see section **3.14.3.2.1**). For the RAN tasks, the *F*-values for between-group statistics showed an overall effect of age in inter-group comparisons. Post-hoc analysis of pair-wise comparisons showed a significant age effect across all RAN tasks except in the case of the Object and Digit Naming tasks when comparing performances by the youngest and intermediate groups. There was no significant difference for both these tasks between these age groups (*p*>0.05), whilst for Colour Naming the youngest group had significantly longer times than those of the intermediate group at the 0.05 significance level. Performance times for the youngest group were significantly longer than those of the eldest group, with a significance level of $p<0.001$ for Colour and Object Naming tasks, and $p<0.01$ for Digit Naming (**Table 3.33**).

A significant difference was noted ($p< 0.05$) in the mean performance times between the intermediate and oldest groups for Colour, Letter and Digit Naming tasks, with a highly significant difference (p<0.001) for Object Naming.

Comparing confidence levels (Appendix M) showed that on the Colour Naming task, there was very little overlap between these for the successive age groups, suggesting a difference in the performance times of all three age groups on this task.

On Object Naming, the confidence intervals for the youngest and intermediate groups overlapped substantially (46.03 to 51.88 and 46.59 to 50.62, respectively), suggesting no significant difference between the age-group means for this task. There was no overlap with the confidence limits for the oldest group.

On the Letter Naming task, there were no valid cases for the youngest participants. Some overlap was found between the intermediate and the oldest groups, with confidence levels of 36.80 to 45.50 and 33.23 to 37.24, respectively.

On the Digit Naming task, the confidence levels for the oldest group (34.99 to 38.65) did not overlap with the two other age groups. There was some overlap however between the youngest and the intermediate groups, with confidence limits of 41.87 to 52.97 and 39.40 to 47.25, respectively, suggesting some similarity in the mean performance of both the younger groups on this test.

With the exception of two medium effect sizes between the intermediate and oldest groups for Colour Naming and between the intermediate and youngest group for Digit Naming, all other effect sizes were found to be large. This excludes the difference between the intermediate and youngest group for Object Naming: as the sample size was very small the test may have been underpowered. Again, these findings confirmed the observations obtained from the confidence intervals (Appendix M).

Table 3.34 Means, SD, ANOVA, effect size, and strength (Cohen's d) and p-values for pair-wise comparisons for age groups on RAN Tasks

<table>
<tr><td></td><td></td><td colspan="6">AGE GROUPS</td><td rowspan="3">F</td><td rowspan="3">Effect size</td><td rowspan="3">Strength of effect size (Cohen's d)</td><td rowspan="3">p value for pair-wise comparison</td></tr>
<tr><td></td><td></td><td colspan="2">4.0 4.11</td><td colspan="2">5.0 5.11</td><td colspan="2">6.0 6.11</td></tr>
<tr><td></td><td>Measure</td><td>M</td><td>SD</td><td>M</td><td>SD</td><td>M</td><td>SD</td></tr>
<tr><td rowspan="4">RAN TASKS</td><td>Colour</td><td>46.68</td><td>7.39</td><td>41.60</td><td>6.25</td><td>38.22</td><td>5.91</td><td>12.29***</td><td>a 0.75
b 1.27
c 0.54</td><td>Large
Large
Medium</td><td>.035**
<.001***
<.045*</td></tr>
<tr><td>Object</td><td>48.96</td><td>4.77</td><td>48.60</td><td>6.69</td><td>42.75</td><td>6.36</td><td>12.94***</td><td>a 0.06
b 1.11
c 0.90</td><td>Small
Large
Large</td><td>1.00
<.001***
<.001***</td></tr>
<tr><td>Letter</td><td>--</td><td>--</td><td>41.15</td><td>6.96</td><td>35.23</td><td>7.27</td><td>6.10**</td><td>--
--
c 0.83</td><td>--
--
Large</td><td>--
--
.016*</td></tr>
<tr><td>Digit</td><td>47.41</td><td>6.90</td><td>43.33</td><td>9.28</td><td>36.82</td><td>6.19</td><td>9.76***</td><td>a 0.51
b 0.62
c 0.84</td><td>Medium
Large
Large</td><td>.702
.002**
.011*</td></tr>
</table>

Age groups: a=youngest vs. intermediate b=youngest vs. oldest c=intermediate vs. oldest

*F-value significant at 0.05 level (*p*≤0.05)** F-value significant at 0.01 level (*p*≤0.01) ***F-value significant at 0.001 level (*p*≤0.001)

In the RAN tasks, the performance times of the youngest group were significantly longer than those of the oldest group for Colour and Object Naming. The intermediate and oldest groups showed a significant difference in mean performance times for Colour, Letter, Digit and Object Naming.

3.14.3.5 Non-parametric tests to investigate between-group differences in performance on RRS tasks by age

3.14.3.5.1 PA and LK tasks

Non-parametric tests were conducted to check the parametric analysis of the effect of age on the five PA tasks (Rhyme Awareness, Syllable Segmentation, Alliteration Awareness, Phoneme Isolation and Blending) and LK to confirm the pair-wise between-group comparisons. The results of these non-parametric analyses using Kruskal-Wallis and Mann-Whitney U tests to investigate the effect of age on the PA tasks and LK are provided in Appendix L.

There were highly significant differences between the youngest and oldest groups for all tasks ($p < 0.001$). The non-parametric findings confirm a significant difference between the youngest and intermediate age groups for the PA tasks (Rhyme Awareness, Syllable Segmentation, Alliteration Awareness and Phoneme Isolation), Blending, and also LK. However, there was no significant difference between the two oldest groups for the Rhyme Awareness task.

3.14.3.5.2 RAN tasks

The non-parametric findings confirm a significant difference ($p < 0.001$) between the youngest and oldest groups for RAN tasks focusing on Colour, Object and Digit Naming. There were less significant differences ($p < 0.01$) on the Digit Naming task between the intermediate and youngest groups and significant differences on the Letter Naming task ($p < 0.05$).

Comparing the intermediate and oldest groups, significant differences were found for the Colour and Object Naming tasks ($p < 0.001$) and the Digit Naming task ($p<0.01$). Significant differences were obtained between the youngest and intermediate age groups for the Object Naming task ($p < 0.001$) and the Colour, Letter and Digit Naming tasks ($p < 0.05$).

3.14.4 Socioeconomic Status and RRS

This section addresses the effect of socioeconomic status (SES) on the composite PA measure, Blending and LK. Although the questionnaire provided many different types of data, the decision was made to use the SES factors most commonly used in previous studies, namely level of parental education, family income, and property type. The remaining data can be used in future studies.

3.14.4.1 Descriptive analysis

Table 3.35 shows the educational qualifications for the parents of the children participating in the study. Most parents have a bachelor's degree: mothers (41.7%) and fathers (34.6%). Some 2.6% of mothers and 8.6% of fathers had postgraduate qualifications such as Masters degrees or PhDs.

EDUCATION	MOTHER	FATHER
No qualifications	4.2	2.4
Elementary education	8.9	6.3
Intermediate education	13.1	11.7
High school education	23.4	26.0
Diploma	6.0	9.9
Bachelor's Degree	41.7	34.6
Postgraduate Education	2.6	8.6
Nil Response	0.3	0.5
TOTAL %	100%	100%

Table 3.35: Parental education profile for the sample

Table 3.36 presents the participant profile according to area in which the school was situated, the type of property where the family reside, and their monthly income. There is an equal spread of children across the four districts of Riyadh. Nearly a third of the participants lived in flats whilst over a half lived in villas (the name given in Saudi Arabia to large detached houses with their own grounds). However, nearly a quarter of these lived in shared residences. Almost 85% of the families surveyed had an income which totalled 6,000 or more Saudi riyals per month.

		VALID PERCENT	TOTAL PERCENT	TOTAL
SCHOOL AREA	North	25	100	384
	South	25		
	East	25		
	West	25		
PROPERTY TYPE	Studio flat	3.9	100	384
	Flat	32.0		
	Multi-occupancy Villa	24.7		
	Villa	37.9		
	Palace	1.0		
	Nil response	.5		
FAMILY INCOME[1]	Less than 3000 SR	5.5	100	384
	3 000-5 999 SR	9.6		
	6 000-8 999 SR	22.9		
	9 000-11 999 SR	19.6		
	12 000-14 999 SR	13.0		
	15 000-17 999 SR	10.7		
	18 000-20 999 SR	10.2		
	21 000 + SR	8.2		
	Nil response	.3		

Table 3.36: Participant profile: School area, property type and family income. [1]*Family income is amount per month in Saudi Riyals (SR)*

3.14.4.2 Correlation between SES variables and age

An initial analysis was carried out to find the relationship between SES variables and age to determine if there was any association between age and the SES variables in order to be able to statistically control for age as a possible additional variable and avoid this influencing the relationship between SES factors. A Pearson correlation (r) was carried out and the results showed no correlation between age of participants and any of the key SES variables.

3.14.4.3 Correlation between SES variables

A correlation analysis was then carried out to attempt to create composite measures for the key SES factors and then use these to determine possible links between these factors and performance in some key RRS. As **Table 3.37** shows, a small significant negative correlation was found between school area and property type ($r=-.146$, $p<0.01$), as well as family income ($r=-.203$, $p< 0.01$).

A moderately significant positive correlation was found between the mother's and father's level of education (*r*=450, p<0.01), suggesting that parents tend to have a similar level of educational qualifications. As might be expected, there was also a moderate positive correlation between property type and family income (*r*=.467, p<0.01) suggesting that the amount of money people have tends to affect the type of property they can afford to rent or buy. Perhaps the most striking finding was that school type (public or private) was not related to any SES variable. In addition, parental education was not correlated with family income or property type.

			SCHOOL TYPE	SCHOOL AREA	EDUCATION MOTHER	EDUCATION FATHER	PROPERTY TYPE	FAMILY INCOME RANGE
SCHOOL TYPE		P	1					
		Sig. (2-tailed)						
SCHOOL AREA		P	.000	1				
		Sig. (2-tailed)	1.000					
EDUCATION	MOTHER	P	.042	.089	1			
		Sig. (2-tailed)	.414	.082				
	FATHER	P	.063	.058	**.450****	1		
		Sig. (2-tailed)	.220	.258	.000			
PROPERTY TYPE		P	.039	**-.146****	-.008	.039	1	
		Sig. (2-tailed)	.452	.004	.878	.450		
FAMILY INCOME		P	.052	**-.203****	-.059	.049	**.476****	1

Table 3.37: Pearson correlations (r) between SES variables.

***Correlation significant at the 0.01 level (2-tailed) ***Correlation significant at the 0.001 level (2-tailed).*

The information regarding parental education, originally provided as two separate figures for the father's and the mother's educational qualifications, was combined to produce a composite category (parental education) given that a moderate correlation was found between these. Likewise, property type and family income were also combined for the same reason.

3.14.4.4 Correlation between SES variables and PA, Blending and LK

The analysis here focuses on determining possible links between four key SES factors (school type, school area, parental education, and a composite category combining property type and family income) and participant performance in three areas: the composite PA measure (as used previously), Blending and LK. Blending has been included because it is a good measure of explicit PA. RAN will not be analysed for all SES factors. The correlation between these SES variables and participant performance in the composite PA measure, the Blending task, and the LK task was investigated, with age partialled out, by performing Pearson correlations (*r*). The results are shown in **Table 3.38**.

Composite PA measures were significantly correlated with the composite measure of family income and property type (*r*=274, p<0.001), school type (*r* =.138, p< 0.01), and parents' level of education (*r*=.131, p<0.05). The Pearson correlation is a measure of the linear correlation between two variables *X* and *Y*, giving a value between +1 and −1 inclusive, where 1 is a positive correlation, 0 is no correlation, and −1 is a negative correlation. All correlations were positive. In the case of the Blending task, a significant positive correlation was only found with school area (*r*=.156, p<001). The LK task was positively and significantly correlated with both school area (*r*=.260, p<0.001) and parental education levels (*r*=.114, p<0.05).

School type was not correlated with any other SES variable. School area was correlated negatively with family income and property type (*r*=-.215, p<0.001) and does not correlate with any other SES variable. It was interesting that parental qualifications were not correlated with all the SES variables.

The overall sizes of correlations were generally small, but varied somewhat according to the tasks involved, with the more comprehensive composite PA measure yielding a weak correlation. The key finding is that the effect of SES appears to be very small whereas in Western studies SES has been shown to play a major role (for example, McDowell et al., 2007; Nobel et al., 2006). It is worth noting that significant correlations imply that there is a relationship between variables, but this does not necessarily imply causation. For this purpose, regression analyses were carried out.

AGE (CONTROL)		PA COMPOSITE MEASURE	BLENDING TASK	LK TASK	SCHOOL TYPE	SCHOOL AREA	PARENTAL EDUCATION	FAMILY INCOME & PROPERTY
PA COMPOSITE MEASURE	*R*	1.000						
	Sig. (2-tailed)	.						
BLENDING TASK	*R*	**.274*** **	1.000					
	Sig. (2-tailed)	.000	.					
LK TASK	*R*	**.428*** **	**.316*** **	1.000				
	Sig. (2-tailed)	.000	.000	.				
SCHOOL TYPE	*R*	**.138** **	-.017	.036	1.000			
	Sig. (2-tailed)	.007	.749	.483	.			
SCHOOL AREA	*R*	.040	**.156** **	**.260*** **	.000	1.000		
	Sig. (2-tailed)	.434	.002	.000	.995	.		
PARENTAL EDUCATION	*R*	**.131* **	-.019	**.114* **	.069	.094	1.000	
	Sig. (2-tailed)	.010	.717	.026	.183	.068	.	
FAMILY INCOME & PROPERTY	*R*	**.274*** **	.027	-.014	.061	**-.219*** **	.000	1.000
	Sig. (2-tailed)	.000	.602	.780	.235	.000	.999	.

Table 3.38: Pearson correlations (r) between SES variables and composite PA tasks, Blending task and LK, controlling for age

**Correlation significant at 0.05 level (2-tailed) **Correlation significant at 0.01 level (2-tailed) ***Correlation significant at 0.001 level (2-tailed)*

3.14.4.5 Regression Analysis

3.14.4.5.1 Contribution of SES-related independent variables

Linear regression analyses were carried out in order to find out if SES measures could predict performance in the composite PA measure, Blending and LK, and to identify which SES measure is the best predictor of performance. Predictor (SES measures) and outcome (test scores) variables were entered into simple and multiple linear models in order to identify any SES measures that might predict performance in RRS tasks. This was done by obtaining the R square (R^2) or the square of the correlation coefficient to find out how much of the variance in the dependent variable (performance on composite PA measure, Blending and LK) is explained by the independent variable (SES measures). It also gives an indication of the relative contribution of each independent variable.

The dependent variables (composite PA measure, Blending and LK) are continuous as opposed to categorical independent variables (school type, parental education, and family income plus property type). The school type was dummy coded, and the degree of parental education was ranked from basic to higher, then a composite score was generated by combining these scores for both parents. Family income and property type were also added together.

Multiple regression analysis was used as a first step then age was added to the analysis using the standard enter method. To compare the contribution of each SES-related independent variable, β values were required. The highest β coefficient was .232, which is for the composite category of family income plus property type. This means that this variable makes the strongest unique contribution to explaining the dependent variable (composite PA measure) when the variance explained by all other variables in the model is controlled for. The β value for school type was lower (.098), indicating that it made less of a contribution to the composite PA measure variance.

The second step involved checking whether each independent variable makes a statistically unique contribution. If the significance value was less than .05, the variable (SES-related variables) is making a significant unique contribution to the prediction of the dependent variable (score for composite PA measure, Blending or LK.); if it is greater than .05 then the converse is true. In this case, the bolded values in **Table 3.39** are those variables that made a unique and statistically significant contribution to the prediction of the performance for composite PA measure, Blending and LK tasks.

The β values in **Table 3.39** are helpful for assessing the contribution of each variable: family income plus property type (=0.232) make the main contribution, followed by parental education (=.140) and school type (=.098). All these results are significant so the factor that is the best predictor of performance in the composite PA measure is family income plus property type.

With regard to the Blending task, β values follow a different order, beginning with school type (=−.019), then family income plus property type (=.018), and finally parental education (=.009). In this case, all results were non-significant.

Finally, with regard to the LK task, the order of β values differs once more: parental education (=.130) is followed by family income plus property type (=−.024), and school type (=.022). As before, all the results are non-significant with exception of parental education, meaning that this is the best predictor of performance in the LK task.

It is possible that one factor in a child's SES background is not enough to determine his/her performance in the different tasks. Therefore, it is essential to determine the combined effect of the predictor variables in order to understand which of the SES background variables in the presence of other variables best predicts performance in the composite PA measure. Therefore, over and above the control variable of age the final step of the multiple regression analysis showed that the significant predictor variables of SES are school type, parental education, and family income plus property type.

Table 3.39 presents the R^2 for these SES predictors, showing that together 8.8% (R^2=.088) of the variation in composite PA measure performance can be accounted for by the SES variables (F=12.070, df=379, p=.000) Furthermore, performance on the Blending task can also be explained by the combined effect of SES variables: 0.1% (R^2=.001) of the variation in the Blending task can be accounted by the SES variables. The same was true for LK, with 1.8% of the variation accounted for by SES variables (R^2=.018).

			UNSTANDARDIZED COEFFICIENT		STANDARDIZED COEFFICIENT			
			B	Std. Error	β	t	Sig.	F
COMPOSITE PA MEASURE	R^2=.088	**(Constant)**	9.664	1.448		.000	.000	F=12.070 df=379 p=.000
		School Type	1.083	.544	.098	1.991	**.047**	
		Income & Property	.505	.107	**.232**	4.708	**.000**	
		Parental Education	.291	.103	.140	2.834	**.005**	
BLENDING TASK	R^2=.001	**(Constant)**	6.664	.752		8.856	.000	F=.086 df=379 p=.968
		School Type	-.101	.283	-.019	-.359	.720	
		Income & Property	.020	.056	.018	.351	.725	
		Parental education	.009	.053	.009	.167	.868	
LK TASK	R^2 =.018	**CONSTANT**	5.529	2.727		2.028	.043	F= 2.332 df= 379 p=.074
		School Type	.437	1.024	.022	.427	.670	
		Income & Property	-.095	.202	-.024	-.470	.639	
		Parental Education	.490	.193	**.130**	2.537	**.012**	

Table 3.39: Combined predictors of performance in composite PA measure, and Blending and LK tasks.

3.14.4.5.2 Modelling age and SES factors as predictors of performance

A hierarchical multiple regression was carried out to evaluate the ability of the model (which includes the SES variables) to predict performance in the composite PA measure, Blending and LK tasks, after controlling for age. The aim is to discover whether the SES variables are still able to predict a significant amount of variance in the scores for the composite PA measure, Blending and LK tasks when the effect of age is controlled.

3.14.4.5.3 Age and SES factors as predictors of performance in the composite PA measure

In **Table 3.40** two models are shown. Model one refers to the first set of variables that were entered (age), while model two includes all variables that were entered in both sets, i.e. age and SES-related (school type, family income plus property type and parental education). After the variables for age had been entered, model one accounted for 18.7 % of the variance (.187 x 100). After the second set of SES-related variables were also included, model two accounted for 27.3% of the variance.

Model one (age alone) and model two (age plus SES-related variables) predicted scores of the dependent variable (composite PA measure) to a statistically significant degree (p-values <.05). Model two, which includes both sets of variables, was significant (F (4, 375) = 35.136, p< .0005).

To determine the extent to which each of the variables contributed to the overall effect, coefficient values are useful if any of the predictors are statistically significant. In this case, age and all SES-related variables (school type, family income plus property type and parental education) made a unique statistically significant contribution (less than .05). In order of importance according to their β values, age (=.431) ranks highest, followed by family income plus property type (=.241), parental education (=.112), and school type (=.102).

			UNSTANDARDIZED COEFFICIENT		*STANDARDIZED COEFFICIENT*			
			B	***Std. Error***	**β**	***t***	***Sig.***	***F***
		(Constant)	2.295	1.673		1.371	.171	*F*=87.057 *df*=378 *p*=.000
COMPOSITE PA MEASURE	**Model 1** **R^2**=.187 **Adjusted** **R^2**=.185	Age	.235	.025	.433	9.330	.000	
	Model 2 **R^2**=.273 **Adjusted** **R^2**=.265	**(Constant)**	-5.352	2.011		-2.661	.008	*F*=35.136 *df*=375 *p*=.000
		Age	.234	.024	**.431**	9.760	**.000**	
		School Type	1.119	.486	**.102**	2.300	**.022**	
		Income & Property	.524	.096	**.241**	5.463	**.000**	
		Parental Education	.232	.092	**.112**	2.524	**.012**	

Table 3.40: Modelling age and SES factors as predictors of performance in the Composite PA measure

3.14.4.5.4 Age and SES factors as performance predictors in Blending task

A similar modelling process was carried out to determine the contribution of variables on the performance of children in the Blending task, the results of which are shown in **Table 3.41**. After the variables for age had been entered, model one accounted for 14.3 % (.143 x 100) of the variance in the Blending task. Once the second set of SES-related variables were also included, model two accounted for 14.4% of the variance.

Model one (age alone) and model two (age plus SES-related variables) predicted scores of the dependent variable (Blending task) to a statistically significant degree (*p*-values <.05). Model two, including both sets of variables, was significant (F (4, 375) = 15.749, *p*<.0005).

As **Table 3.41** shows, age was the only variable that made a unique statistically significant contribution to scores in the Blending task: *p*-values less than .05, with a β value of .379. None of the SES-related variables made an individual contribution to performance in this particular RRS.

			UNSTANDARDIZED COEFFICIENT		STANDARDIZED COEFFICIENT			
			B	*Std. Error*	β	*t*	*Sig.*	*F*
BLENDING TASK		**(Constant)**	.054	.853		.063	.950	F=62.898 df=378 *p*=.000
	Model 1 R^2=.143 **Adjusted** R^2=.140	Age	.102	.013	.378	7.931	.000	
	Model 2 R^2=.144 **Adjusted** R^2=.135	**(Constant)**	.104	1.083		.096	.924	F=15.749 df=375 *p*=.000
		Age	.102	.013	**.379**	7.918	.000	
		School Type	-.086	.262	-.016	-.328	.743	
		Income & Property	.028	.052	.026	.540	.589	
		Parental Education	-.017	.049	-.16	-.338	.735	

Table 3.41: Modelling age and SES factors as predictors of performance in the Blending task

For a third time, the same modelling process was carried out to determine the contribution of variables on the performance of children in the LK task, the results of which are shown in **Table 3.42**.

Once the variables for age were entered, model one accounted for 24.6% (.246 x 100) of the variance in the LK task. Once the second set of SES-related variables were also included, model two then accounted for some 25.7% of the variance.

Model one (age alone) and model two (age plus SES-related variables) predicted scores of the dependent variable (LK task) to a statistically significant degree (*p*-values $<.05$). Model two, including both sets of variables, was again significant ($F(4, 375) = 32.424$, $p < .0005$).

Table 3.42 shows that only two variables, namely parental education and age, made a unique statistically significant contribution (less than .05) as predictors. According to their β values, age ranked first (=.490) followed by parental education (=.098).

			UNSTANDARDIZED COEFFICIENT		STANDARDIZED COEFFICIENT			
			B	Std. Error	B	t	Sig.	F
LK		(Constant)	-22.095	2.925		-7.555	.000	F=123.527d f=378 p=.000
	Model 1 R^2=.246 Adjusted R^2=.244	Age	.488	.044	.496	11.114	.000	
	Model 2 R^2=.257 Adjusted R^2=.249	(Constant)	-25.448	3.689		-6.898	.000	F= 32.424 df=375 p=.000
		Age	.482	.044	**.490**	10.976	.000	
		School Type	.510	.892	.026	.572	.568	
		Income & Property	-.055	.176	**-.**014	-.315	.753	
		Parental Education	.369	.169	**.098**	2.188	.029	

Table 3.42: Modelling age and SES factors as predictors of performance in the LK task

3.14.5 Exposure to additional tuition in Qur'anic recitation

This section addresses whether levels of exposure to tuition in Qur'anic recitation can affect children's performance in all the measures of RRS. A composite PA measure was once again used comprising Rhyme Awareness, Syllable Segmentation, Alliteration Awareness, and Phoneme Isolation.

3.14.5.1 Participant characteristics

Firstly, the questionnaire results were used to establish the number of children having extra Qur'anic recitation tuition. In terms of age profile, in public sector schools 12 (36.4%) were in the youngest group, 13 (39.4%) in the intermediate group and 8 (24.2%) in the oldest group. In the private sector, these figures were 6 (27.3%), 8 (36.4%), receptively.

Table 3.43 shows that 55 children or 14.32% of the total sample (n=384) received extra Qur'anic recitation tuition, some 33 (60%) at public schools and 22 (40%) at private schools. Participants were fairly evenly spread across age groups within each school sector. As for gender, 56% were male and 44% female. Numbers of males and females were fairly evenly spread across age groups within the private school sector, but boys predominated in the public schools.

In terms of age profile, in public sector schools 12 (36.4%) were in the youngest group, 13 (39.4%) in the intermediate group and 8 (24.2%) in the oldest group. In the private sector, these figures were 6 (27.3%), 8 (36.4%) and 8 (36.3%), respectively.

PARTICIPANTS EXPOSED TO EXTRA QUR'ANIC RECITATION TUITION											
Public schools (*n*=33)						Private schools (*n*=22)					
4;0-4;11		5;0-5;11		6.0-7		4;0-4;11		5;0-5;11		6.0-7	
M	F	M	F	M	F	M	F	M	F	M	F
8	4	7	6	5	3	3	3	4	4	4	4
12		13		8		6		8		8	
TOTAL = 55											

Table 3.43: Age, gender and school type for participants exposed to extra Qur'anic recitation tuition

The 55 participants from the main sample exposed to extra Qur'anic recitation tuition were identified and their existing scores for the composite PA measure, Blending, LK and RAN tasks were used for statistical analysis.

3.14.5.2 Descriptive analysis

Table 3.44 shows the mean performance in the RRS test battery (with the composite PA measure) by age group and level of exposure to Qur'anic recitation tuition, whether standard or extra. The means were very similar for participants who had standard tuition and those who were exposed to extra tuition. This is because the SD in each case is so wide, it covers the means for both scores. As a result, there is an overlap for all tasks: PA, Blending, LK and RAN (see **Table 3.44**)

		STANDARD TUITION *(n=329)*			EXTRA TUITION *(n=55)*		
RRS TASKS	**Age Group**	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**
Composite PA measure	4;0-4;11	14.02	4.23	106	16.88	4.70	17
	5;0-5;11	17.54	4.89		19.60	5.69	
	6;0-7;0	20.71	4.94		20.61	5.96	
Blending	4;0-4;11	5.71	2.34	108	5.71	1.96	20
	5;0-5;11	6.30	2.23		5.85	2.85	
	6;0-7;0	8.18	2.77		8.00	3.58	
LK	4;0-4;11	3.87	6.01	115	4.53	6.97	18
	5;0-5;11	8.16	8.54		12.95	11.70	
	6;0-7;0	16.79	9.66		18.22	10.99	
RAN Colour naming average time	4;0-4;11	52.91	6.99	61	53.56	7.438	10
	5;0-5;11	47.57	7.48	91	50.03	8.282	13
	6;0-7;0	42.24	7.96	102	42.06	6.401	18
RAN Object naming average time	4;0-4;11	54.46	5.47	86	55.73	7.44	15
	5;0-5;11	51.75	6.11	95	52.37	6.42	17
	6;0-7;0	46.29	7.17	98	46.96	7.60	17
RAN Letter naming average time	4;0-4;11	51.10	.	1	59.50	.	1
	5;0-5;11	49.89	10.14	25	53.25	.	1
	6;0-7;0	38.37	9.12	60	41.72	11.88	12

RAN Digit naming average time	4;0-4;11	48.85	6.64	5	49.13	12.55	2
	5;0-5;11	47.33	9.62	15	43.30	12.86	3
	6;0-7;0	39.46	7.83	64	41.95	7.556	12

Table 3.44: Mean and SD for results of RRS tasks by age group and exposure to Qur'anic recitation tuition

The results in **Table 3.45** report only valid performances for the RAN tasks. There are only negligible differences between the mean performance times across the four RAN tasks between those who had extra exposure to tuition in Qur'anic recitation and those who did not. The level of exposure did not appear to have had any effect on performance times in the RAN tasks.

VALID RAN TASK	LEVEL OF EXPOSURE TO TUITION	N	Mean	SD
Colour naming	Standard	60	38.10	6.25
	Extra	11	38.93	3.64
Object naming	Standard	51	42.72	6.67
	Extra	7	42.96	3.69
Letter naming	Standard	46	35.22	7.13
	Extra	6	35.37	8.76
Digit naming	Standard	47	36.40	6.13
	Extra	8	39.28	6.30

Table 3.45: Mean and SD for performance times (in seconds) on valid RAN tasks, and exposure to Qur'anic recitation tuition.

3.14.5.3 Inferential analysis

First, a univariate ANOVA was carried out to investigate the effect of extra tuition on the composite PA measure, Blending and LK, by age group. Between-group effect size (partial eta squared) was then calculated for participants with standard and extra exposure to tuition, and also by age group (**Table 3.46**).

As these results indicate, with regards to the effect of age, for the Composite PA measure and the Blending task there was a medium but significant effect. For the LK task, there was a significant large effect. The same table also shows significant effects for the level of Qur'anic recitation tuition on the composite PA measure (F (1, 53) = 5.21, p<0.05) and the LK task (F (1, 53) = 3.45, p<0.05). However the effects were small, at 0.01 for the composite PA measure and 0.5 for LK of the overall variation in performance on Composite PA and LK respectively attributable to the influence of extra exposure. This tuition produced no significant effect on the Blending task, suggesting it has no impact on this task. There was no significant interaction effect on any of the tasks, suggesting that the effect of the level of exposure to Qur'anic recitation tuition applies equally across all age groups.

		COMPOSITE PA	BLENDING	LK
Age	F	**17.78*** **	**15.93*** **	**37.38*** **
	P	.00	.00	.00
	Partial Eta Squared	.09	.08	.17
Extra Qur'anic recitation tuition	F	**5.21***	.32	**3.45***
	P	.02	.57	.05
	Partial Eta Squared	.014	.001	.01
Interaction	F	1.55	.13	1.09
	P	.22	.88	.34
	Partial Eta Squared	.01	.00	.01

Table 3.46: Effect of age and level of exposure to Qur'anic recitation tuition and intera0ction effect on the composite PA measure, Blending and LK tasks.

**F-value significant at 0.05 level ($p < 0.05$) **F-value significant at 0.01 level ($p < 0.01$) ***F-value significant at 0.001 level ($p < 0.001$)*

Graphical displays of performance of the age groups with/without extra tuition for the composite PA measure, Blending and LK tasks are shown in **Figure 3.6-Figure 3.7 and 3.8**. As previously noted, graphical displays do not provide concrete information about significant differences between groups and an ANOVA test was used for this purpose. These findings confirmed the observations obtained from the confidence intervals (Appendix N).

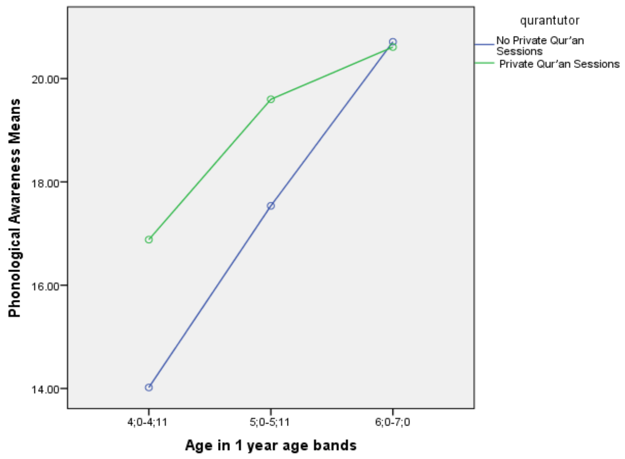


Figure 3.6: Mean performance on composite PA tasks by age group with/without extra tuition

For the composite PA measure, the greatest difference was to be found between the performances of the youngest and the intermediate groups with no difference for the older group.

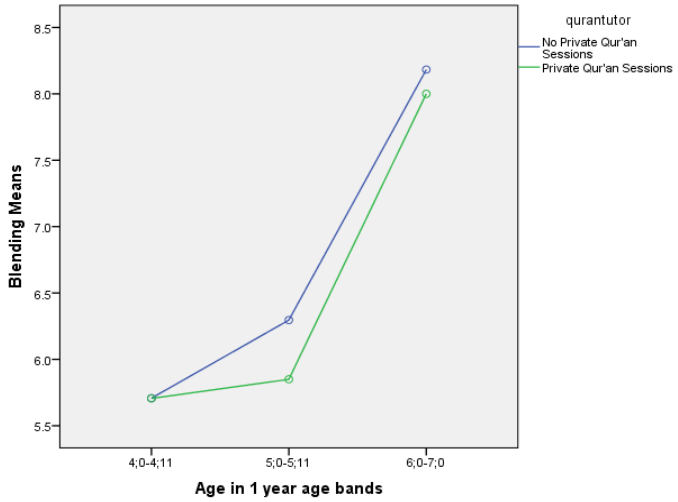


Figure 3.7: Mean performance on Blending task by age group with/without extra tuition

The greatest difference in performance on the Blending task was seen in the results for the intermediate group (5;0-5;11), as shown in **Figure 3.7**. In this figure, those not exposed to extra tuition (represented by the blue line) perform better than those taking extra tuition (green line). This is typically the age at which children start to learn to read individual words so one might expect this to be reflected in the results. However, these differences are less noticeable by the time children start learning to read formally when they enter school at the age of six. Again, no interaction effect was found using ANOVA.

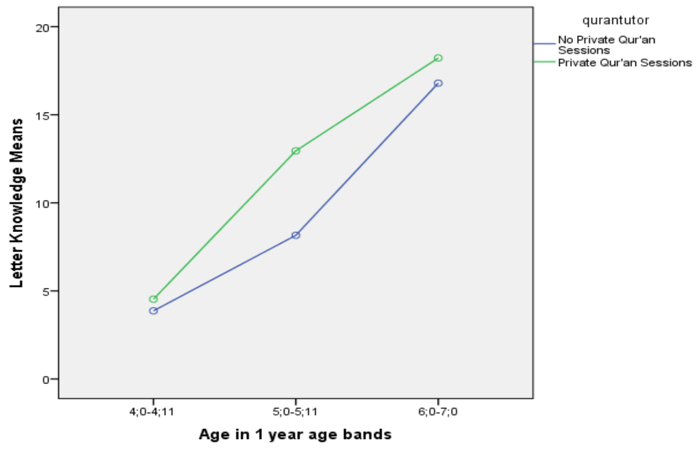


Figure 3.8: Mean performance on the LK task by age group with/without extra tuition

Finally, the most noticeable difference in performance on LK tasks was once again found in the intermediate group (5;0-5;11) as seen in **Figure 3.8**.

Figure 3.7 - 3.8 suggest that extra exposure to Qur'anic recitation tuition was more directly connected to literacy skills, since the Blending task and LK are more directly connected to this. Again, no interaction effect was found when ANOVA was used.

The test statistics for a one-way independent ANOVA for the four RAN tasks, including values of partial eta squared as a measure of effect size, are shown in **Table 3.47**. There was no significant effect of the level of exposure of Qur'anic recitation tuition on any of the four RAN tasks.

		RAN TASKS			
		Colour Naming	**Object Naming**	**Letter Naming**	**Digit Naming**
Age	**F**	3.076	7.465	5.901	3.107
	P	.050	.001	.018	.05
	Partial Eta Squared	.000	.121	.090	.085
Extra Quranic recitation tuition	**F**	.023	.645	.002	.396
	P	.879	.424	.962	.531
	Partial Eta Squared	.053	.006	.000	.006
Interaction	**F**	1.403	.132	.	1.096
	P	.250	.876	.	.340
	Partial Eta Squared	.025	.002	.000	.032

Table 3.47: Effect of age and level of exposure to Qur'anic recitation tuition and interaction effect on the RAN tasks with values of partial eta squared.

**F-value significant at 0.05 level ($p < 0.05$) **F-value significant at 0.01 level ($p < 0.01$) ***F-value significant at 0.001 level ($p < 0.001$)*

When Levene's Test of equality of error variance was applied to all the RAN tasks the assumptions were met (<.05) for all of these except the Colour-Naming task. This means that the age groups are equal for Object, Letter and Digit Naming tasks (see **Table 3.48**).

RAN Tasks	Sig.
Colour Naming	.003
Object Naming	.407
Letter Naming	.831
Digit Naming	.263

Table 3.48: Levene's Test of equality of error variance for RAN tasks.

3.15 Conclusion

This chapter has explained the methodological approach adopted in this research and provided a detailed description of the process of developing the instruments, materials and testing procedures employed in Study One to capture the required data. The methods employed to codify and analyse the data collected, including any ethical dimensions to be considered in research of this type, were also presented together with the results of statistical analysis of the data gathered. The following chapter, Chapter Four, focuses on Study Two which was designed to validate the Arabic RRS test battery developed in Study One by carrying out an evaluative test-retest procedure.

4. CHAPTER FOUR: STUDY TWO

4.1 Introduction

Study Two was designed to validate the RRS test battery by carrying out an evaluative test-retest procedure. To this end, the same test was administered on two occasions to a new sample of participants. This study also built on the results of Study One by investigating the predictive abilities of the RRS test battery and exploring the relationship between participants’ performance in the RRS test battery and their reading-related abilities as rated by the children’s teachers. Study Two employed the same instruments as those used in Study One, namely the RRS test battery and the SES parental questionnaire on a new sample of participants, so there was no need in this case for a pilot study. However, in Study Two the RAN task was only administered to the oldest group (i.e. 6;0-7;0), as this had proved to be an unreliable measure for assessing the two younger age groups in Study One. An additional measure of reading-related abilities was also incorporated which required teachers to provide input to the study.

This chapter provides a detailed account of the design of Study Two and the testing procedures employed (**section 4.2**). The methods used to codify and analyse the data collected are also explained (**section 4.3**) together with the ethical dimensions which need to be considered in research of this type (**sections 4.4**). The results of the statistical analysis for Study Two are presented in **section 4.5**. Both descriptive and inferential statistics are presented here.

4.2 Research Design

	DATE	EVENT	PURPOSE
Recruitment Phase	May 2012	Contact with administrators for participant schools	Teacher identification of top/bottom ranking pupils from targeted schools
Data Collection	May 2012 – July 2012	Data collection (x 2)	Testing/re-testing participants

Table 4.1: An outline of the research design

4.2.1 Sampling method

Two schools, both situated in the North area of Riyadh, were chosen, one public and one private, in order to target children from different socioeconomic backgrounds. The reason for selecting this area was that Study One revealed it had recorded the greatest spread of performances of all four areas previously sampled and in addition, children there had displayed the lowest levels of attainment. It was decided therefore that the spread of results provided by this sample would provide the best foundation for Study Two as it would ensure that a broad range of abilities would be represented. The intention was also to concentrate on the extremes within this population by focusing on the highest and lowest performers in the sample as identified by their teachers in the hope that this might enable further investigation of the impact of SES and more clearly reveal the statistical significance of variables against teacher ratings. All participants also had to meet the inclusion criteria specified in Study One (**section 3.6**) and, as previously, this was verified before testing commenced.

4.2.2 Consent

As with Study One, before beginning this investigation, informed consent was sought from all those identified as key stakeholders following the process outlined in **section 3.5** and using slightly adapted versions of the original letters, information sheets and consent forms (Appendices B-D).

4.2.3 Participants

The key characteristics of the participants of this additional component are summarized in **Table 4.2**. A total of 60 participants were tested, half from public schools and half from private schools. Participants were also divided equally by age group: 20 in the youngest group, 20 in the intermediate group and 20 in the oldest group (see **Table 4.2**). Since Study One had revealed that gender was not a significant variable in children's performance, this was not factored into later calculations and teachers were asked to identify the five most able and the five least able performers in reading ability from each age group amongst the children, regardless of gender. This judgement was based on teachers' classroom observations and the most recent monthly evaluation they had carried out.

Type of School	Number of Children	4.0-4.11 (*n*= 20)		5.0-5.11 (*n*=20)		6;0-7;0 (*n*= 20)	
		Top	Bottom	Top	Bottom	Top	Bottom
Public	30	5 (4M 1F)	5 (2M 3F)	5 (0M 5F)	5 (0M 5F)	5 (0M 5F)	5 (0M 5F)
Private	30	5 (3M 2F)	5 (2M 3F)	5 (3M 2F)	5 (1M 4F)	5 (0M 5F)	5 (4M 1F)
Totals	**60**	**10**	**10**	**10**	**10**	**10**	**10**

Table 4.2: Study Two participant profile: school type, age and teacher-evaluated reading ability (Top/Bottom).

4.2.4 Procedures

The same procedures were followed for the testing of RRS using the battery developed for Study One, as detailed above. However, the RAN task was only administered to the oldest group since results from Study One suggested it had proved too difficult for the youngest and intermediate age groups. The tasks were administered in the same order in both the first and the second session to ensure that conditions remained as consistent as possible. When carrying out the testing, the researcher was unaware of the teacher rating of the children in order to avoid bias in delivering or scoring assessments. In Study Two the same sample of children was re-tested within 10 days of the first test for test-retest reliability checking.

4.2.5 Measuring children's reading-related abilities

As noted above, an additional measure was required in Study Two for identifying children's level of performance in reading-related tasks in the classroom environment both for the purposes of ranking participants in relation to their peers and also for investigating the degree of correlation between the results obtained from RRS test battery and children's general reading-related abilities. Since there was no existing standardized Arabic literacy scale suitable for use with the age groups involved in this study, it was decided to make use of teacher ratings which form part of the assessment process in the Saudi system.

4.2.5.1 <u>Assessment for kindergarten students</u>

The youngest (4;0-4;11) and intermediate (5;0-5;11) groups had not yet reached the stage of reading independently, so it was decided to take a measure of their reading-related abilities, as assessed by their kindergarten teachers. The teacher was asked to provide a summative assessment for each of the participants which could be used to gauge their general reading-related abilities. At kindergarten level, children are assessed at regular intervals in class on several aspects of their language abilities. Four types of language assessments are routinely carried out with these age groups:

1. Letter knowledge: The child is presented with a single letter on a flashcard and asked to give the name of a letter or pronounce the sound that it makes

2. Letter copying: The child is asked to copy a selection of letters of the alphabet from a printed chart

3. Decoding task The child is asked to decode some simple handwritten words for example: /ba:b/ /بـــا ب/ /da:r/ /دار/. The word recognition tasks are based on words made up of letters in isolation shapes only.

4. Word re-call: The teacher asks the child to listen to a phoneme and to say a word that begins with the same sound. The teacher asks the child "*Can you give me an example of a word beginning with the sound* /says phoneme/?" Items are chosen from a list of letters and words which children have already encountered in class. In each case, the selection of test items is made by the teacher and may differ for each child but each child is assessed on a fixed number of these.

In public schools, the teacher records the results of all these in-class assessments undertaken throughout the school year using a standard Ministry of Education chart which lists all the assessments undertaken, and also indicates the child's level of performance in each task, using the scale: 1 = successful, 2 = partially successful or 3 = unsuccessful. In addition to the language-specific tasks outlined above, this chart is used to record all the tasks that children are required to complete across all areas of the curriculum. Using this information, teachers were asked to identify the top five and bottom five performers in the group who would take part in the study.

Each teacher was also asked to provide the scores for the latest set of summative assessments related specifically to language abilities. For the purposes of this study the scores for the language assessments one, three and four mentioned above were used as a measure of general language abilities and were used to investigate the possible correlation between students' performance in the RRS test battery and their general reading-related abilities. Teachers in private kindergarten were asked to employ a similar system to rate and identify children's performance.

4.2.5.2 Assessment for first grade students

In the first grade (children aged from six to seven years old), students are taught reading and writing skills formally and their reading ability is assessed by a summative assessment at four points over the course of the academic year. The teacher evaluates the child's proficiency in various skills. For the purposes of this study, three key areas were of interest:

1. The ability to decode letters attached to short and long vowels at the syllable level.
2. The ability to read simple words derived from letters which the children have already learned, and
3. The ability to read short, simple sentences.

Since Study Two took place after the third assessment period of the year, the teacher's summative assessments of all of these skill areas was used to identify the top five and bottom five performers in the group who would take part in the study.

This summative assessment used involved the following components.

1. Children are asked to read words that include short vowels. They then move on to simple words including long vowels. This move from short to long vowels represents an incremental transition from shallow to deep orthography in Arabic.
2. Children are presented with words marked with various diacritics and asked to read these words aloud. This assessment component includes 20 items (nouns and verbs) consisting of two or three syllables which contain commonly occurring Arabic graphemes. These have been

selected on the basis of the frequency of words in the Arabic curriculum. This also entails children being able to recognise Arabic letters presented in their initial, middle, final and isolated shapes. Word difficulty within this assessment component is graduated by reducing the marking of short vowels (devocalization).

3. Finally, children are asked to read sentences, the difficulty of which is gradually increased by lengthening the numbers of words in the sentence. In these reading tests, children are scored on the ability to pronounce consonants and vowels correctly.

A sample of the Ministry of Education reading test used in all public schools was provided to private school staff and they also used this test to evaluate students. This helped to provide a greater degree of consistency in the form of assessment and to improve the comparability of outcomes.

Each teacher was asked to provide the scores based on the summative assessment for reading ability which students had taken at the end of the third assessment period. For the purposes of this study the scores for this reading assessment were used as a measure of reading ability and were used to investigate the possible correlation between this score and students' performance in the RRS test battery.

4.3 Data Coding and Analysis

Similar data coding and statistical analysis to that used in Study One (see **section 3.12**) was performed on all the data collected from the validation process carried out in Study Two. All data collected from RRS tasks, questionnaires and teachers were also anonymized as previously described for Study One (**section 3.13**). The test/re-test reliability of the RRS battery was calculated using an intra-class correlation coefficient and the results of these analyses are presented below (**section 4.5.2**).

4.4 Ethical Considerations

As with Study One, all the necessary ethical clearance was received and similar processes were followed with regard to obtaining informed consent and data protection protocols (**section 3.13**).

4.5 Results of Statistical Analysis

4.5.1 Participant characteristics

A total of 60 participants were tested: 30 (50%) attended public schools and 30 (50%) attended private schools. Participants were also divided equally for both public and private sector by age group: 20 (33.3%) in the youngest group, 20 (33.3%) in the intermediate group and 20 (33.3%) in the oldest group (see **Table 4.3**). Teachers were asked to identify the five most able and the five least able performers in reading ability from each age group among the children (16.7%). This judgement was based on their classroom observations and the most recent monthly evaluation they had carried out.

		AGE GROUPS					
Type of School	**Number of Children**	**4;0-4;11** (*n*= 20)		**5;0-5;11** (*n*=20)		**6;0-7;0** (*n*= 20)	
		Top	**Bottom**	**Top**	**Bottom**	**Top**	**Bottom**
Public	30	5	5	5	5	5	5
Private	30	5	5	5	5	5	5
Totals	**60**	**10**	**10**	**10**	**10**	**10**	**10**

Table 4.3: Participant profile: school type, age and teacher-evaluated reading ability (top/bottom).

4.5.2 Reliability of the five PA tasks and the LK task

All data collected using the RRS test battery in Study Two were used to calculate the test/re-test reliability of the battery using an intra-class correlation coefficient (ICC). In this statistical calculation, any tasks that are found to have Cronbach's alpha values above 0.7 can be considered to be reliable (Kline, 1999). The test re-test reliability was determined by testing all children on all measures with a one- to two-week interval between initial and second testing.

		INTRA-CLASS CORRELATION COEFFICIENT					
	Tasks	**Rhyme Awareness**	**Syllable Segmentation**	**Alliteration Awareness**	**Phoneme Isolation**	**Blending**	**LK**
Age group	4;0-4;11	.92	.91	.92	.94	.96	.99
	5;0-5;11	.82	.89	.89	.96	.96	.99
	6;0-7;0	.84	.90	.95	.96	.99	.99

Table 4.4: Reliability of five PA tasks and LK tasks using ICC

Table 4. 4 shows that all the Cronbach's alpha values obtained for the five PA tasks and the LK task were greater than 0.8 across all age groups, which according to Kline (1999) suggests an acceptable level of consistency for these tasks in the current study, with the highest reliability values for alpha (above 0.9) being observed in the youngest age group. Hence, it can be concluded that the tests were reliable and robust and the items in each of these tasks were reliably measuring the same skills in the children assessed.

4.5.3 Reliability of the RAN Tasks

In the case of the RAN tasks, since Study One revealed that these were only suitable for the oldest group, the figures here refer only to this group. As **Table 4.5** shows, Cronbach's alpha coefficients for the RAN tasks were all greater than 0.9, with many values close to 1, suggesting that the items have a high internal consistency.

INTRA-CLASS CORRELATION COEFFICIENT							
Colour Naming A (*n*=20)		Colour Naming B (*n*=20)		Object Naming A (*n*=20)		Object Naming B (*n*=20)	
Time	Error	Time	Error	Time	Error	Time	Error
.96	.98	.98	.96	.97	.91	.99	.94
Letter Naming A (*n*=13)		Letter Naming B (n=13)		Digit Naming A (*n*=14)		Digit Naming B (*n*=14)	
Time	Error	Time	Error	Time	Error	Time	Error
.98	.94	.98	.97	.98	.99	.98	.97

Table 4.5: Reliability of the RAN tasks using ICC and Cronbach's alpha coefficient (oldest group).

4.5.4 Descriptive Analysis

The means and standard deviations for composite PA measure, Blending, LK and RAN tasks were again extracted and this time, the children's performance in each test was segregated by age group.

Table 4.6 presents the mean performance of the top/bottom ranking groups for the composite PA measure, Blending and LK tasks, by age group. In the top ranking group means show that by age those in the intermediate group (5;0-5;11) performed better than the youngest (4;0-4;11) whilst the oldest group (6.0-6.11) performed better than both of the other age groups on the composite PA measure and LK. In the Blending task, the youngest group performed better than the intermediate group amongst the top ranking children but once again the oldest group outperformed both the other groups on this task.

Amongst the bottom-ranking students, the youngest group performed better than the other two age groups in composite PA. The intermediate group performed slightly better than their older counterparts. In the Blending task the bottom-ranking students in the intermediate age group performed better than their younger and older counterparts, with children in both these age groups performing similarly. In LK tasks, the intermediate group again outperformed the youngest and the performances of those in the oldest group were better than those in both the other groups for this task.

In the case of the RAN tasks, as previously noted, Study One revealed that these were only suitable for the oldest group; hence the figures here refer only to this group. The values of the means, standard deviations and number of valid cases (indicated by N) are illustrated in **Table 4.7**. Results show that children in the top ranking group performed faster on all the RAN tasks. The bottom ranking group produced the smallest number of valid cases across all RAN tasks. Colour Naming produced the highest number of valid cases.

	4;0-4;11				5;0-5;11				6;0-6;11			
	RANKING				RANKING				RANKING			
	Top *(n=10)*		**Bottom** *(n=10)*		**Top** *(n=10)*		**Bottom** *(n=10)*		**Top** *(n=10)*		**Bottom** *(n=10)*	
TASKS	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Composite PA measure	23.20	3.08	17.50	2.80	25.20	3.46	13.60	3.17	28.70	1.64	13.50	2.47
Blending	7.30	2.58	5.30	1.65	6.80	2.20	6.20	2.31	10.10	3.34	5.30	2.16
LK	13.20	7.38	3.50	1.51	14.50	10.34	7.30	8.91	24.90	2.99	11.80	7.73

Table 4.6: Mean and SD for the Composite PA measure, Blending and LK tasks by reading ability ranking for all age groups.

	TOP RANKING			BOTTOM RANKING		
TASKS[1]	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**
Colour Naming	36.33	4.89	9	44.13	2.10	4
Object Naming	40.25	7.39	6	47.25	1.71	6
Letter Naming	34.62	7.79	8	43.67	6.42	3
Digit Naming	35.75	6.75	6	43.00	6.76	3

Table 4.7: Mean and standard deviation for all RAN tasks by reading ability ranking (top/bottom) for the oldest group only.

[1] *Values are given in seconds taken to complete the task*

4.5.5 Inferential analysis

Data was analysed using a univariate ANOVA to investigate the effect of between-group differences (i.e. ranking top/bottom) on the composite PA measure, Blending and LK tasks, by age group.

Results in **Table 4.8** showed a significant effect of age on the LK task (F (2.57) = 10.40, $p<0.05$). A significant effect was also found for the teacher rating of reading ability ranking on the composite PA measure (F (1.57=219.7, $p<0.05$), Blending (F (1.57)=15.48, $p<0.05$) and LK tasks (F(1.57)=28.88, $p<0.05$). These effects were large, with 80%, 22% and 35% of the overall variation in performance on the composite PA measure, Blending and LK tasks, respectively, attributable to the influence of the reading ability ranking. Interaction effect was also significant for the composite PA measure and Blending task but not for the LK task.

		COMPOSITE PA MEASURE	BLENDING	LK
Age	**F**	1.8	1.95	10.40*
	P	.17	.15	.00
	Partial Eta Squared	.06	.08	.28
Reading Ability Ranking	**F**	219.70*	15.48*	28.88*
	P	.00	.00	.00
	Partial Eta Squared	.80	.22	.35
Interaction	**F**	14.35*	3.88*	.84
	P	.00	.03	.45
	Partial Eta Squared	.35	.13	.03

Table 4.8: F tests for the effect of age, reading ability ranking (top/bottom) and the interaction effect, with partial eta squared values

**F-value significant at 0.05 level (p < 0.05) **F-value significant at 0.01 level (p < 0.01) ***F-value significant at 0.001 level (p < 0.001)*

A graphical display of the performance of both top and bottom ranking groups on the Composite PA, Blending and LK tasks is provided in **Figure 4.1, Figure 4.2 and Figure 4.2**: *Performance in the Blending task by reading ability ranking (top/bottom)*

3 respectively.

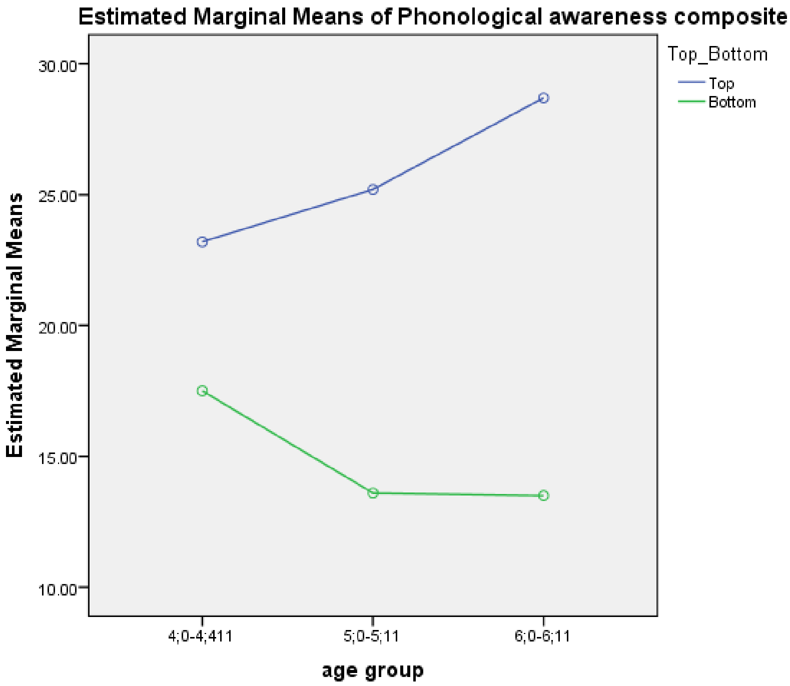


Figure 4.1: Performance on the composite PA measure by reading ability ranking (top/bottom).

Figure 4.1 reveals that the composite PA measure performance for the top-ranking group children increases as they get older and their abilities in these areas may develop as a result of increasingly formal exposure to reading instruction. Among the top-ranking children, the lowest scores came from the youngest whilst the oldest group (6;0-7;0) achieved the highest. However, the pattern for those in the bottom-ranking group was completely different since the mean scores of the bottom-ranking group for the intermediate and oldest group were lower than those of the youngest. In the case of their bottom-ranking counterparts, the highest scores for the PA composite measure came from the youngest group (4;0-4;11), and there was no difference in performance in either the intermediate or oldest groups, indicating that selection may not have been as effective in their case.

Figure 4.2 shows that the children in the top-ranking group achieved the highest scores in the Blending task. Conversely, those in the bottom-ranking group obtained the lowest scores in this task across all age groups. In the top-ranking group, the oldest group achieved the highest scores, followed by the youngest group and the intermediate group. This was not the case in the bottom-ranking group where the intermediate group achieved the highest mean scores. Conversely the youngest and the oldest groups both performed poorly in comparison with the others.

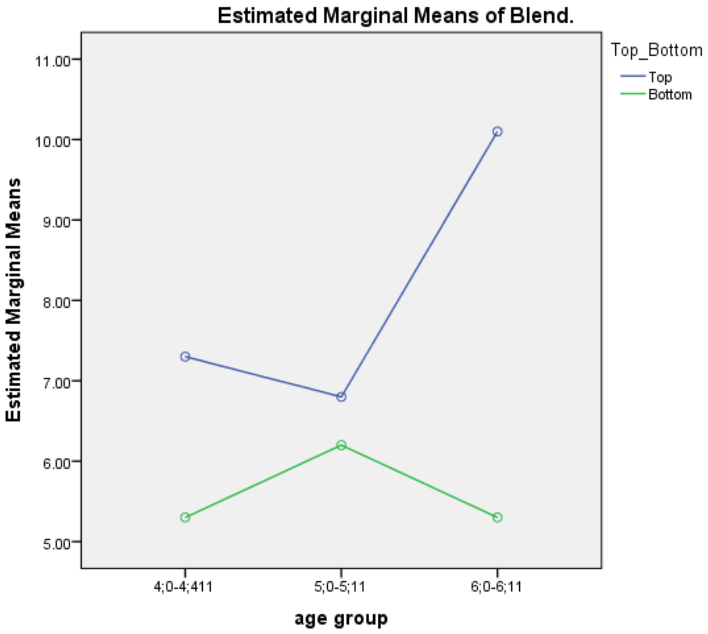


Figure 4.2: Performance in the Blending task by reading ability ranking (top/bottom)

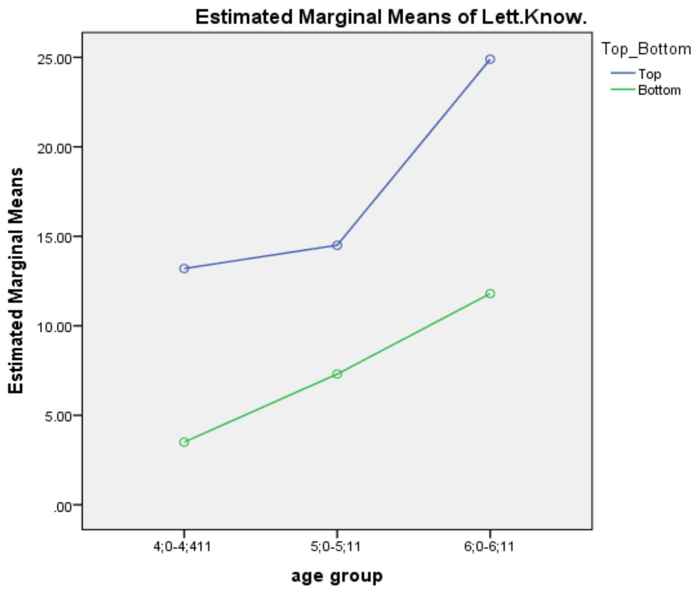


Figure 4.3: Performance in the LK task by reading ability ranking (top/bottom)

Figure 4.2: *Performance in the Blending task by reading ability ranking (top/bottom)*

3 shows that children in the top-ranking groups achieved the highest scores in the LK task. Conversely, children in the bottom-ranking group obtained the lowest scores in the LK task across all age groups. In both the top-ranking and bottom-ranking groups the younger children performed worse than the older children, indicating that there was both an age effect and an ability effect.

In the case of the RAN tasks, given the reduced sample size, data is not statistically valid for predictive purposes.

	Colour Naming	Object Naming	Letter Naming	Digit Naming
F	3.05	.82	1.32	.05
P	.11	.39	.28	.82
T	-3.01	-2.02	-1.78	-1.52
Df	11	10	9	7
Sig. (2-tailed)	.01	.07	.11	.17

Table 4.9: ANOVA results for performance in RAN tasks by reading ability ranking (top/bottom) (oldest group only).

*****F-value significant at 0.001 level ($p < 0.001$) ** F-value significant at 0.05 level ($p < 0.05$)*

In **Table 4.9** the significance (2-tailed) value is greater than .05 for three of the RAN tasks (object naming, letter naming, and digit naming), indicating there is no significant difference between the two groups (top and bottom-ranking) for these constructs. However, the significance (2-tailed) value is smaller than .05 in the case of colour naming, meaning there is a significant difference between the top and bottom-ranking groups in this task, in favour of the top-ranking group.

4.6 Conclusion

In order to fully address the research questions, this investigation was divided into two parts. Study One required the development of a test battery that could be used to identify the performance of Arabic-speaking children in a number of

RRS using PA, LK and RAN tasks. It was complemented by a questionnaire to gather information about their socioeconomic background. Both instruments were used to investigate the correlation between performance in the test battery and a range of factors hypothesized to influence levels of RRS in young Arabic-speaking children. Study Two was designed to investigate the reliability of the RRS test battery and its usefulness as a predictive tool of teacher rating of reading/reading ability in young Arabic-speaking children.

Having presented the results of the statistical analysis of data gathered from Studies One and Two, these findings will be discussed in detail in Chapter Five and their relevance to the research questions presented at the start of this thesis and the hypotheses stated in the Literature Review will be assessed. Consideration will be given to the extent to which these findings from a specific context shed light more generally on Frith's model and on the factors that can impact on children's development of RRS.

5 CHAPTER FIVE: DISCUSSION

5.1 Introduction

As noted previously, research suggests that PA, LK and RAN are good predictors of the development of reading ability for large numbers of children across languages (e.g. Georgiou, Torppa, Manolitsis, Lyytinen, and Parrila, 2012; Georgiou, Parrila, and Liao, 2008; Caravolas et al., 2005; Caravolas, et al.,2012). There is little research available about the predictors of reading in Arabic (Saiegh-Haddad, 2005; Taibah and Haynes, 2011; Tibi and Kirby, 2019). Consequently, there is lack of evidence to guide remedial programs which would help to address problems with literacy in Saudi children, and there is no systematic approach taken to identify those children who are failing to make progress with their reading skills. So, the aim of this research was to develop a comprehensive battery of age-sensitive tests which could be used to assess reading-related skills (RRS) in Arabic-speaking children aged 4;0-7;0 years with a view to producing a solid foundation for the future development of a comprehensive standardized RRS assessment battery designed for use in the Saudi context.

As a clinician in Saudi Arabia, I am aware of the urgent need for this type of assessment that could be used for early diagnosis of potential literacy problems. Part of my professional role is to identify children who are at risk of developing persistent literacy (reading/writing) difficulties so that they can be offered appropriate special educational support with the collaboration of the teacher. A diagnostic test for evaluating RRS development in children who not appear to be making any progress in their early reading and writing experiences does not currently exist for Arabic. So, developing such an assessment for children across the age range 4;0-7;0 years will help to produce a profile of a child’s strengths and weaknesses in PA and other RRS in order to devise appropriate intervention goals. The earlier children can be tested, the more chance there is that appropriate targeted interventions can be used to help prevent later problems.[18]

As indicated in Chapter Two, various studies have also established that the development of RRS can be shaped and affected by factors including gender, age and SES (Gillon, 2012; McDowell et al. 2007). The battery of tests that was developed for this research was used to obtain normative data, and to explore the influence of these factors on the development of RRS in a sample of typically developing Arabic-speaking children in Riyadh, aged 4;0-7;0 years. In addition, Frith's model (1995) suggests that it is necessary to look not only at these areas but also at other environmentally determined factors such as SES, home literacy and socio-emotional problems, which may influence RRS. In this thesis, the decision was taken to examine also if there is any evidence that the emphasis placed on the study of Qur'anic recitation skills (*tajwīd*) from a very early age in the Saudi educational system influences the development of RRS by looking at the impact of exposure to additional tuition in this area.

Hypotheses were then formulated that related to gender, age, socioeconomic status, and levels of exposure to *tajwid* tuition and these were tested using the assessment battery. The extent to which the RRS assessment battery functions in the Arabic-speaking context as a reliable predictor of teacher rating of reading ability was tested in the second study. The results obtained from these two studies were presented in Chapters Three and Four respectively and in this chapter the implications of these results are discussed in detail, both in relation specifically to the hypotheses and also in broader terms by comparing and contrasting them with previous research findings.

This chapter was intended to provide an analytical reflection on the data which was presented in the previous chapters and begins by considering the five hypotheses H1-5 which the two studies conducted in this research were designed to test. It determines the extent to which these can be said to be supported or not on the basis of the findings. At the same time, it compares and contrasts findings from this research conducted in an Arabic-speaking Saudi

[18] For a discussion of appropriate targeted interventions relating to PA see Chapter Seven "Phonological Awareness intervention: guiding principles" and Chapter Eight "Phonological Awareness intervention: instructional Framework" in Gillon (2012).

context with evidence from other cultural and linguistic settings, exploring how the results from this research appear to complement or contradict previous studies in this area. In order to facilitate reference to statistical evidence which was presented in previous chapters, cross referencing is used throughout.

5.2 Gender and PA, LK and RAN

This section discusses the findings in relation to the following hypothesis:

H1: Gender will affect children's performance in RRS, with girls gaining significantly higher scores in PA, LK and RAN tasks.

The first issue which was addressed was whether gender affects children's performance on all RRS measures (PA, LK and RAN tasks). Gender was considered before age effects because if significant gender effects are found, there could be an argument for creating separate norms for boys and girls. The composite PA results for participants' performance on RRS tasks according to gender were not significant but show slightly higher means for the girls across all measures. When these same results were separated by age group, the pattern was somewhat more varied, with boys in the youngest group performing slightly better than the girls in two of the PA tasks (rhyme awareness and syllable segmentation) and two of the RAN tasks (Colour Naming and Digit Naming). In the blending task, the boys in the youngest group achieved the same mean as their female counterparts. In the intermediate group, boys completed the RAN letter naming task with a lower mean time in seconds than the girls whilst boys in the oldest group produced better mean scores than their female peers for Rhyme Awareness. None of the participants in the youngest group was able to produce valid test results for the RAN letter naming task. (More is said about the difficulties which children experienced with the RAN tasks later in **section 5.3.4**) However, the t-tests for gender differences on each task taken by age group were not statistically significant.

Initially, the overall mean scores for two tasks appeared to differentiate between males and females at a statistically significant level. These were phoneme isolation ($p < 0.05$) and RAN (Object Task) ($p \leq 0.01$) with females performing better than males on both these measures (see **Table 3.16**). However, once a Bonferroni correction for multiple comparisons was applied, neither of these results was significant. No significantly higher scores were achieved by girls in any of the PA tasks, LK or RAN and therefore H1 was largely not met.

Findings from studies focusing on gender in Western countries have produced variable findings. Schaefer et al.'s (2009) study in German speaking children concluded that differences in the PA performance between male and female were small and not statistically significant. In contrast, a recent study by Mohamed et al. (2019) in Arabic pre-schoolers showed that girls had higher values in most PA tasks than boys, significant results was observed only on rhyme discrimination, phoneme segmentation and phoneme isolation task (small to medium effect size). It is interesting to note, therefore, that this study did not find gender effects on RRS and performance levels for girls and boys were similar despite the gender segregation that operates throughout the Saudi education system (even at pre-school level). As noted in the literature review, since the 1960s, gender factor was often affect literacy development and that girls generally perform better than boys in terms of verbal and linguistic functions (Maccoby and Jacklin 1974; Halpern, 1986). More specifically, researchers have found that girls performed better than boys on phonemic awareness tasks (Moura et al., 2009). Also, girls outperformed boys on phoneme segmentation fluency; correct letter sound fluency and whole word reading fluency of non-words (Chipere, 2014) and rapid naming (Burman et al., 2008). Although Lundberg et al.'s (2012) study showed a clear gender effect, with girls achieving higher scores for PA at initial testing and greater rates of improvement after training, they highlight the inconsistency of the results of studies on gender. They point out that the fairly limited sample sizes used in previous empirical research on PA and gender difference made it difficult to reveal significant statistical differences and consequently they recruited a sample of over 2000 participants.

As detailed above, numerous small differences between the mean scores of girls and boys, consistently favouring the former, were noted across all measures in the present study. However, these were not found to be statistically significant. In support of findings, no gender differences have been observed in some studies spanning different language contexts, for example, in Schaefer et al.'s study (2009) in German, and Krishna Priya et al.'s (2018) study of phonological skills of Malayalam language. However, Krishna Priya et al.'s participants were grouped based on their socio economic status, and the study included only syllable level tasks. Although the number of participants recruited for the present study (*n*=384; 192 boys and 192 girls) was substantially larger than the cohorts used in some previous studies (Burt et al. (1999) (*n*=57; 28 boys and 29 girls) and Schaefer et al. (2009) (*n*=55; 30 boys and 25 girls), the sample for the present study was smaller than that tested by Lundberg et al. (2012). In terms of the age range these studies overlapped with our research. Participants in Schaefer et al. (2009) were aged 4;0-6;0 while for Lundberg et al. (2012) was 5.75-6.67.

Since Study One demonstrated that gender was not a significant variable in children's performance, this result did not support the need for separate analyses in respect of the norms for boys and girls; therefore, it was not factored into calculations for Study Two which focused on the extremes of performance, testing the most and least able children in teacher-rated reading ability from each age group, regardless of gender. (The gender breakdown for participants in Study Two is shown in **Table 4.2**. It would be interesting to investigate gender in a future study with a larger sample of children at the extremes of the ability range, comparing this with findings for the full ability range sampled in Study One.

5.3 Age and RRS

This section discusses the findings in relation to the following hypothesis:

H2: Age will affect children's performance in RRS, with older age groups gaining significantly higher scores in PA, LK and RAN tasks than younger age groups.

This hypothesis was intended to address the effect of age on RRS as measured by children's performance in PA, LK and RAN tasks. In the manual used for Frederickson et al.'s (1997) test the norms for ages are presented at six-monthly intervals. However in this study, using 6-month bands didn't show a significant impact on Study One. The reason why these differences were not in evidence at six-monthly intervals in Study One is unclear. This may simply have been a result of the age profile of the sample which was used unlike study of

Frederickson et al. (1997). Significant age-related differences in performance were only revealed for 12-month age bands (**section 3.14.3.2**) so these were used rather than the six-month age cohorts.

Differences in outcomes across the three one-year age groups revealed a clear age effect on PA and LK tasks. Overall, when statistical analysis was run using one-year bands, the test was found to be age-sensitive, an important finding in relation to the usefulness of the test battery because tasks designed to measure children's development need to be age-sensitive to show the developmental changes in their performance over time and this emerged clearly.

5.3.1 Age and PA Tasks

This sub-section discusses age effects in task performance for the four PA tasks developed for this study (Rhyme awareness, syllable segmentation, alliteration awareness, phoneme isolation), and composite PA as a combined measure of these four PA tasks. Then the focus shifts to blending (taken from Taibah, 2006), LK and RAN, in turn.

The children's performance in all the PA tasks improved with age, showing a clear age effect and signalling the developmental progression of PA with age as hypothesized, and **Table 3.24** showed the location and strength of these effects. Nevertheless, there were important differences in the distribution of scores across subtests. The results of this study support findings of Mohamed and his associates (2019) on PA developmental trajectories in pre-schoolers. They divided participants into two groups based on their age: KG1and KG2 with mean age range between 3.75 – 4.8 years and 5.5 – 5.9 years respectively. KG2 children outperformed KG1 children and a significant difference was observed in almost all the tasks except blending (syllables and non-words), and non-word repetition (Mohamed et al., 2019). Similarly, Al-Sulaihim and Theo (2017) compared PA skills with reading ability over time in Arabic speaking children. They measured the PA skills at the beginning of grade one and at the end of the school year. Al-Sulaihim and Theo found that children improved in PA task scores over time. They suggested that literacy training, formally introduced in the first grade, has a positive effect on PA skills. Similar studies by McDowell et al. (2007) and Duranovic et al. (2012) also concluded that age is one of the important predictors of children's PA. McDowell et al. examined children (age range = 2-5 years). Results showed that age contributed unique variance to the prediction of PA.

PA task acquisition rate varies and rhyming and alliteration are early reading skills while tasks required for sound manipulation like phoneme deletion, substitution and reversal develop at a later stage (Chafouleas et al., 1997; Mohamed et al., 2019; and Moyle et al., 2013).

For rhyme awareness, pairwise comparison showed no significant difference between the oldest and the intermediate groups, both of which significantly outperformed the youngest group. The positively skewed distribution on this task shows a ceiling effect, often observed when a task is too easy for a target population (see Coolican, 2013); many of the older children scored the maximum or very near to it on this task. The study findings are in the line with the previous studies, such as Al-Sulaihim and Theo (2017) study, which examined the changes in Arabic group scores over time on a rhyme awareness task and found a significant difference was found by the end of the year. Results of this study were also supported by a recent study which showed that rhyming discrimination and production were positively associated with age, KG2 children outperformed KG1 children and a significant difference was observed in both tasks (Mohamed et al., 2019).

For syllable segmentation, results were more evenly spread. Although there were significant differences between the youngest and intermediate groups, and between the intermediate and oldest groups, effect sizes were small in each case. Mohamed et al. (2019) report similar results for syllable segmentation and observed a development trend across grade levels.

Alliteration awareness again showed a fairly even spread, with significant differences between the children in the youngest and the intermediate groups, and between those in the intermediate and the oldest groups. In this case, the effect size between the youngest and the oldest groups was large, showing a wide spread of responses and a larger age-related difference than for the previous tasks. This was again in the line with Sulaihim and Theo (2017) who found children improved in initial sound matching task scores when tested at the end of the school year (T2). Nation and Hulme (1997) found that children Performance at different PA tasks including alliteration sound categorisation improved with age.

Of the four PA tasks, phoneme isolation showed the greatest impact of age, with significance differences between the youngest and the oldest groups, and also between the intermediate and oldest groups (see **Table 3.24**). This finding may be due to the fact that by the time Arabic speaking students reach first grade, they have learned the alphabetic principle and together with auditory discrimination skills this allows them to identify the first sound in a word. This result is consistent with previous Arabic studies, for example, Saeigh-Haddad (2003) investigated initial and final phoneme isolation and diglossic variables in kindergarteners (n=23; mean age=5.9) and first graders (n=42; mean age 7). Saeigh-Haddad concluded that first graders' PA was better than kindergarteners'. Results are also consistent with Mohamed et al. (2019) who found that KG2 (mean age 5.5 – 5.9 years) children had higher scores than KG1 (mean age 3.75 - 4.8 years) children in first, final and medial phoneme isolation. Likewise, Asadi and Abu-Rabia (2019) found developmental changes between K2 and K3 for a phonemic isolation task that were highly significant.

Our results showing developmental progression in PA are consistent with the previous literature (Treiman and Zukowski, 1991, 1996; Kertoy, 2005; Schaefer, 2009). Previous study results also confirm that age-related differences reflect the level of the phonological unit that is the focus of the task, with skills developing from large to smaller linguistic units – PA is a multi-level skill (Anthony and Francis, 2005). Very few young children were able to complete tasks at the phoneme level, whilst a considerable number of older children were successful in completing the small segment tasks (i.e. alliteration awareness, phoneme isolation and blending), showing they had started to focus on smaller units and were attempting to break the words down into phonemes.[19] In contrast, results indicate that the youngest children were performing at the level that would be expected if they were answering by chance. For example, in the

[19] For further information about the nature of these tasks see Chapter Three, sections 3.8.5.3 (Alliteration Awareness), 3.8.5.4 (Phoneme Isolation Task) and 3.8.5.5 (Blending Task).

alliteration awareness task when children receive three options, 33% of responses would be right without any knowledge. For phoneme isolation and blending, the child was required to produce a phoneme (phoneme isolation) or sequence of syllable(s) (blending), so chance did not enter into these tasks.

The findings of this study support Stackhouse and Wells' (1997) contention that different aspects of PA develop at different ages and rates, indicating a developmental hierarchy. The results indicate that some PA tasks are easier than others. Rhyme awareness and syllable segmentation tasks are easy tasks when compared to blending. This is consistent with the pattern of development reported from literature on the development of PA in English-speaking children (Burt et al., 1999; Anthony and Francis, 2005) and in Arabic-speaking children (Tibi, 2010; Abou-Elsaad, Ali, and Eel-Hamid, 2016).

One striking feature of the findings is the indication of significant development in PA abilities when children start school. This has some bearing on the ongoing debate about the nature of the link between PA and literacy, and the cause and effect relationship between skills, i.e. whether PA is an antecedent to or consequence of successful literacy acquisition (Castles and Coltheart, 2004; Hulme et al., 2005). There are suggestions that while the shift from implicit to explicit PA (McGuinness, 2005) can occur spontaneously with age (Burt et al., 1999; Lonigan et al., 1998), reading training in formal education may contribute to this development (Ehri et al., 2001). There is evidence that the preschool children (the youngest age group) had also developed an awareness of the units of rime and onset. The mean scores for these tasks indicate that the children in the youngest and intermediate groups were more aware of rhyme than phonemes (as evidenced by alliteration and phoneme isolation and blending tasks). In the battery developed for this research, rhyme awareness and syllable segmentation can be considered implicit. Alliteration awareness is situated between implicit to explicit while phoneme isolation and blending are explicit tasks. These are key measurements of RRS.

In Saudi Arabia, formal reading instruction (including explicit PA training) normally starts at the age of six when children generally enter the first grade of elementary school; however, some private kindergartens provide this form of instruction for younger children. Since participants' prior exposure to formal reading instruction was not established and controlled for in this study, it is not possible to draw conclusions about causation, but there is a very clear indication that those tasks that require more explicit knowledge at the phoneme level develop when formal literacy training begins at school, rather than before (see **2.3.1.1** for further information about PA and developmental progression).

5.3.2 Age and Blending

For the blending task no significant difference was noted between the youngest and the intermediate groups (p>0.05), although the task did produce significant differences between the youngest and oldest groups (p<0.001) with a large effect size (see **Table 3.24**). Similarly, Taibah and Haynes (2011) studied Arabic phonological skills and its contribution to reading ability. They included children ranged in age from 6.33 to 9.18 years and found higher scores in blending task in third graders than in early grade children. The reason for insignificant results in the intermediate age group may be due to the form of presentation used in this task in that no pictures were provided, so single segments had to be kept in the child's working memory before blending (see Stackhouse et al., 2007). This might account for the task proving easier for those participants who had received more exposure to explicit PA training. In similar to findings of this study, Mohamed et al., 2019 did not find any significant difference in syllable blending and blending non-words between the two age groups (KG1:4-5 years and KG2: 5-6 years). As suggested by Ziegler and Goswami (2005), PA and reading skills develop hand-in-hand and the blending task may require formal and alphabetic reading tuition. For Saudi children this normally starts at the age of six. Thus the findings in this study for the blending task follow the expected pattern, and show it to be developmentally sensitive.

5.3.3 Age and LK

As noted earlier (**section 3.8.6**), children in the LK task were asked to give the name or sound of the displayed letter. Data analysis and results revealed that the largest differences in the children's performance on the LK tasks were found between the youngest (4;0-4;11) and the oldest (6;0-7;0) age groups (see **Table 3.28**). Analysis showed that the LK task had a significant age effect with a medium to large effect size. The oldest group performed better than both the other age groups with large effect sizes. This finding is consistent with previous evidence that LK performance improves as children get older (Carroll et al., 2003). Also in an Arabic context, Al-Sulaihim and Theo (2017) found a significant improvement for LK in children when tested at the beginning and by the end of the academic year.

As with blending the noticeable improvement in the oldest group corresponds with entry into the first grade of elementary schooling and therefore may reflect the impact of formal Phonics instruction. This is consistent with previous indications of an association between PA and LK. Piasta and Wagner (2010) found that among American pre-school children exposure to PA information (albeit without explicit training) was associated with higher LK; Cardoso-Martins et al. (2011) found that when Portuguese-speaking pre-schoolers in Brazil were given explicit PA training they acquired LK more easily than a control group who were not given this training. Similarly, phonological awareness aLK are linked to literacy in Arabic language. Al-Sulaihim and Theo (2017) examined the relationship between PA, LK and reading abilities after formal education was introduced. They included first graders and measured the variables over time. At the end of the academic year, a significant association between LK and reading ability was found while there was no significant relationship at the start of the year.

5.3.4 Age and RAN

In our study, RAN tasks included series of colours, objects, letters and digits. There were no valid responses for letter and few responses (n=6) for digit naming task in the youngest group (4;0-4;11),) which proved that these tasks were too difficult for them. Valid responses for the other RAN tasks were also fewer in this group (see **Table 3.29**).

In all RAN tasks, children performed better with increasing age and a significant difference was observed between the two older groups (see T**able 3.34**). Researchers have argued that both RAN and PA are good and early predictors of children’s reading ability even in pre-schoolers (Kirby et al., 2003; Rathvon, 2004; Flora 2011; Taub and Szente, 2012; Norton and Wolf, 2012). RAN is also considered to be one of the best predictors of reading fluency across different orthographies (Georgiou et al., 2008a, Landeral and Wimmer, 2008) and has been used in a variety of language settings, including Arabic (Taibah and Haynes, 2011; Tibi and Kirby, (2019). Saiegh-Haddad (2005) found that letter-sound- naming speed was a good predictor of pseudoword reading fluency.

Since kindergarten is not obligatory in Saudi education, some children in the oldest group in this study may have been better trained than others in recognising colours, objects, letters and digits, but this was not controlled for, and may have been another important source of influence.

As explained in the methodology chapter (**section 3.8.7**), RAN is a lexical retrieval task. These results reflect a key problem with the suitability of the four RAN tests for the intermediate and youngest groups. The test battery lasted 30 to 45 minutes and the RAN tasks were delivered last to the participants. Following recommended practice, all participants were given a mini-training session to ensure they understood the task. By that stage, the youngest children were noticeably tiring and whilst all possible attempts were made to ensure that they were assisted to achieve their optimal performance, this combination of the attention span required and the difficulty of the full battery of tests proved too challenging for the younger participants.

As Lane and Bundy note “it is challenging to administer standardized tests to preschool-age children because their attention spans tend to be quite limited, as is their ability to follow directions, and to persist when challenged (2012: 78). Although in theory two separate sessions of tests might have helped to solve this problem, it would have been very difficult to obtain permission from schools to recruit children and administer test battery in two sessions.

Researchers have used specific RAN tasks in specific age groups. For example, Taub and Szente (2012) included six- to ten-year-olds, whereas Asadi and Khateb (2017) studied a total of 458 Arabic children of age range between 6.99 and 7.96 years to assess letter and object RAN ability. In Assad and Eviatar (2014), children were from three grades. First grade (n=31; mean age= 7.02 years), third grade (n=16; mean age= 8.94 years) and fifth grade (n=17;

mean age=10.88 years) and were tested on letter naming task. Tibi and Kirby (2019) tested children (n=201; mean age =8.1 years) on digit and object naming tasks. The youngest participants tested in Arabic studies, for example in Taibah and Haynes' (2011), were 6.33 years old, and children were tested on the alpha-numeric tasks as these are more appropriate for older children. The results of the present study show that older children (6.0-7.0) performed better in RAN tasks than younger children and are in line with the previous studies (Asaad and Eviatar, 2014). Asaad and Eviatar (2014) reported that naming speed has a stronger effect of age for first and fifth graders than third graders, arguing that the first-grade effect was related to letter learning, whereas the fifth-grade effect was related to orthographic pattern learning.

Another point which merits discussion here is the fact that participants in the two younger age groups seemed to find it much harder to complete the Letter Naming component of the RAN tasks than the LK component of the battery, even though both activities involved an element of letter recognition. It is interesting here to compare and contrast the results for the LK task in the battery and those for the Letter Naming element of the RAN tasks and to consider firstly, potential explanations for this difference and secondly, what, if anything, these differences and similarities can potentially tell us about their predictiveness regarding reading ability.

In the case of the LK task, children were required to identify the name or sound of all 28 letters of the Arabic alphabet, presented as isolates, one after another, whilst in the Letter Naming element of the RAN task, only six individual Arabic letters were presented in series, as follows:

ن /no:n/ ل /la:m/ ي /ya:ʔ/ م /me:m/ ص /sa:d/ ك /ka:f/

See **sections 3.8.6 and 3.8.7** for further details.

Results for LK and RAN Letter Naming in this study support previous findings that have suggested that the task involving single-item naming of letters (LK) is very different to the added demands imposed by serial naming of letters in RAN (Norton and Wolf, 2012: 436). Across several studies, single-item and serial naming have been found to be only moderately correlated (Logan et al., 2009) and according to Meyer et al. (1998), the added demands associated with the continuous, serial nature of RAN letter naming make it a better predictor of reading than LK. Moreover, Logan et al. (2009) found that single-item naming and serial naming speed grew at different rates as children got older, adding weight to the claim that RAN should not be considered as simply another version of single-item naming, as it involves other processes.

These results confirm Hypothesis 2 that the age of the children being tested affected their performance in RRS tasks, showing that these tasks are sensitive to age. However, as is the case for findings in other languages, this study demonstrates that some tasks are more age-sensitive than others. There is little evidence that the more transparent script of Arabic makes a particular difference. Children's attendance at kindergarten may have been an important factor in their performance since it increases exposure to PA and LK, affecting their skills at isolating and blending phonemes in particular, but this was not controlled for. However, there is a strong indication that in the case of the phoneme-level tasks being exposed to formal literacy teaching at school makes a real difference to performance levels.

Essential PA skills and sound structure of language learning develops during preschool age. At the age of six children start to take reading instructions Therefore, the battery is useful in showing age effects and also supports previous findings that recommended the use of age-differentiated RAN tasks, reserving alpha-numeric tasks (Digit Naming and Letter Naming) for children aged six and over (Troia et al., 1996; Wagner et al., 1999; Kirby et al., 2003; Taibah and Haynes, 2011; Norton and Wolf, 2012). Age effects on some measures may also reflect the effect of starting formal literacy teaching, since results for the PA tasks, especially in phoneme isolation task, show variable magnitudes of difference between the age groups. However, the mechanism underlying these predictors is unclear and needs further investigation.

5.4 Socioeconomic Status and PA and LK

This section discusses the findings in relation to the following hypothesis:

H3: Socioeconomic status will predict children's performance in RRS, with those from higher socioeconomic backgrounds achieving

significantly higher scores in PA, Blending and LK tasks than those from lower socioeconomic backgrounds.

This study aimed to investigate the impact of SES on PA, LK and blending tasks. Before conducting the various analyses to answer this research question, Pearson correlation analysis was done to check the association between SES factors (school type, school area, parental education, and a composite category combining property type and family income) and children's performance in three areas: the composite PA measure (as used previously), blending and LK. Blending was included because it is a good measure of explicit PA.

Linear regression analyses were then carried out to find out if SES measures could predict performance in the composite PA measure, blending and LK, and to identify which SES measure was the best predictor of performance. This was done by obtaining (R^2) or the square of the correlation coefficient to find out how much of the variance in the dependent variable (performance on composite PA measure, blending and LK) was explained by the independent variable (SES measures). It also gave an indication of the relative contribution of each independent variable. Multiple regression analysis was conducted using the standard entry method, taking age into account. To compare the contribution of each SES-related independent variable, β values were required. The highest β coefficient was.241, which was for the contribution of the composite category of family income plus property type to the composite PA score. This means that model 2 made a significant unique contribution to explain the dependent variable (composite PA measure) when the variance explained by all other variables in the model was controlled for. The β value for school type was lower (.102), indicating that it made less contribution to the composite PA measure variance.

The R^2_{Adj} for these SES predictors accounted for 26.5% (R^2_{Adj} =.265) of the variation in the composite PA measure performance (F (4,375) =136, p=.000), while for the blending task the combined effect of SES variables accounted for 13.5% (R^2_{Adj} =.135) of the variation (F=15.749, df= 375, p=.000). None of the SES variables made a significant contribution to performance in this task. For LK, the effect of all variables accounted for 24.9% (R^2_{Adj} 249) of the variation (F

(4,375) = 32.424, p=.000). Parental education and age, made a significant contribution (less than .05) as predictors.

5.4.1 SES and PA

A number of previous studies found a moderate positive association between social background and PA, which is in line with our results (e.g. Aram, et al., 2013; Arafat et al., 2017) with children from higher SES backgrounds out-performing those from lower SES backgrounds. The study by McDowell et al. (2007) found that the effect of SES on PA increased with age. Their cross-sectional study of children aged 2;0-5;0 years was a much larger sample (*n*=700) than the one tested for this study, but there is some potential overlap between their results and those in the youngest group (4;0-4;11) in this study. However, participants in McDowell et al.'s (2007) study represented both lower and higher SES whilst those in the Riyadh study were largely middle class. Both studies included PA as one of a range of measures. When McDowell et al. (2007) used multiple regression analysis to determine the amount of variance in PA explained by SES and age, their results indicated that SES, age, and other variables each contributed unique variance to the prediction of PA. They found that age moderates the relation between SES and PA (i.e., relations between SES and PA are amplified with increases in age). However, the present study did not examine the effects of SES in different age groups, or the interaction between SES and age. Another study by Noble et al. (2006) found a highly significant correlation between SES and PA ($p < .001$, r=.44) stronger than the current study ($p < .05$, r=.13). This may be because the study was conducted in a different linguistic, cultural and socioeconomic context. In Arabic-speaking context, Arafat et al. (2017) found also strong correlations between SES and PA (phoneme isolation task) in kindergarteners ($p < .01$, .38) and in first graders ($p < .01$, .31). However, participants in Arafat et al.'s study differed in SES background from the present study in that their participants mostly were with low SES. Although many studies have shown that SES and PA are associated (e.g. Noble et al., 2005; Arafat et al, 2017), the nature of this relationship is less clear. Lundberg et al. (2012) studied the influence of SES on PA among preschool children. The children were tested on two occasions: first at the start of the pre-school year (T1) and second after 8 months at the end of the academic year (T2). PA was related to SES between the levels, with children from low socio economic backgrounds showed lower PA scores. Lundberg et al. suggested that a stimulating home environment is important for teaching and modifying PA in children. They also argued that children in high SES schools had higher parental pressure for systematic language training.

5.4.2 SES and Blending

Results showed that none of SES variables made a significant contribution to performance in Blending task. In contrast to our study, Noble and colleagues (2006) tested reading ability of English speaking first grade children. They assessed a range of reading skills including blending and found that SES had a systematic effect on reading skills (Nobel et al., 2006). Moreover, Krishna Priya et al. (2018) carried out a study to assess PA in children from different socio economic backgrounds and age (n= 480, age= 3-7 years). Participants were divided into four age groups (i.e. group I (3-4 years), group II (4-5 years), group III (5-6 years) and group IV(6-7 years), which were further divided into groups (mid and high) based on their SES. Children in the age range of 5-6 years in the high SES group had higher scores on syllable blending task (96.83%), while scores were lower (66.16%) in the mid SES group. Krishna Priya et al. concluded that children from high socioeconomic background had higher scores in most of the reading tasks when compared to the scores of children with mid socioeconomic status. Also, several studies indicate that children entering the schools in areas of low SES have poor PA and are consistent with delay written word recognition (e.g. Duncan and Seymur, 2000).

5.4.3 SES and LK

Results of the present study showed that only two SES variables: school area ($r=0.26$ at $p\leq0.01$) and parental education ($r=0.11$ at $p\leq0.05$) correlated significantly with the LK task, with 1.8% of the variation accounted for by SES variables ($R^2=.018$). Similarly, Arafat et al. (2017) explored Arabic speaking children's literacy skills from kindergarten to first grade (n= 109; mean age =5.7 years). In kindergarten Arafat et al. found a significant but not very strong correlation between SES and letter names ($r =.22$; $p <.05$) while in the case of letter sounds there was a stronger correlation ($r=.32$; $p <.01$).

5.4.4 Overall Findings on SES

Although the sample in this study is comparable in size to those used in previous studies, it does not exhibit the same diversity in terms of SES found in earlier studies which were designed to compare participants from a representative range of socio-economic backgrounds. Arafat et al., 2017 found that SES has a direct effect on literacy skills in Kindergarten children, while an indirect effect on text reading ability in first graders. Another study by Aram et al. (2013) examined the relationship between SES, home literacy environment (HLE) and literacy skills in Arabic speaking children. Aram et al. investigated the link between child's literacy skills controlling for both SES and HLE. The study was conducted in 89 Israeli children of age range 5-8 years. Finding indicated that SES was a unique predictor with 18% variance in regression analysis. It was concluded that early literacy skills were significantly associated with SES and HLE. Similar to findings of this study, Aram et al. found that LK significantly correlated with parental education showing that children whose parents are educated demonstrate higher early literacy skills. Hence, low SES is considered one of the risk factors related to child's early literacy skills. This studies highlight the significance of family background in an Arab community. However, Arafat et al.'s study was conducted in Israel and most of the Arab families had low socio economic status, unlike the present study sample where the majority were from middle income background. This may account for the smaller effects of SES found in this study compared with studies in other Arabic speaking and other language contexts.

A sample containing private and public schools was chosen for this study as it was thought that this would be likely to provide a more varied profile of pupils. In some countries one might expect a noticeable socio-economic division in the background of students attending fee-paying private schools and those attending free public sector schools.

A study conducted in Riyadh (Alsuiadi, 2016) which examined parental reasons for choice of private versus public school highlighted that such decisions were not necessarily based exclusively on economic factors. Some high-income families chose the public sector as they thought teaching, discipline and instilling of religious values were better there. Other families at the lower end of the income range relied on financial help from relatives to send the eldest son to private school due to their “parental desires and ambitions to ensure their children were given every possible opportunity to compete effectively in modern life” (Alsuiadi, 2016: 211). This illustrates socio-cultural differences which mean that the factors typically used by Western researchers to determine socioeconomic background may function differently in the Saudi context. In this case, the final decision concerning the choice of schools lay outside the researcher’s control as this was made by the education authorities on her behalf, although the need for a range of socio-economic backgrounds to be reflected in the sample was made clear.

Findings relating to level of parental education (which has frequently been used as an indicator of income levels and class in Western research) also illustrate the difficulties of comparing research findings across cultures. Although a moderately significant positive correlation was found between the mother’s and father’s level of education ($r = 450$, $p < 0.01$) revealing that parents tend to have a similar level of educational qualifications, parental education as a composite category was not correlated with family income. Although increasing numbers of Saudi women now attend university, which is free for all Saudi citizens, it is still common for married women not to work outside the home even when they are qualified to do so. This means that although both husband and wife could be graduates, this would not necessarily translate into two salaries.

There are a number of ways in which the sample might not be considered representative. Firstly, its homogeneity in terms of class. Alnuaim’s (2013) research on class in Saudi society had not been published at the time the questionnaires for this study were being developed and thus was not able to be considered in relation to SES status. However, his study offers important insights into the nature of class and SES in contemporary Saudi society.

Alnuaim indicates that in economic terms it is now necessary to think of Saudi society in terms of ‘middle classes’ (lower middle class, neo-middle class and upper middle class), a category which now covers a very wide income range. e sets the lower middle class boundary at a monthly salary of 3,900 Saudi riyals (SR), ascending to 38,000 for the boundary between upper middle class and upper class. These boundaries do not correspond exactly with those used in the questionnaire in this study as shown in **Table 5.1**.

QUESTIONNAIRE CATEGORIES		ALNUAIM
Monthly earnings	**% of sample**	**Class**
Less than 3000 SR	5.5	**Lower (less than 3 900 SR)**
3 000-5 999 SR	9.6	**lower middle**
6 000-8 999 SR	22.9	
9 000-11 999 SR	19.6	**neo-middle**
12 000-14 999 SR	13.0	
15 000-17 999 SR	10.7	**upper middle (from 3 900 to 38 000 SR)**
18 000-20 999 SR	10.2	
21 000 + SR	8.2	**Upper class (over 38 000 SR)**

Table 5.1: Alnuaim's categories of class mapped onto those of the study questionnaire.

Using Alnuaim's divisions, this would mean that at least 75% of the participants in the sample would be categorised as belonging to the "middle classes", being spread across Alnuaim's sub-categories whilst less than 15% came from families in the lower class bracket. Nearly 85% of families in the sample have an income which totals 6,000 SR or more per month. For the purposes of comparison, at the time of writing, the starting salary for a Saudi teacher with an undergraduate degree in the private sector is 5,000 SR, and in the public sector, 8,060 SR.

Alnuaim's (2013) model of the Saudi class pyramid based on these figures shows the following class breakdown for contemporary Saudi society (**Figure 5.1**) If the figures for income used in the questionnaire are compared to Alnuaim's class categories and to his class pyramid, the percentage of those participants who could be described as coming from a lower class background was low in comparison to their representation in Saudi society. However, it must be remembered that Alnuaim's categories also include those living in rural areas who often fall into the lower income bracket. The fact that this study focused solely on an urban area and specifically excluded children who did not have Saudi citizenship in order to try to control for possible influence of linguistic variation somewhat complicates the issue of representativity.

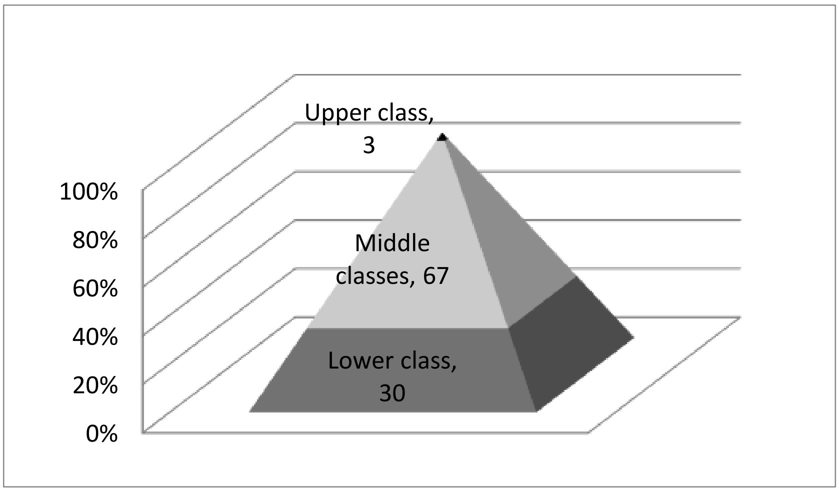


Figure 5.1: Saudi class pyramid (Alnuaim, 2013)

The second way in which the sample could be considered non-representative was in terms of the profile of educational attainment for the parents in the sample who might be considered to be exceptionally well educated. According to Alnuaim (2013), based on official statistics, in the Kingdom as a whole, currently 18% of the population have completed undergraduate studies with 45% having completed only elementary education. However, given the accessibility of educational provision at university level in Riyadh percentages might be expected to be higher.

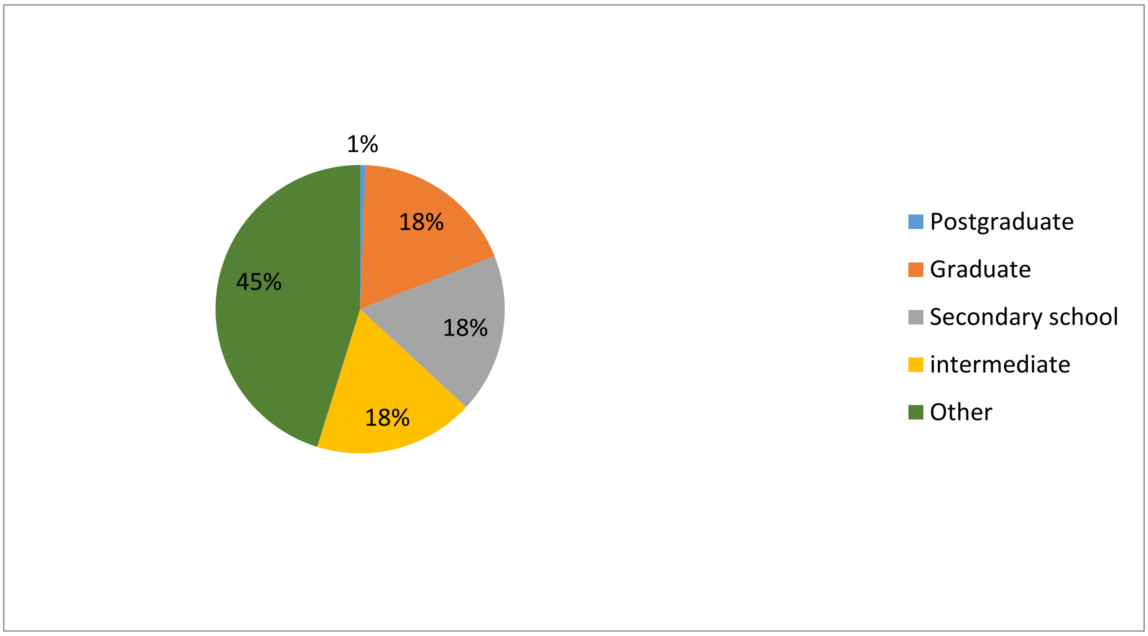


Figure 5.2: Educational profile of Saudi Arabia

(Source: based on data from Alnuaim 2013: 24)

Figures for this study revealed that slightly more mothers (46.7%) had achieved qualifications at HE level (diploma or bachelors degree) than fathers (44.5%). In addition 2.6% of mothers and 8.6% of fathers had postgraduate qualifications (Masters/PhD). Nonetheless, it is striking that, at the opposite extreme, the parental education profile for the sample indicates that 13.1% of females and 8.7% males had received no formal education or only elementary level. The most recent figures on the Riyadh City website (www.arriyadh.com) estimate that 8% of the capital's population are illiterate.

Another relevant factor in terms of the study's lack of representativeness is the fact that the participants are exclusively from Riyadh—the Saudi capital—which is more developed than would be the case elsewhere in the Kingdom. Ideally samples from different regions including more rural areas should have been used for comparative purposes to investigate the influence of SES on RRS. However, over 80% of the Saudi population now lives in an urban environment (Alnuaim, 2013).

All of the issues concerning the lack of representativeness of the sample discussed here can be considered as limitations of the study and need to be addressed in a further study to provide normative data for the wider Saudi population.

It is also worth noting one final factor which may have influenced the results relating to SES, namely, the survey method which was chosen to collect the data from participants' parents. In order to guarantee a high return the researcher administered the questionnaire herself by telephone. De Leeuwu (2005) found that when the researcher is present during collection of questionnaire data, even over the telephone, participants tend to give more socially desirable responses, meaning they may inflate their actual income or intellectual achievements. However, the researcher felt that it was worthwhile risking this potential effect on results in order to obtain a larger sample since the response rate from the pilot study had been relatively low. Many Saudis are still wary of providing personal information and the researcher was able to provide further personal reassurance concerning the purpose for collecting the data and the procedures used to ensure its confidentiality. As previously noted (see **section 3.10.3** and **3.10.4**), this strategy produced an excellent response rate. It is difficult to draw wholly meaningful comparisons between SES and RRS with research in other Arabic-speaking settings since the limited studies to date have often been carried out in conditions of extreme poverty and social deprivation of a kind not typically found in Riyadh (Arafat et al., 2017; Aram et al., 2013; Levin et al., 2008).

Statistically significant results were obtained in our study one, but the actual differences between the groups was very small, as reflected in the effect sizes for PA composite, blending and LK tasks (see **Table 3.38- 3.39**). With a large sample (in this case, n= 384) quite small associations can become statistically significant, even if the differences between the groups are of little practical importance.

In summary, only a weak correlation was found between participants' SES and their performance in RRS tasks, and the contribution of SES was limited. Since these results differ from previous findings regarding SES factors carried out in a Western context (for example, Lundberg et al. 2012), it would be interesting to investigate SES in a future study with a larger sample of children from different cultures.

5.5 The Effects of Exposure to Qur'anic Recitation

This section examines the findings in relation to the hypothesis:

H4: Children who have extra exposure to study of Qur'anic recitation will perform better in RRS (PA, LK and RAN), when controlling for age

Comparing the mean scores of the participants that had had extra exposure to Qur'anic recitation (*tajwīd*) tuition with those who had not been taking tuition showed that the former group performed better on all the PA measures, including the composite measure and LK tasks, whereas in the blending task the performance of both groups was similar. There appeared to be some differences in mean scores for performance in the Composite PA and LK (**Table 3.44**). ANOVA results showed that extra exposure to tajwid tuition had a small but significant effect on the Composite PA measure scores and LK (see **Table 3.46**). Also, the results showed negligible differences between the mean performance times across the four RAN tasks between the groups. The assumption is that the skills acquired in Qur'anic recitation do not relate to those tested in RAN tasks.

As noted in the Literature Review, previous studies (e.g. Rosowsky, 2001; Robertson, 2002) have attributed various linguistic and educational benefits to the impact of *tajwid* tuition but none of these studies investigated this phenomenon in very young children with Arabic as their mother tongue. As noted above, in this study, children who had extra *tajwid* sessions at the pre-school stage (4;0-4;11 and 5;0-5;11) performed better on the PA and LK tasks. A possible explanation for this is that during *tajwid* sessions children are exposed to print and letters/graphemes since some of them follow the text with their fingers at the same time as listening to the relevant section from the Qur'an.

When thinking about the possible skills which might be developed in *tajwid* classes, it is important to consider the form which this tuition takes. *Tajwid* systematizes in a rigorous and meticulous fashion how sounds, words and phrases should be articulated when reciting verses of the Qur'an on the basis of an oral tradition built up over the course of the centuries (Brierley et al., online). Tuition in *tajwid* covers many different components of pronunciation, developing students' awareness of particular phonological features which include contextual allophonic variation, stress patterns, prolongation, and special articulatory effects (Brierley et al., online).

Pupils are assessed on their ability to memorize specific verses of the Qur'an, their ability to recognize these and on their skills of articulation. In both public and private Saudi schools, children receive 45 minutes of *tajwid* tuition every day of the school week. During this time, they hear the same brief passage of the Qur'an (sometimes only four or five words of a *surah* (verse) at a time) repeated several times from an audio recording which is played to the whole class. The teacher then recites the same passage and the children repeat this all together as a class or as individuals. While they are listening to the Qur'anic text being recited, children normally have the written version of the same passage in front of them and some follow this with their finger. A similar method is employed by private tutors in the home environment or in sessions held at the mosque.

Given the range of different skills involved in these sessions, including auditory discrimination, memorisation, and articulatory production, the relationship with PA development is not a simple one. Furthermore, older children may develop decoding skills, learning how to link graphemes to phonemes without actually comprehending what is written in the text or recited, which is relevant to alliteration, phoneme isolation tasks and LK. Surprisingly, extra exposure to *tajwid* tuition did not appear to impact on performances in the blending task which requires participants to blend phonological units (syllables, onset-rime, or phonemes) into words, this is perhaps because when children are taught to recite the Qur'an they may learn to recognise phrases in the text and associate them with what is recited but they are not taught to analyse the text in terms of its micro-level components i.e. phoneme by phoneme which is the method used for teaching reading skills at this age. *Tajwid* can be thought of as a form of auditory memorisation and as such it does not place emphasis on distinguishing individual phonemes or lexemes nor does it teach children to develop techniques for identifying these. This suggests that any effects on PA must be implicit, since the teaching does not help children explicitly identify syllables or phonemes.

One of the shortcomings of this component of Study One makes it is difficult to draw and definitive conclusions concerning H4. In this study those who received extra exposure to *tajwid* tuition were identified on the basis of the questionnaire response provided by parents. However, parents were not asked to specify the actual quantity of extra tuition, the frequency of this and the content of these sessions. Given the lack of information about the amount, frequency and content of the extra *tajwid* tuition, it is not possible to claim that extra exposure to *tajwid* tuition accounts for the difference in performance in the LK tasks. Exposure to additional tuition may influence the children's RRS ability but further studies are needed to check the validity of this result, and second to investigate the source of any differences observed. It is possible that other factors not investigated in this study may account for group differences, including attention span, attendance patterns at school, and the reasons why children had extra tuition e.g. they might come from a home where a greater emphasis is placed on acquiring literacy skills in Arabic.

In the case of the oldest group, a further factor makes it difficult to draw any direct links between the level of exposure to *tajwid* tuition and levels of performance in the RRS tasks. Although there was no interaction with age, the difference between children with and without tajwid tuition is small in the oldest group, and this may be due to children in the first grade receiving tajwid as part of the school curriculum and formal reading instruction, including learning letters. This is an additional influence on this group which may impact on their PA, and LK skills. Nonetheless, since there were group differences in the oldest students on LK it is possible that the extra exposure to *tajwid* tuition (and hence to the written language) was helping the oldest group.

Whilst H4 cannot be fully evaluated due to the shortcomings of the study outlined above, there are indications that those children exposed to extra *tajwid* tuition may benefit from this at the start of the learning process as it may improve children's PA and help them when they are learning to read. However, the effect of tajwid tuition needs to be checked more rigorously by controlling other important factors, for example, by including children attending the same nursery and therefore receiving the same classroom input, and taking into account amount and type of extra tuition. If effects of extra tuition are confirmed, future studies could go on to investigate the reason for these effects, e.g. whether they impact on auditory discrimination, listening, attention, memory, and articulatory production.

5.6 RRS Tasks and Relationship to Teacher-Rated Reading Ability

As previously noted, Study Two aimed to evaluate the reliability and validity of the RRS test battery which was developed for Study One. Teachers were asked to identify the top five and bottom five performing study participants in the youngest and the intermediate groups on the basis of assessing their progress in reading-related abilities. This teacher-rated assessment of reading ability for the youngest and the intermediate age groups was a broad test of reading-related abilities including some LK, decoding and copying of letters due to the fact that there was no existing standardized Arabic literacy scale suitable for use with the age groups involved in this study. This ranking clearly involves a substantial element of subjectivity. Moreover, it probably reflected children's learning abilities more generally and a range of subskills, rather than their RRS specifically.

For the children in the oldest group (6;00-7;00), who were in first grade of elementary school, the teacher-rated assessment was a more objective measure of reading ability since it was tested by actual reading tasks including reading words with both short and long vowels, words consisting of two or three syllables and even short sentences (see **4.2.5.1** and **4.2.5.2** for details of these assessments). This reflects the fact that formal instruction in reading and testing progress in literacy starts at this stage within the Saudi system. Thus, although the teacher-rated assessment of the older group involved elements other than reading, is more subjective than a reading test and without a fixed format, it is nonetheless more of a measure of reading than the less specified and more subjective assessment used for the other two younger groups.

This needs to be taken into account when comparing the results of Study Two with previous studies that used a standardised reading measure. Moreover, since there was no prior screening in this study for children with dyslexia, it is possible that some of those that were ranked as the poorest performers may have been in this category. This is an area that would benefit from further exploration.

The ranking of children's reading ability by teachers, especially that of the younger and intermediate age groups (4;11-5;0 and 5;11-6;0), may also reflect their levels of prior exposure to LK and reading and this study did not include a measure of this. It should be remembered that children who had attended kindergarten would have had greater levels of exposure to LK. The researcher observed that entering school at different levels in the early stage could have impacted on the children's performance in RRS and would need to be controlled in future studies.

The following sub-sections concern the findings in relation to the hypothesis:

H5: Children with higher teacher ratings of reading abilities will perform significantly better on all RRS tasks than children with lower teacher ratings.

Findings for PA, LK and RAN are discussed but since Study One revealed that the RAN tasks were considered ineffective for use with the two younger groups, in Study Two the RAN tasks were only administered to the oldest participants (6;0-7;0).

5.6.1 <u>The predictiveness of PA tasks for teacher rating of reading ability</u>

Study Two findings show that the PA tasks that were used in this study have potential to predict reading ability since they were able to distinguish between the highest and lowest-rated participants, particularly in the case of the oldest group. However, for the youngest and the intermediate age groups this element of predictiveness was smaller though still significant. This means that H5 was broadly supported.

Descriptive analysis was done for the performance of the top/bottom ranking groups for the composite PA measure. In the top ranked group, the mean scores (SDs) were 23.20 (3.08), 25.20 (3.46), and 28.70 (1.64) for the age groups youngest (4;0-4;11), intermediate (5;0-5;11), and oldest (6;0-6;11) respectively (**Table 4.6**). Further analysis using univariate ANOVA showed that composite PA measure performance increased as the children got older (**Fig. 4.1**), which means that, as expected, the intermediate group performed better than the youngest, and the oldest group performed better than the other two groups. However, the pattern for those in the bottom-ranking group was different in that the mean scores of the bottom-ranking group for the intermediate [13.60 (3.17)] and oldest group [13.50 (2.47)] were lower than those of the youngest [17.50 (2.80)] (**Table 4.6**). Estimated marginal means of the PA composite measure revealed that there was no difference in performance between the intermediate and oldest groups, indicating that selection may not have been effective in their case (**Fig 4.1**). There was no significant effect of age on PA performance in both top and bottom groups, while a significant effect was found for the teacher rating of reading ability ranking on the composite PA measure (F (1.57=219.7, $p<0.05$).The interaction effect between age and reading ability was also significant for PA (F=14.35;p=0.00) (**Table 4.8**). The PA improves as the children get older and their abilities in this area may develop as a result of increased formal exposure to reading instruction. Findings of Al-Sulaihim, and Theo's longitudinal study (2017) showed an improvement in PA skills once formal literacy training had been introduced.

The results of the Blending task were interesting. Firstly, in the top-ranking group, the oldest group achieved the highest scores, followed by the youngest group and the intermediate group. This was not the case in the bottom-ranking group where the intermediate group achieved the highest mean scores. Conversely the youngest and the oldest groups both performed poorly in comparison with the intermediatiate group (see **Table 4.6**). The finding that the differences between top and bottom groups may be greater in the oldest group than the younger groups may be attributed to the less subjective nature of teacher rating such that this is more firmly based on reading abilities.

It is well known that PA develops in a continuum ranges from syllables toward phonemes level (Goswami and Bryant, 1990; Stanovich, 1992). In other words, syllabic awareness appears before phoneme awareness, and both predict reading skills independently (Engen and Hoien 2002; Muter, et al.; 2004; Aidinis, 2012). Also, some studies found phoneme skills was better to predict word recognition than onset-rime skills (Muter, Hulme, Snowling, and Taylor, 1998). In the current study the items of blending task were also ordered by level of difficulty (i.e. syllable, onset-rime and phoneme). Therefore, in our results top and bottom ranked groups were compared, it was the bottom oldest group that were at floor on the phoneme items, while the top oldest group were not at floor. Also, both the intermediate and the youngest children performed poorly on this element of the task.

In blending task, it is possible that the top intermediate group would do better on the syllable blending than the bottom group (with no difference on phoneme blending), which would suggest that syllable blending is an important indicator of RRS according to teacher rating in the younger age group. Therefore, it could be argued that some sub-tests should be excluded from the final battery for the youngest age group.

The study findings are consistent with Schaefer et al. (2009) and Taibah and Haynes (2011). The latter study showed that kindergarten (age= 6.33 years) and first grade participants (age= 7.17 years) had mastery only to syllable level that is consistent with the period of reading prior to the introduction of phonic reading of shallow, vowelized Arabic text. These data from Arabic support previous findings from English language research that the awareness of larger phonological units, not just phoneme-level awareness, is linked with early reading development (Anthony and Lonigan, 2004). This is also consistent with study one results which showed that children in the intermediate age group performed better in the Blending task at syllable rather than phoneme level. Overall, taking into account the similarities and differences in the samples involved and the measures used, the findings obtained in the present study for the oldest age group (6.00-7.00) are consistent with previous studies that demonstrated the existence of a moderate to high correlation and predictive relationship between performance in PA and reading development (word recognition and non-word reading). These results have been found in both English-speaking contexts (Wagner et al., 1993, 1994; Torgesen et al., 1997). Predictive ability of PA has been documented across orthographies (Caravolas et al., 2012) and in Arabic-speaking contexts, whether monolingual (e.g Al-Manni and Everatt, 2005; Taibah and Haynes, 2011, Tibi and Kirby, 2018; Tibi and Kirby, 2019) or bilingual (e.g Saiegh-Haddad and Geva, 2008).

Specifically in the Saudi context with respect to the Blending task and its predictiveness of reading ability, Taibah and Haynes (2011) found significant moderate to high correlations between the scores for their PA test battery and the early reading development tasks such as word decoding and text reading fluency tested in each school grade. The most important difference, however, was that in Taibah and Haynes' (2011) study all the children had begun formal instruction in reading and so they were able to be tested on their actual reading skills (word recognition, word reading, reading comprehension and fluency), using the same test for all groups. For the one age group that corresponds between the studies (first grade of elementary school 6;00-7;00), the finding that the top- and bottom-ranking groups is in line with Taibah and Haynes' (2011) findings that PA is a predictor of reading ability. (The evidence here is from comparison between two extreme groups, and this is not the same as looking at correlations in the whole group).

5.6.2 The predictiveness of LK tasks for teacher rating of reading ability

The data analysis has shown that children in the oldest group have the highest mean scores in the LK task which tested participants' knowledge of letter names and sounds, while children in the youngest group achieved the lowest mean scores (see **Table 4.**). When the mean scores in relation to reading ability classification (top and bottom ranking) were compared for all three age groups some interesting patterns emerged.

As expected, mean scores increased with age and for LK, as for Blending, there is only a small difference in means between the youngest and the intermediate groups for the performance by the top-ranking group of participants, but a greater increase of the mean scores between the intermediate and oldest groups. It is likely that this reflects the transition from pre-school to elementary school in the Saudi system and the start of formal training in reading at this age. In contrast, mean scores for the performances by those in the bottom-ranking groups show a more gradual improvement across all three groups by age. Strikingly, the mean for the oldest group fell marginally below the mean for the youngest top-ranking group (see **Figure 4.2***: Performance in the Blending task by reading ability ranking (top/bottom)*

).

A significant effect was found for reading ability ranking on the LK (see **Table 4.8**) demonstrating that the measure is valid for the older age group and that it can discriminate between those children who are likely to display strong and weak reading ability (the top- and bottom-ranking groups). There is also a difference, though smaller, in the youngest (4;0-4;11) and intermediate groups (5;0-5;11) supporting the potential of LK as an indicator of teacher ratings. These findings emphasise the role that LK can play in children who are learning to read and supports H5 in that the children's performance in LK tasks is a predictor of their reading ability, without assuming causal relations between these tasks.

The relationship between LK and teacher-rated reading ability might be partly due to both of these elements being affected by experiences that were not examined in this study as opposed to solely reflecting the child's ability. Attending kindergarten is thought to be a key experience in this respect. Children placed in the bottom group due to their poor performance according to the teacher rating may be as a result of their more limited experience of LK and reading (due to differences in input through kindergarten or at home).

Foy and Mann (2003: 65) note:

> The acquisition of letter names, like the acquisition of vocabulary, is clearly something that requires exposure. To learn letters a child must be exposed to them and their names and sounds. Aside from classroom experiences, the home literacy environment is surely the most important source of such exposure.

The original questionnaire gathered some relevant information regarding the home literacy environment of the participants, including, for example, availability of books in the home, and this information could be used to explore possible effects on LK specifically and RRS more generally. However, there is strong evidence for the predictive accuracy of teacher rating and RRS in early elementary school children and a strong correlation between teacher judgement and student’s performance on standardized reading tests (Hecht et al., 2001). Findings of this study suggest that teacher rating together with assessment of RRS and LK could help to identify children at risk of reading difficulties and dyslexia, enabling them to receive early, targeted educational interventions to support their reading development.

5.6.3 <u>The predictiveness of RAN tasks for teacher rating of reading ability.</u>

As discussed in the section on age and RAN tasks (**section 3.14.3.4**), the sample of data was reduced because of the number of invalid cases particularly in the younger age groups and this means that H5 could not be fully evaluated in the way that was planned. However, data for RAN tasks was considered for the oldest group of participants, for which it had proved to be valid. There was no significant difference between the two groups (top and bottom-ranking) for these constructs. However, the significance (2-tailed) value is smaller than .05 in the case of colour naming, meaning there was a significant difference between the top and bottom-ranking groups in this task, in favour of the top-ranking group.

It is worth considering any patterns reflected in the various tasks which were administered (**Table 4.7**). When the means are considered, as might be expected, there was a considerable difference between the top- and the bottom-ranking performances reflected in the times which participants recorded for completing this series of naming-speed tasks. However, the other outcome worthy of note here is the difference in the number of participants who succeeded in producing valid performances for this series of tasks. Whilst similar numbers of individuals in both top- and bottom-ranking groups were able to complete the Object Naming tasks satisfactorily (although the latter group recorded slower times), for all the other naming tasks, there is a striking difference in the numbers from each group who managed to complete the tasks which involved rapid naming of colours, letters and digits. This, in addition to the previous finding that many of the participants found the full set of all four RAN tasks overly taxing, adds further support to the recommendations by Taibah (2006) that alpha-numeric RAN tasks should be reserved for use with those who have already started formal education (6;0-7;0).

Taibah and Haynes (2011) noted that the predictive power of RAN remained lower than that of PA overall, but increased with age, appearing to correspond with increasing automaticity of decoding in Grade 2 of elementary school and increased demands on orthographic processing in the non-vowelized word recognition stimuli encountered in Grade 3. This may help to explain why the RAN tasks were more suitable for the highest age group (Grade 1) but did not produce significant intergroup differences for other age groups. However, literature indicated that predictive patterns for reading ability differ across orthographies (Georgiou, Parrila, and Liao, 2008a). While, Landerl et al. (2019), investigated the predictors of reading ability across five languages and found that predictive patterns were similar across orthographies for RAN, but not for PA in reading. These differences may be attributed not only to the depth of the orthography, but also to the age group assessed and the outcome measures used in the studies. In this study, we included Arabic speaking children of age range 4 to 7 years, while other similar studies involved children of different age groups (Landerl et al., 2019; Georgiou, Parrila, and Liao, 2008a).

5.7 Frith's Model and the Results of the Studies

The discussion concludes by drawing some tentative links between aspects of the two studies and Frith's model (1995), and also sheds light on the broader issues which this raises concerning the difficulties of attempting to measure children's reading-related abilities.

It is important to emphasise that this research was not intended to establish the relations between the different factors in the model theorised by Frith (1995) (namely, environmental, biological, cognitive, behavioural in relation to PA, LK and RAN) or to attempt to determine the degree to which those different factors are causative (see **Figure 2.6**). Rather, Frith's model was intended to serve as a framework for thinking about RRS and guided the selection of factors considered in the two studies. Thus, hypothesised effects of gender and age (H1 and H2) were related to Frith's biological factors of gender and age respectively whilst performance in RRS (PA, LK and RAN) combined what she described as cognitive and behavioural factors. In addition, Frith's model (1995) indicates the need to look carefully at environmentally determined factors that may influence RRS, hence the focus on SES (H3) and exposure to Qur'anic recitation (H4). Most crucially, Frith's model (1995) highlighted the need to consider multiple factors (biological, cognitive, behavioural and environmental) when developing an assessment designed to identify children at risk of developing reading difficulties and also to consider the factors that may influence children's performance in RRS in an assessment of this kind.

Beginning with the factors which Frith identified as biological, findings showed that the RRS tasks were sensitive to age since there is a correlation between the age of the participants and their performance in tasks designed to measure PA, LK, and RAN. Clearly, there will always be individual differences in developmental levels (as the outlier data shows) but in general terms, there is evidence that RRS skills develop in a particular order and that PA is a multi-level skill, and that as children develop they are able to progress to dealing with smaller linguistic units.

In the case of gender, however, results in this study were less conclusive in statistical terms, despite the body of literature that has pointed to girls generally having superior language and reading abilities to boys (see **section 2.5.1**).This finding is inconsistent with previous studies (Lundberg, 2012; Moura et al. 2009).

It is worth highlighting here the observation by Lundberg et al. (2012: 308) that "Despite decades of intense research, we have still no clear explanation to the basis of a putative gender difference in the kind of language skills involved in phonological awareness". The same authors also make an interesting general point which perhaps indicates the need to rethink gender somewhat differently in relation to Frith's (1995) model: "Is there a neuro-biologically based difference? Or is the pattern of socialization of girls different such that a more communicative and social orientation is encouraged?" (Lundberg et al. 2012: 308).

Analysis of the data for Study Two has shown that whilst PA, LK and RAN are indicators of reading ability in young children as rated by teachers, their strength as indicators may vary according to age. This may reflect changes in the teacher ratings used across this age range. Literature showed that studies used methods of teacher rating and it was moderately accurate in identifying children with reading difficulties (Bates and Nettlebeck 2001; Eckert et al. 2006; Flynn and Rabbar 1998). In Saudi Arabia, teacher ratings play a crucial role in identifying children at risk of developing reading difficulties as there is no reliable standardised reading assessment for Saudi children. Children are assessed at regular intervals in class on different language learning abilities. The results of the assessment are recorded on a chart provided for every student by the Ministry of Education. Based on the assessment results during the year, the teacher categorises the child as successful, partially successful or unsuccessful. It is possible that the use of a standardised, or at least consistent, direct assessment of children's abilities, would provide more accurate evidence of children's reading abilities and that selection of top- and bottom-ranking groups based on this might reveal more consistent effects of age within these groups (see unexpected findings for the three age bands in the bottom-ranking group as identified by teaching ratings, see **Figure 4.1, 4.2 and section 5.6.1**).

Study One results from the RRS test battery show that the LK of participants in the oldest group (6;0-7;0) showed marked improvement. Given that acquisition of these skills relies on exposure (**see 5.6.2**), in terms of Frith's model, it is likely that environmental input was responsible for this profile: participants in the oldest group developed grapheme-phoneme skills during the process of learning to read formally, and this in turn aided the development of their phonological processing to the level displayed in the study where their scores are consistently higher than those of the other two groups.

Moving on to other environmental factors which can impact on development of RRS, firstly, in the case of SES, statistical data did not indicate any strong positive correlation between the background of participants and RRS but this discussion has highlighted the problems of attempting to compare certain categories across cultures, the difficulties of applying understanding gained in one cultural context to another, and the importance of considering socioeconomic and sociocultural differences between cultures in investigation of environmental factors.

Secondly, the study of *tajwid* skills from a very early age within the Saudi educational system is a unique socio-cultural feature and a potential environmental factor which merits further exploration. Tajwid requires auditory discrimination, phonological storage, and articulatory production. It may also involve recognition of letter-sound relations and blending of sounds (if children follow the script with their fingers). This research provides some indication that the processes involved in *tajwid* and its practice might be helpful for PA, but a number of limitations in the investigation of tajwid in this study limit the conclusions that can be drawn from the findings. To check and extend these, it would be important in future studies to measure or control the duration of exposure to extra tuition in Qur'an recitation, and the type(s) or amount of *tajwid* skills targeted, as well as other environmental differences (e.g. nursery attendance, parental attitudes). Such studies would clarify whether and to what extent these environmental variables impact on children's performance in PA measures and the LK test.

General intelligence was not tested directly in my study but children were selected on the basis that they were considered to be developing normally. An even more comprehensive investigation of RRS using the Frith framework would include a measure of general intellectual abilities (g).

6 CHAPTER SIX: CONCLUSION

6.1 Overview of the Chapter

This chapter presents the conclusions that can be drawn from the research findings summarised and discussed in Chapter Five. It also outlines the limitations of the studies undertaken in this research, provides implications for assessment and intervention of children with reading difficulties and recommendations for future research.

6.2 Summary of Findings

Chapter Five presented the findings of the study in relation to the hypotheses that were formulated on the basis of findings from previous literature. It also briefly examined how these relate to Frith's (1995) model. This concluding chapter considers the results of those studies in the light of the primary aims of this thesis together with their implications. The primary aims of the thesis were to develop and trial a RRS battery and obtain normative data, and to explore the influence of various factors (gender, age, socioeconomic status, and additional exposure to Qur'anic recitation tuition) on the development of RRS in a sample of typically developing 4;0-7;0 year-old Arabic-speaking children in Saudi Arabia. This research was intended to contribute to the future development of an effective, comprehensive, standardized reading-skills assessment battery for use with Saudi children aged 4;0-7;0 years old. As previously noted, this type of tool is vital for early diagnosis of reading-related difficulties, allowing timely interventions and targeted remedial help to be made available when needed.

The effect of gender on RRS as measured by children's performance in PA, LK and RAN tasks was found to be minimal whereas in the case of age, the first two of these tasks were found to be age sensitive and helpful for discriminating between abilities whilst RAN proved to be useful only in the case of the oldest group. These results are in line with previous research and demonstrate that PA and LK are reliable and informative measures of abilities in Arabic-speaking young children. In the case of the RAN tasks, results supported recommendations made by previous researchers that the predictive ability of these tasks applies principally to older children (e.g. Kirby et al., 2003; Taibah and Haynes, 2011). These points would need to be incorporated into any future RRS test battery.

As noted in Chapter Three, although some specific socioeconomic factors were found to correlate with the RRS measured, this appears not to play the role previously noted in some studies such as those by Wallach et al. (1977), and Aram, et al. (2013).It is unclear if this was the result of the particular SES profile of the chosen sample or whether other issues need to be considered in making comparisons across cultures. Certainly, closer attention needs to be given to the factors used to define SES and to the child's experiences outside the formal teaching environment of the classroom.

Findings were inconclusive concerning the effect of levels of exposure to tuition in *tajwid* (Qur'anic recitation) but indicate that this culturally specific topic merited further investigation, particularly given that in the Saudi context it plays a uniquely important role in the lives of children both in the school environment and outside of it. In more general terms, the continuing central importance of memorisation in the education system in many Arabic-speaking countries also points to the need for more work in this area.

6.3 Implications of this Rresearch

The development of this RRS battery is the first step toward developing the provision of a benchmark for local professionals in the field of literacy to use with young Saudi students. The most robust result to emerge from this thesis is that this test battery for PA, LK and RAN is age-sensitive, revealing the developmental trajectory of RRS in Arabic-speaking children across age groups from 4;0 to 7;0 years.

This thesis highlighted the importance of assessing RRS using tests that capture the developmental sequence. Adopting this approach led to the design of developmentally orientated tests. This study has helped to provide a better picture of what to expect of Arabic-speaking children aged 4;0-7;0 years in terms of RRS including PA, LK and RAN.

In terms of discriminating between children's levels of abilities in RRS in the Saudi education system, the findings showed that rhyme awareness and syllable segmentation are useful for testing 4;0-4;11 year-olds, while for 5;0-5;11 year-olds it might be better to start with alliteration and phoneme Isolation as they are getting closer to decoding skills. For children aged 6;0-7;0 years in the first grade of elementary school, who are starting formal reading instruction, the blending task is useful. Blending is particularly important for using grapheme-phoneme correspondences to read new words, a route to reading that is most available in languages with a relatively transparent orthography, as in the case of the vowelized Arabic script to which children are first introduced. In order to be informative, RAN needs to be task-differentiated by age group, with alpha-numeric tasks reserved for children aged 6;0 or over.

Based on the results of Study One,[20] PA in young children follows the predicted developmental sequence. This suggests that assessments carried out by teachers and speech and language therapists should be developmentally organised (progressing from syllable, to rime, to phoneme) to ensure that emerging PA is identified and used as a platform from which further developments can be supported. Teachers of children aged 4;00-4;11 should focus on rhyme awareness and syllable segmentation activities, while those aged from 5;00-5;11 should in addition focus on alliteration, the earliest phoneme awareness task. When teachers are teaching phonics for children in Grade 1 (6;00-7;00) and the learning of letter-sound association, it is appropriate to focus on phoneme level (such as phoneme isolation and blending tasks) because this is required in the case of reading Arabic script which is largely mapping between phoneme and grapheme, particularly at the early stage when vowels are represented. However, results of this research suggest that even at the early stage teachers should work with children on awareness at the level of larger units. Thus, for example, syllable segmentation may be more informative about abilities in younger children at the initial stage,

[20] See boxplots for Tables 3.18-3.22 and Table 2.24 showing pair-wise comparisons between age groups for PA tasks.

when many children are unable to do blending. Research remains to be done on the relative efficacy of different phonological training strategies for remediation in young children at risk of developing reading difficulties in the Arabic context, drawing on evidence from research on remediation in other language contexts.

Following future standardization of the assessment developed in this study, it could be used for diagnosing those with difficulties with RRS at risk of developing into poor readers. Being aware of the pattern of performance between ability groups could be used to help diagnosticians determine whether or not the performance of an individual is typical. With an adequately sensitive tool, assessors will be able to (1) identify children needing support with decoding or reading that is affecting their educational achievement; (2) feel confident in their diagnosis of weaknesses that may pose a problem; (3) and attend to children’s educational needs in relation to reading by providing specific recommendations to parents and carers regarding support, and possibly classroom modifications and exam arrangements. An appropriate diagnosis that can identify areas of strength and weaknesses in RRS could contribute to a personalized and appropriate intervention plan.

6.4 Scope and Limitations of the Battery Developed

Despite the limitations noted below, the battery developed as the basis of this doctoral research is able to provide a comprehensive assessment of a range of RRS relating to PA, LK and RAN which is unusual in one battery. It appears to be able to tap into skills that previous literature has identified as being related to the acquisition of reading skills. It also gave results that were found relate to teacher ratings of individual abilities in RRS. With some modifications, it could be used with Arabic-speaking children in Riyadh, Saudi Arabia, to identify those individuals who are most at risk of potentially developing reading problems with the aim of providing timely intervention.

Based on the results and analysis of the two studies undertaken with a sample of children in Riyadh, clinicians and specialized assessors involved in diagnosing difficulties in RRS of reading disability who wish to use this test should pay attention to the following:

1. Caution needs to be exercised in interpreting the scores due to the fact that spoken Arabic has many different dialects (see **section 1.4**). This dialectal variation in Arabic means many different words might be used to refer to the same object which might affect scoring in a test. Therefore some of the items used in this test would need to be adapted for use with Arabic-speaking children from countries other than Saudi Arabia.

2. Caution must also be taken in interpreting scores that have been normed using children from other Arabic-speaking populations. Ideally, norms that have been derived from Arabic-speakers of a different nationality should not be used at all. It is important for those wishing to use this test in the future to adapt the Arabic RRS assessment test for local usage in order to build the norms that are needed to interpret scores with children who are speakers of Arabic dialects other than Saudi. No single assessment will be capable of determining children's level of RRS in Saudi Arabia. RRS assessment must therefore take into account the variety of Arabic that the child speaks to judge whether the RRS battery is applicable.

3. RRS assessment should also consider the possible differences in the amount of PA training children have received in school which could affect performance. Questionnaires and/or interviews with teachers and/or parents can be used to obtain information about the linguistic and literacy environment.

4. Although the battery is quite comprehensive, it is not an assessment of all the aspects of child development relevant to pupils at elementary school. A comprehensive RRS assessment of the Arabic-speaking child should also incorporate other measures of phonological skills such as phonological working memory (Taibah and Haynes, 2011) in addition to language abilities and knowledge, such as receptive vocabulary (Burt et al., 1999) to check whether the child has any language difficulties. It should also draw on teachers' classroom observations of executive functions such as attention and verbal memory that might contribute to the child's performance when testing for PA.

5. The RRS battery can be used to raise professionals' understanding of phonological awareness and its importance for literacy development, and their recognition of the need for early identification of a problem with RRS. Targeting PA and LK skills in the first grade may significantly improve the ability to read Arabic and literacy interventions should include aspects of RRS. Decoding (i.e. grapheme-phoneme mapping) can be part of an intervention programme for any child, irrespective of their academic level (Levin, Saiegh-Haddad, Hende, Ziv, 2008, Schneider, Roth, Ennemoser,2000; Hulme, C., Bowyer- Crane, C., Carroll, J., Duff, F. and Snowling, M. , 2012).

6.5 Limitations of the Research

Although an extensive piloting phase was conducted prior to this investigation, there were some methodological limitations to this study.

With regards to the sampling process, the selection of schools was limited to one urban context, namely Riyadh, and although there is considerable socioeconomic and ethnic diversity within the capital, the requirement for participants of Saudi nationality limited the range of backgrounds of participants. Whilst this does not invalidate the results of this investigation, testing participants from a more diverse range of backgrounds would have been helpful in terms of assessing the test battery.

When this study took place, it was not compulsory for all Saudi children to attend kindergarten, and as previously noted, this meant that not all first grade participants would have received the same amount of tuition at the point at which the study took place. In particular, some would already have received training on some of the RRS being assessed by the text battery. This potential advantage was not factored into this study.

For the study on the possible benefits of *tajwid* (Qur'anic recitation) on acquisition of RRS, participants were divided into just two groups: those who had been exposed to extra tuition in Qur'anic recitation and those who had not. However, it was not possible to factor in the exact nature of this exposure in terms of the amount received or the level of complexity of the material on which children had been trained during extra tuition in terms of phonemes; both of these elements might have had a bearing on their performance. Moreover, there was no baseline indicating the level of children's performance in RRS before starting extra tuition and, thus, measures taken indicate the child's proficiency at the time of the test, but not the impact of the tuition; that is, the pre- vs. post-tuition gain.

In addition to sampling, limitations exist in relation to the tasks of the test battery caused by the nature of Arabic and its alphabet. In the Phoneme Isolation task, children were asked to identify the first sound in a word that they heard (the words used for this task can be found in **Table 3.10**). However, Arabic has only three vowels (/a/, /i/, and /u/) with each having a short or long form. Consequently many words in Arabic have an initial syllable containing the short vowel /a/ for example, /ta**ʔ**/. Many of the names of letters in the Arabic alphabet also contain the same /a/ sound, such as, for example, /ba**ʔ/** or ta**ʔ/**. All but two of the items in the Phoneme Isolation task required an answer including an /a/ sound. As a result, it was often unclear whether the child had correctly isolated the required phoneme or was simply providing the first syllable of the word. In theory, this limitation could have been addressed by using words containing /i/ or /u/ but in practice this was not feasible since a high percentage of the most common Arabic words suitable for use with this age group feature /a/.

The current absence of a standardised reading measure for Arabic-speaking children was an important limitation in Study Two. This meant that the reading rating obtained from the teachers was based on a largely subjective judgement of performance in summative assessments carried out by the teacher at fixed points during the academic year. The fact that students were asked to perform different tasks at different times means that the teacher-rated reading assessment was not unified across all the children within each cohort. At kindergarten level, particularly, assessment gives only a general idea about the child's language ability. Moreover, in all cases, teachers' results may be influenced by observer bias, leading to an overly negative or positive view of a child's abilities. This limitation highlights the need for a standardised reading test that would take into account the specific characteristics of Arabic rather than being based on existing models originally developed for other languages and that can be used to assess children's reading at different ages and stages.

A final point concerns the absence of contextual information on health or cognitive difficulties that might have affected the children's performance. For example, at the time of testing and collecting data, the research participants did not undergo an auditory test to check whether children's level of hearing was within normal limits. If any children had an ongoing or temporary hearing impairment, such as might happen following a virus, this may have affected their hearing. This, in turn, might have affected their performance on the test battery by impairing their ability to discriminate between sounds.

6.6 Recommendation for Further Research

1. This study was carried out with children in the age group 4;0-7;0 years covering kindergarten, pre-school and first year of elementary school. Further studies are required which extend this age range to compare and contrast the role of RRS in older school children. Since many previous studies examined performance in RRS among children in the later years of elementary school (e.g. Taibah and Haynes, 2011, Al-Mannai and Everatt, 2005), this would also allow for more meaningful comparisons with studies done on other nationalities and language groups.

2. This research was devised as a cross-sectional study. Conducting longitudinal studies is necessary to track the development of RRS in children, the factors that hinder or foster this process and the link which these have with reading skills. A longitudinal mixed design (incorporating within and between subject factors) would have been a more informative research design since subjects could then be examined at different ages over a period of time. This type of study would provide useful information concerning the types and speed of changes. Most importantly, it would make it possible to evaluate the predictiveness of the test battery.

3. Given the amount of linguistic variation to be found within Arabic it would be particularly important to collect more representative samples of data from different areas or regions with Saudi Arabia and then see if it was valid to pool them. This would make it possible to select measures that are appropriate for individual children at risk of developing reading difficulties. Participants in the study were limited to children of largely middle-class

Saudi nationals in this urban environment. This yields limited results. To obtain more valid and comprehensive results future research should carry out studies on children from more diverse backgrounds in terms of ethnicity, social and economic background and place of residence.

4. As noted in this thesis, the learning of *tajwid* (Qur'anic recitation skills) begins at a very early age within the Saudi educational system and this unique socio-cultural feature merits further exploration, as this study was unable to provide any conclusive results in this area. This study did not measure or control for the duration of exposure to extra tuition in *tajwīd*, and the type or amount of *tajwid* skills. These variables may have had an important impact on children's performance in PA measures and the LK test. Collecting information on these variables in future research might show that *tajwid* experience contributes to children's performance in RRS. It is important to determine whether the same pattern of results would emerge using a stronger research design that controlled for the amount of *tajwīd* and for other factors that might be associated with extra *tajwid* tuition.

5. The study could also be extended by investigating performance on the RRS battery in groups of children who are at risk of reading difficulties or known to have reading difficulties. Such research would reveal whether the RRS battery is effective in identifying these children. It would also enable investigation of relations between their RRS and their reading difficulties.

6. As previously noted, the questionnaire designed for use in this study included items on children's home literacy environment which could be used in future data analyses.

6.7 Contribution of this Research

Although children's performance across the board may be influenced by factors that were not measured in this study and despite any limitations caused by the lack of a measure of reading, there is evidence of the battery's usefulness as a tool for practitioners attempting to identify those at risk of developing reading-related difficulties. The research reported in this thesis shows that the RRS battery is able to discriminate among young Arabic-speaking participants of different ages and indicates what skills are typical at what age. There are indications that those children exposed to extra *tajwid* tuition may benefit at the start of the learning process as it may improve children's PA and help them when they are learning to read. This study also offers important insights into the nature of class and SES in contemporary Saudi society. Furthermore, it is informative about skills known to be relevant in Arabic language, and finally, this research showed a relationship with teacher rating of children's performance in RRS skills. The nature of reading instruction in Arabic classrooms needs to be studied, and longitudinal studies would be helpful to further inform teaching/intervention programs PA and decoding.

7 BIBLIOGRAPHY

Abou-Elsaad, T., Ali, R. and El-Hamid, H. A. (2016). Assessment of Arabic phonological awareness and its relation to word reading ability. *Logopedics Phoniatrics Vocology*, *41*(4), 174-180.

Abu Ahmad, H., Ibrahim, R. and Share, D. L. (2014). Cognitive predictors of early reading ability in Arabic: A longitudinal study from kindergarten to grade 2. In E. Saiegh-Haddad and M. Joshi (Eds.), *Handbook of Arabic literacy: Insights and perspectives* (pp. 171–194). New York, NY: Springer. https://doi. org/10.1007/978-94-017-8545-7_8.

Abu-Rabia, S. (2001). The role of vowels in reading Semitic scripts: Data from Arabic and Hebrew. *Reading and writing: An Interdisciplinary Journal*, *14*(1), 39-59. doi:10.1023/A:1008147606320.

Abu-Rabia, S. (2000). Effects of exposure to literary Arabic on reading comprehension in a diglossic situation. *Reading and Writing: An Interdisciplinary Journal*, *13*, 147–157.

Abu-Rabia, S. (2007). The role of morphology and short vowelization in reading Arabic among normal and dyslexic readers in grades 3, 6, 9, And 12. *Journal of Psycholinguistic Research*, *36,* 89-106.

Abu-Rabia, S. and Siegel, L. S. (2002). Reading, syntactic, orthographic, and working memory skills of bilingual Arabic-English speaking Canadian children. *Journal of Psycholinguistic Research, 31*, 661–678.

Adams, M.J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Aidinis, A. and Nunes, T. (2001). The role of different levels of phonological awareness in the development of reading and spelling in Greek. *Reading and Writing: An Interdisciplinary Journal, 14*: 145–177, 2001.

Al- Asmari, M.G.H. (2008). Saudi Labor Force: Challenges and Ambitions. *Arts and Humanities, 16* (2), 19-59.

Al Suiadi, F. (2016). Reasons influencing decision making of parental choice of school. *International Journal of Research in Education and Science, 2*(1), 201-211.

Al-Arabiya, (2014). Enrolment rates at kindergarten Available online at: http://www.alarabiya.net/ar/saudi-today/2014/08/29/-أحوال-تعليم-لطلاب-بالمدارس-الحكومية-في-2014.html [accessed: 15/05/2015].

Al-Batal, M. (1992). “Diglossia proficiency: The need for an alternative approach to teaching”. *In* Rouchdy, A (ed.) *The Arabic language in America.* Detroit, MI: Wayne State Press, pp. 284-304.

Alloway T. P., Gathercole, S. E., Willis, C. and Adams, A. M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology*, *87*, 85-106.

Al-Mannai, H.A. and Everatt, J. (2005). Phonological processing skills as predictors of literacy among Arabic-speaking Bahraini school children. *Dyslexia, 11*, 269–291.

Alnuaim, M. (2013). *The composition of the Saudi middle classes: A preliminary study* Riyadh: gulf research centre available online at: http://www.files.ethz.ch/isn/171150/Middle_Class_8157.pdf [accessed on 24/03/2015].

Al-Shboul, O. Q. R. “The Arabic language: Young learners and literacy” in *Los efectos de los trastornos lingüísticos (lenguaje escrito) sobre el rendimiento del alumnado de primaria en Jordania*, Granada: University of Granada, pp. 280-310.

Al-Sulaihim, N. and Theo, M. (2017). Phonological awareness and literacy in Arabic-speaking children. *Arab Journal of Applied Linguistics, 2*(1), 74-90.

Al-Toma, S. J. (1969). *The problem of diglossia in Arabic: A comparative study of classical and Iraqi Arabic*. Cambridge, MA: Harvard University Press.

Ammar, M. and Ben Maad, M. R. (2013). The role of Arabic orthographic literacy in the phonological awareness of Tunisian children. *International Journal on New Trends in Education and their Implications, 4*(2), 15-23.

Anthony, J. L., and Lonigan, C. J. (2004). The nature of phonological awareness: Converging evidence from four studies of preschool and early grade school children. Journal ofEducational Psychology, 96, 43–55.

Anthony, J. L., Lonigan, C. J., Burgess, S. R., Driscoll, K., Phillips, B. M. and Cantor, B. G. (2002). Structure of preschool phonological sensitivity: overlapping sensitivity to rhyme, words, syllables, and phonemes. *Journal of Experimental Child Psychology, 82,* 65-92.

Anthony, J.L. and Francis, D.J. (2005). Development of phonological awareness. *Current Directions in Psychological Science, 14 (5)* 255-259.

Arab-Moghaddam, N. and Sénéchal, M. (2001). Orthographic and phonological processing skills in reading and spelling in Persian/English bilinguals. *International Journal of Behavioral Development*, 25(2), 140-147.

Arafat, A. H, Korat, O, Aram, D. and Saiegh-Haddad. E (2017). Continuity in literacy achievements from kindergarten to first grade: A longitudinal study of Arabic-speaking children. *Reading and Writing,* 30:989–1007.

Aram, D., Korat, O., Saiegh-Haddad, E., Arafat, S. H., Khoury, R. and Abu Elhiji, J. (2013). Early literacy among Arabic- speaking kindergarteners: The role of socioeconomic status, home literacy environment and maternal mediation of writing. *Cognitive Development. 28,* 193-208.

Asaad, H. and Eviatar, Z. (2014). Learning to read in Arabic: The long and winding road. *Reading and Writing*, *27*(4), 649-664.

Asadi, I. A. and Khateb, A. (2017). Predicting reading in vowelized and unvowelized Arabic script: An investigation of reading in first and second grades. *Reading Psychology*, *38*(5), 486-505.

Asadi, I. A., Khateb, A., Ibrahim, R. and Taha, H. (2017a). How do the different cognitive and linguistic variables contribute to reading in Arabic? Across-sectional study from first to sixth grade. *Reading and Writing: An Interdisciplinary Journal, 30*, 1835–1867.

Ball, E.W. and Blachman, B.A. (1991). Does phoneme awareness training in kindergarten make a difference in early word recognition and developmental spelling? *Reading Research Quarterly*, 49-66

Bates,C., and Nettelbeck, T.(2001).Primary school teachers' judgments of reading achievement. *Educational Psychology*, *21*, 177 –187.

Beauchat, K.A., Blamey, K. L. and Phillipakos, Z. A. (2012). *Effective* Read-Al ouds For Early Literacy. New York, NY: Guilford Press.

Below, J.L., Skinner, C.H., Fearrington, J.Y. and Sorrell, C.A. (2010). Gender differences in early literacy: An analysis of kindergarten through fifth-grade DIBELS probes. *School Psychology Review, 39,* 240–257.

Bentin, S. and Ibrahim, R. (1996). New evidence for phonological processing during visual word recognition: The case of Arabic. *Journal of Experimental Psychology Learning Memory and Cognition, 22*(2):309-23.

Berglund, J. (2010). Hidden knowledge: The value or cost of Islamic religious education in secular schooling. In Berglund, J. (ed) *European perspectives on islamic education and public schooling.* Sheffield: Equinox.

Bjaalid, Inger-Kristin, Høien, Torleiv and Lundberg, Ingvar. (1997). Dual - route and connectionist models: a step towards a combined model. Scandinavian Journal of Psychology 38: 73–82.

Blachman, B. A., Ball, E., Black, R. S. and Tangel, D. M. (1994). Kindergarten teachers develop phoneme awareness in low-income, inner-city classrooms: does it make a difference?. *Reading and Writing: An Interdisciplinary Journal, 6*, 1-18.

Blaiklock, K. (2004). The importance of letter knowledge in the relationship between phonological awareness and reading. *Journal of Research in Reading, 27*(1), 36–57.

Borzone de Manrique, A. M. and Signorini, A. (2002). El aprendizaje inicial de la lectura. Incidencia de las habilidades fonológicas de la estructura de la lengua, de la consistencia de la ortografíay delmétodo de enseñanza. Lingüística en el aula, 5(1), 29-48.

Boullata, I. J. (1989). Arabic oral traditions. *Oral Tradition, 4*(1-2), 1-3.

Bowey, J. (1995). Socioeconomic status differences in preschool phonological awareness and first-grade reading achievement. *Journal of Educational Psychology*, *87*, 476–487.

Bowey, J. A., McGuigan, M. and Ruschena, A. (2005). On the association between serial naming speed for letters and digits and word-reading skill: Towards a developmental account. *Journal of Research in Reading, 28*, 400-422.

Bowey, J.A., Cain, M. T. and Ryan, S. M. (1992). A Reading-level design study of phonological skills underlying fourth grade children's word reading difficulties*. Child Development, 63*, 999-1011.

Bowling, A. (2009). *Research methods in* health: *Investigating health and health services*. New York: McGraw-Hill.

Boyle, H. N. (2004). *Qur'anic Schools: Agents of preservation and change.* London: Routledge.

Bradley, L. and Bryant, P. E. (1978). Difficulties in auditory organization as a possible cause of reading backwards, *Nature, 271*, 746-747.

Bradley**,** L. and Bryant**,** P. E. (1983). Categorizing sounds and learning to read—A causal connection. *Nature, 301*, 419-421.

Brierley, C., Sawalha, M., Heselwood, B. and Atwell, E. (2016). A Verified Arabic-IPA mapping for Arabic transcription technology, informed by quranic recitation, traditional Arabiclinguistics, and modern phonetics. *Journal of Semitic studies*. LXI/1. doi: 10.1093/jss/fgv035.

British Dyslexia Association. .2007. Definition. Available online at: http://www.bdadyslexia.org.uk/dyslexic/definitions [accessed on 24/03/2015].

Bryant, P., Bradly, L., MacLean, M. and Crosslan, J.D. (1989). Nursery rhymes, phonological skills and reading. *Journal of child language*, *16*, 407-428.

Bryman, A. (2012). *Social Research Methods*.4th edn. Oxford: Oxford University Press.
Burgess, S.R. and Lonigan, C.J. (1998). Bidirectional relations of phonological sensitivity and pre-reading abilities: Evidence from a preschool sample. *Journal of Experimental Child Psychology, 70,* 117–141.

Burgess, S.R., Hecht, S.A. and Lonigan (2002). Relations of the home literacy environment (HLE) to the development of reading-related abilities: A one year longitudinal study. *Reading Research Quarterly*, *37,* 408-426.

Burgoyne, K., Kelly J.M., Whiteley, H E. and Spooner, A. (2009). The comprehension skills of children learning English as an additional language. *British Journal of Educational Psychology, 79,* 735–747.

Burman, D.D., Bitan, T. and Booth, J.R., (2008). Sex differences in neural processing of language among children, *Neuropsychologia 46*(5), 1349–1362.

Burt, L., Holm, A., and Dodd, B. (1999). Phonological awareness skills of 4-year-old British children: An assessment and developmental data. *International Journal of Language and Communication Disorders, 34*, 311–335.

Cain, K. (2010). *Reading development and difficulties* (Vol. 8). London John Wiley and Sons.

Caravolas, M., Hulme, C. and Snowling, M.J. (2001). The foundations of spelling ability: Evidence from a 3-year longitudinal study. *Journal of Memory and Language*, *45*, 751–774.

Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litavský, M., Onochie-Quintanilla, E., Salas, N., Schöffelová, M., Defior, S., Mikulajová, M. and Seidlová-Málková, G. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychological science*, *23*(6), 678-686.

Caravolas, M., Volin, J. and Hulme, C. (2005). Phoneme awareness is a key component of alphabetic literacy skills in consistent and inconsistent orthographies: evidence from Czech and English children. *Journal of Experimental Child Psychology, 2* (2), 107-39.

Cardoso-Martins, C, Lara Mesquita, T.C. and Ehri, L. (2011). Letter names and phonological awareness help children to learn letter-sound relations. *Journal of Experimental Child Psychology, 109*, 25–38.

Carroll, J.M., Snowling, M.J, Hulme, C. and Stevenson, J. (2003). The development of phonological awareness in preschool children. *Developmental Psychology, 39*(5), 913-923.

Castles, A. and Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read?. *Cognition, 91*, 77–111.

Chafouleas, S. M., Lewandowski, L. J., Smith, C. R. and Blachman, B. A. (1997). Phonological awareness skills in children: Examining performance across tasks and ages. *Journal of Psychoeducational Assessment, 15*(4), 334-347.

Chipere, N. (2014). 'Sex differences in phonological awareness and reading ability'. *Language Awareness 23*(3), 275–289. https://doi.org/10.1080/09658416.2013.7 74007.

Churchill, G. A. and Iacobucci, D. (2002). *Marketing research: Methodological foundations,* 8th edn. Mason, OH: South-Western.

Cirino, P. T., Israelian, M. K., Morris, M. K. and Morris, R. D. (2005). Evaluation of the doubledeficit hypothesis in college students referred for learning difficulties. *Journal of Learning Disabilities, 38*, 29–44.

Cohen, V. and Cowen, J. (2007). *Literacy for children in an information age: teaching reading, writing and thinking.* Belmont, CA: Wadsworth.

Coltheart, M., Rastle, K., Perry, C., Langdon, R. and Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*: 204–256.

Coolican, H. (2013). *Research methods and statistics in psychology.* New York: Routledge.
Cooter, R. B. and Reutzel, D. R. (2004). Teaching children to read: Putting the pieces together. Upper Sadle River, NJ: Pearson Education, Inc

Cossu, D., Shankweller, G., Liberman, I.W., Katz, L. and Tola, G. (1988). Awareness of phonological segments and reading ability in Italian children. *Applied Psycholinguistics, 9*, 1-16.

Creswell, J. (2003). *Research design: qualitative, quantitative and mixed methods approaches*. 2nd ed. Thousand Oaks, CA: Sage.

Cronin, V. and Carver, P. (1998). Phonological Sensitivity, Rapid naming, and beginning reading. *Applied Psycholinguistics*. *19,* 447-461.

Daly, E. J., Neugebauer, S., Chafouleas, S. and Skinner, C. H. (2015). *Interventions for reading problems*, 2nd edn. London; New York: The Guilford Press.

Daniels, P. and Bright, W. (1996). *The world's writing systems*. New York, NY: Oxford University Press.

De Jong, P. F. (2011). What discrete and serial rapid automatized naming can reveal about reading. *Scientific Studies of Reading*, *15* (4), 314-337.

De Leeuwu, E. J. (2005). To mix or not to mix data collection modes in surveys. *Journal of Official Statistics*, *21*(2), 233-255.

Deacon, H. and Kirby, J. (2004). Morphological awareness: Just 'more phonological? The roles of morphological and phonological awareness in reading development. *Applied Psycholinguistics, 25*, 223-238.

Defior, S. (2004). Phonological awareness and learning to read: A cross-linguistic perspective. In P. Bryant and T. Nunes (Eds.), *Handbook on children's literacy*. Pp. 163–174. Dordrecht: Kluwer.

Denckla, M.B. and Rudel, R.G. (1976). Rapid automatized naming: Dyslexia differentiated from other learning disabilities. *Neuropsychologia, 14*, 471–479.

Denton, C. A., Hasbrouck, J. E., Weaver, L. R. and Riccio, C. A. (2000). What do we know about phonological awareness in Spanish?. *Reading Psychology, 21*(4), 335-352.

DGER, (2010). *Directory of general education 2010,* Riyadh: Ministry of Education.

Di Filippo, G., Brizzolara, D., Chilosi, A., De Luca, M., Judica, A., Pecini, C., et al. (2005). Rapid naming, not cancellation speed or articulation rate, predicts reading in an orthographically regular language (Italian). Child Neuropsychology, 11, 349–361

Dickinson, D. K. and Tabors, P. O. (2001). *Beginning literacy with language: Young children learning at home and school.* Baltimore, MD, US: Paul H.

Dodd, B. A., Oerlemans, M. and McCormick, M. (1996). *Queensland university inventory of literacy*. University of Queensland, Australia: Department of Speech Pathology and Audiology.

Dodd, B., Crosbie, S., MacIntosh, B., Teitzel, T. and Ozanne, A. (2000). *Preschool and primary inventory of phonological awareness*. London: The Psychological Corporation.

Duncan, L.G. and Seymour, P.H.K. (2000). Socioeconomic differences in foundation-level literacy. *British Journal of Psychology, 91*, 145–166.

Duranovic, M., Huseinbasic, M. and Tinjak, S. (2012). Development of phonological awareness and letter knowledge in Bosnian preschool children. *International Journal of Linguistic*s, *4*(2), 18–34.

Durgunoğlu, A.Y. and Oney, B. (1999). A Cross-linguistic comparison of phonological awareness and word recognition. *Reading and Writing, 11*, 281-299.

Ehri, L. C. (1997). Sight word learning in normal readers and dyslexics. In B. Blachman (ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 163–189). Mahwah, NJ: Erlbaum.

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z. and Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, *36*(3), 250–287.

Ehri, L., and Wilce, L. (1980). The influence of orthography on readers' conceptualization of the phonemic structure of words. *Applied Psycholinguistics*, 1, 371-385.

Ehri, Linnea C. and Robbins, Claudia. (1992). Beginners need some decoding skill to read words by analogy. Reading Research Quarterly 27: 12–26.

Elbeheri, G. and Everatt, J. (2007). Literacy ability and phonological processing skills among dyslexic and non-dyslexic speakers of Arabic. *Reading and Writing: An Interdisciplinary Journal*, *20*, 273–294.

Elbeheri, G., Everatt, J., Mahfoudhi, A., Abu Al-Diyar, M. and Taibah, N. (2011). Orthographic processing and reading comprehension among Arabic speaking mainstream and LD children. *Dyslexia*, 1*7* (2), 123-142.

Elmonayer, R. (2013). Promoting phonological awareness skills of Egyptian kindergarteners through dialogic reading. *Early Child Development and Care, 183*(9):1229-1241.

Engen, L. and Hoien, T. (2002). Phonological skills and reading comprehension. *Reading and Writing: An Interdisciplinary Journal*, *15*(7-8), 613-631.

Escribano, C. L. (2007). Evaluation of the double-deficit hypothesis subtype classification of readers in Spanish. *Journal of Learning Disabilities, 40*, 319–330.

Ethnologue. Arabic: A macrolanguage of Saudi Arabia available online at: http://www.ethnologue.com [accessed on 25/01/15].

Faber, A., 1992. Phonemic segmentation as epiphenomenon. *The Linguistics of Literacy*, *21*, pp.111-134.

Farsoun, S.K. (2006). Class structure and social change in the Arab world. *In* N. S. Hopkins, and I. Saad Eddin (eds) *Arab society: Class, gender, power and development*. Cairo, New York: American University in Cairo Press, pp.11-28.

Fatani, A. H. (2006). Language and the Qur'an. In O. Leaman, *The Qur'an: An encyclopedia.* London; New York: Routledge, pp.356-372.

Ferguson, C. A. (1959). Diglossia. *Word, 15*, 325-340.

Field, A. (2009). *Discovering statistics using SPSS*, 3rd edn. London: SAGE.

Fischer, R. and Milfont, T.L. (2010). Standardisation in Psychological Research. *International Journal of Psychological Research, 3*(1), 88-96.

Flora, S., B. (2011). *Early literacy intervention activities, grades PK – K*. Greensboro, NC: Carson-Dellosa.

Flynn, J. M., and Rabbar, M. H. (1998). Improving teacher prediction of children at-risk for reading failure. *Psychology in the Schools*, 35, 163–172

Forrest-Pressley, D-L. and Waller, T. G. (2013). *Cognition, metacognition and reading*. New York: Springer-Verlag.

Foulin, J.P. (2005). Why is letter-name knowledge such a good predictor of learning to read?. *Reading and Writing*, *18*, 129–155.

Foy, J. G. and Mann, V. A. (2003). Home literacy environment and phonological awareness: Differential effects for phoneme awareness and rhyme awareness. *Applied Psycholinguistics*, *24*, 59-88.

Foy, J. G. and Mann, V. A. (2006). Changes in letter-sound knowledge are associated with development of phonological awareness in pre-school children, *Journal of Research in reading. 29,* 143-161.

Frederickson, N., Frith, U. and Reason, R. (Eds) (1997). *Phonological assessment battery (PhAB).* Windsor: NFER-Nelson.

Frith, U. (1995). Dyslexia: Can we have a shared theoretical framework?. *Educational and Child Psychology, 12*, 6-17.

Furness, B. and Samuelsson, S. (2009). Preschool cognitive and language skills predicting Kindergarten and Grade 1 reading and spelling: A cross-linguistic comparison. *Journal of Research in Reading*, *32*(3), 275–292.

Furness, B. and Samuelsson, S. (2011). Phonological awareness and rapid automized naming predicting early development in reading and spelling: Results from a cross-linguistic longitudinal study. *Learning and Individual Differences, 21*(1):85-95.

Garcia, J. R. and Cain, K. (2014). Decoding and reading comprehension: A meta-analysis toidentify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research, 84(*1), 74–111.

Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instrument. *Psychological Assessment, 6*, 304-312.

Gent, B. (2002). *The Basic Steps to Learning.* London: Qabas.

Georgiou, G. K., Papadopoulos, T. C., Fella, A. and Parrila, R. (2012). Rapid naming speed components and reading development in a consistent orthography. *Journal of Experimental Child Psychology, 112,* 1–17. https://doi.org/10.1016/j.jecp.2011.11.006.

Georgiou, G., Parrila, R. and Liao, C.-H. (2008a). Rapid naming speed and reading across languages that vary in orthographic consistency. *Reading and Writing, 21*, 885–903.

Georgiou, G. K., Parrila, R. and Papadopoulos, T. C. (2008). Predictors of word decoding and reading fluency across languages varying in orthographic consistency. *Journal of Educational Psychology, 100*(3), 566–580.

Georgiou, G.K., Torppa, M., Manolitsis, G., Lyytinen, H. and Parrila, R. (2012). Longitudinal predictors of reading and spelling across languages varying in orthographic consistency. *Reading and Writing*, *25*(2), 321-346.

Gibbs, S. (2003). Do pictures make a difference? A test of the hypothesis that performance in tests of phonological awareness is eased by the presence of pictures. *Educational Psychology in Practice*, *19*, 219–228.

Gillon, G. T. (2012). *Phonological awareness: from research to practice*, 2nd edition. New York: Guilford Press.

Goswami, U. and Bryant, P. (1990). *Phonological skills and learning to read*. Hove: Psychology Press.

Gough, P.B. and Tunmer, W. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*, 6–10.

Grainger, J. (2010). *Reading Foundations: A Structured Program for Teaching Essential Reading Skills.* Victoria: Australian Council for Educational Research.

Gral-Azulay, I. (2015). *Phonological and morphological skills in emerging English-Hebrew bilinguals* (Doctoral dissertation, University of Toronto).

Gullberg, J. and Granholm, J. (2010). *A Phonic stick study: Investigating the effectiveness of a phonological awareness intervention in children with Down's syndrome.* Master's dissertation, Uppsala University.

Haeri, F. (2011). *The essential message of the Qur'an.* London: Zahra.

Haeri, N. (2000). Form and ideology: Arabic sociolinguistics and beyond. *Annual Review of Anthropology,* 29, 61-87.

Halpern, D. (1986). *Sex differences in cognitive abilities*. Hillsdale, NJ: Erlbaum.

Harm, M. W. and Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review. 106*: 491–528.

Harm, M. W. and Seidenberg, M. S. (2004). Computing the meaning of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review. 111*: 662–720.

Harmmersley, M. (2013). *What is Qualitative Research?* London: Bloomsbury Academic.

Hatcher, P. J., Hulme, C. and Ellis, A. W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills the phonological linkage hypothesis. *Child Development, 65*(1), 41-57.

Hecht, S. A. and Greenfield, D. B. (2001). Comparing the predictive validity of first grade teacher ratings and reading-related tests on third grade levels of reading skills in young children exposed to poverty. *School Psychology Review*, *30*(1), 50.

Hecht, S. A., Burgess, S. R., Torgesen, J. K., Wagner, R. K. and Rashotte, C. A. (2000). Explaining social class differences in growth of reading skills from beginning kindergarten through fourth grade: The role of phonological awareness, rate of access, and print knowledge. *Reading and Writing: An Interdisciplinary Journal*, *12*(1-2), 99–127.

Hogan, T. P., Catts, H. W. and Little, T. D. (2005). The relationship between phonological awareness and reading: Implications for the assessment of phonological awareness. *Language, Speech, and Hearing Services in Schools, 36*, 285–293.

Hoien, T. and Lundberg, I. (1988). Stages of word recognition in early reading development. *Scandinavian Journal of Educational Research*, *32,* 163-182.

Holes, C. (1995; 2004). *Modern Arabic: Structures, functions, and varieties.* London: Longman.

Holopainen, L., Ahonen, T. and Lyytinen, H. (2001). Predicting delay in reading achievement in a highly transparent language. *Journal of Learning Disabilities, 34*, 401–413.

Houser, R. A. (2014). *Counselling and educational research: Evaluation and application*, 2nd edn. Thousand Oaks, CA: Sage.

Huang, F.L., Tortorelli, L.S. and invernizzi, M.A. (2014). An investigation of factor associated with letter-sound knowledge at kindergarten entry. *Early Childhood Research Quarterly, 29*, 182-192.

Hulme, C., Bowyer- Crane, C., Carroll, J., Duff, F. and Snowling, M. (2012). The causal role of phoneme awareness and letter-sound knowledge in learning to read: Combining intervention studies with mediation analyses. *Psychological Science, 23*(6), 572–577.

Hulme, C., Hatcher, P., Nation, K., Brown, A., Adams, J. and Stuart, G. (2002). Phoneme awareness is a better predictor of early reading skill than onset-rime awareness. *Journal of Experimental Child Psychology, 82*, 2- 28.

Hulme, C., Snowling, M., Caravolas, M. and Carroll, J. (2005). Phonological skills are probably one cause of success in learning to read: A comment on Castles and Coltheart. *Scientific Studies of Reading, 9*(4), 351-365.

Hutzler, F., and Wimmer, H. (2004). Eye movements of dyslexic children when reading in a regular orthography. *Brain and Language*, 89 (1), 235-242.

Ibrahim A. Asadi, I.A. and Abu-Rabia, S. (2019). The Impact of the Position of Phonemes and Lexical Status on Phonological Awareness in the Diglossic Arabic Language. *Journal of Psycholinguistic Research, 48*:1051–1062

Jenkins, J.R.B. and Jenkins, L.M. (1972). Comparison of letter name and letter sound training as transfer variables. *American Educational Research Journal, 9,* 75-86.

Johnston, R. S., Anderson, M. and Holligan, C. (1996). Knowledge of the alphabet and explicit awareness of phonemes in pre-readers the nature of the relationship. *Reading and Writing*, *8*, 217-234.

Juel, C. (1988). Leaning to read and write: a longitudinal study of fifty four children from first through fourth grade. *Journal of Educational Psychology*, *80*(4), 437-447.

Kahteran, N. (2006). Tajwid. In Leaman, O. *The Qur'an: An encyclopaedia* London; New York: Routledge, pp.635-638.

Kats, L. and Frost, R. (1992).The reading process is different for orthographies: The orthographic depth hypothesis. In: L. Katz and R. Frost (Eds.), Orthography, phonology, morphology, and meaning, 67-84. Amsterdam: North-Holland.

Kawamoto, A.H., Liu, Q., Lee, R.J. and Grebe, P.R. (2014). The segment as the minimal planning unit in speech production: Evidence based on absolute response latencies. *The Quarterly Journal of Experimental Psychology*, *67*(12), 2340-2359.

Kennedy, E. (2012) Literacy in Early Childhood and Primary Education (3-8 years), Research Report commissioned by Ireland National Council for Curriculum and Assessment: Dublin.

Kertoy, M. (2005). Phonological awareness development in preschool children: Factors associated with performance. Poster presented at convention of American Language Hearing Association, held in San Diego, CA.

Kim, Y. S., Petscher, Y., Foorman, B. R. and Zhou, C. (2010). The contributions of phonological awareness and letter-name knowledge to letter-sound acquisition—a cross-classified multilevel model approach. *Journal of Educational Psychology*, *102*(2), 313.

Kirby, J. R., Parilla, R. K. and Pfeiffer, S. L. (2003). Naming speed and phonological awareness as predictors of reading development. *Journal of Educational Psychology,* 95, 453–464.

Kirby, J.R. and Savage, R.S. (2008). Can the simple view deal with the complexities of reading?. *Literacy*, *42*(2), 75–82.

Klenke, K. (2008). *Qualitative Research in the Study of Leadership.* Chicago: Emerald Group Publishing.
Kline, P. (1999). *The handbook of psychological testing* (2nd ed.). London: Routledge.

Korat, O., Klein, S. P. and Segal-Drori, O. (2007). Maternal mediation in book reading, home literacy environment, and children's emergent literacy: A comparison between two social groups. *Reading and Writing: An International Journal, 20,* 361–398.

Krishna Priya, .G., Chacko, G. and Avilala, V. K. (2018). "Phonological awareness in different socioeconomic and age groups". *International Journal of Current Research, 10*, (03), 66829-66838.

Landerl, K. and Wimmer, H. (2000). Deficits in phoneme segmentation are not the core problem of dyslexia: Evidence from German and English children. *Applied Psycholinguistics, 21*, 243–262. https://doi.org/10.1017/S0142716400002058.

Landerl, K. and Wimmer, H. (2008). Development of word reading fluency and spelling in a consistent orthography: An 8-year follow-up. *Journal of Educational Psychology, 100*(1), 150–161.

Landerla, K., Freudenthalera, H.H., Heeneb, M., De Jongc, P.F., Desrochersd, A., Manolitsise, G., Parrila. R.F. and Georgioug, G.k. (2019). Phonological awareness and rapid automatized naming as longitudinal predictors of reading in five alphabetic orthographies with varying degrees of consistency. *Scientific Studies of Reading*. 23, (3), 220–234.

Landy, F. J. and Conte, J. M. (2010). *Work in the 21st century: An introduction to industrial and organizational psychology*. Hoboken, NJ: Wiley-Blackwell.

Lane, S. J. and Bundy, A. C. (2012). *Kids can be Kids: A Childhood occupations approach.* Philadelphia, PA: F. A. Davis.

Layes, S., Lalond, R. and Rebai, M. (2015). Effectiveness of a phonological awareness training for Arabic disabled reading children: Insights on metalinguistic benefits. *Bellaterra Journal of Teaching and Learning Language and Literature, 8*(4), 24-42.

Layes, S., Lalonde, R. and Rebai M. (2017). Automatized naming through word reading and comprehension in normal and disabled reading Arabic-speaking children. *Reading and Writing Quarterly. 33*, (2), 123–140.

Layes, S., Lalonde, R. and Rebaï, M. (2015). Reading speed and phonological awareness deficits among Arabic-speaking children with dyslexia. *Dyslexia, 21*, 80 –95.

Leopla, J., Poskiparta, E., Laakkonen, E. and Niemi, P. (2005). Development of and relationship between phonological and motivational processes and naming speed in predicting word recognition in grade 1. *Scientific Studies of Reading, 9*, 367–399.

Levin, I., Saiegh-Haddad, E., Hende, N. and Ziv, M. (2008). Early literacy in Arabic: An intervention study among Israeli Palestinian kindergartners. *Applied Psycholinguistics, 29,* 413-436.

Lieberman, I., Shankweiler, D., Fischer, F.W. and Carter, B. (1974). Reading and the awareness of linguistic segments. *Journal of Experimental Child Psychology, 18*, 201-212.

Literacy Trust. Illiterate adult in English, Literacytrust.org.uk, available online at: http://www.literacytrust.org.uk/adult_literacy/illiterate_adults_in_england [accessed: 18/05/2015].

Logan, J. A. R., Schatschneider, C. and Wagner, R.K. (2009). Rapid serial naming and reading ability: the role of lexical access. *Reading and Writing, 24*, 1–25.

Lonigan, C. J., Burgess, S., Anthony, J. and Barker, T. (1998). Development of phonological sensitivity in 2- to 5-year old children. *Journal of Educational Psychology, 90*, 294–311.

Lovett, M. W., Steinbach, K. A. and Frijters, J. C. (2000). Remediating the core deficits of developmental reading disability: A double-deficit perspective. *Journal of Learning Disabilities, 33*, 334–358.

Lundberg, I., Frost, J. and Petersen, Q P. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly, 23*, 263—284.

Lundberg, I., Larsman, P. and Strid, A. (2012). Development of phonological awareness during the preschool year: The influence of gender and socio-economic status. *Reading and Writing, 25*(2), 305–320.

Lyon, G. R., Shaywitz, S. E. and Shaywitz, B. A. (2003). Defining dyslexia, comorbidity, teachers' knowledge of language and reading: A definition of dyslexia. *Annals Of Dyslexia, 53,* 1-14.

Maccoby, E. and Jacklin, C. (1974). *The psychology of sex differences*. London: Oxford University Press.

Maclean, M., Bryant, P. and Bradley, L. (1987). Rhymes, nursery rhymes, and reading in early childhood. *Merrill-Palmer Quarterly, 33*, 255-281.

Mahfoudhi, A., Everatt, J. and Elbeheri, G. (2011). Introduction to the special issue on literacy in Arabic. *Reading and Writing, 24*,011–1018.

Makhoul, B. (2017). Moving beyond phonological awareness: The role of phonological awareness skills in Arabic reading development. *Journal of psycholinguistic research*, *46*(2), 469-480.

Manis, F. R., Doi, L. M. and Bhadha, B. (2000). Naming speed, phonological awareness, and orthographic knowledge in second graders. *Journal of Learning Disabilities, 33*, 325–333.

Mann V. and Wimmer H. (2002). Phoneme awareness and pathways into literacy: A comparison of German and American children. *Reading and Writing: An Interdisciplinary Journal. 15* (7–8), 653–682.

Mann, V. A. and Foy, J. G. (2003). Phonological awareness, speech development, and letter knowledge in preschool children. *Annals of Dyslexia, 53*, 149-173.
Foy, J. G., and Mann, V. A. (2001). Does Strength of Phonological Representations Predict Phonological Awareness in Preschool Children? Applied Psycholinguistics, 22, 301-325.

Matthew, J. S., Ponitz, C. C., and Morrison, F. J. (2009). Early gender differences in self-regulation and academic achievement. Journal of Educational Psychology, 101, 689–704.

Mattingly, I. G. (1972). Reading: the linguistic process and linguistic awareness. In J. F. Kavanagh and I. G. Mattingly (Eds.), *language by ear and by eye: the relationship between speech and reading.* Cambridge, Massachusetts: M.I.T. Press, 1972.

Mayringer, H., Wimmer, W., and Landerl, K. (1998). Phonological skills and literacy acquisition in German. In P. Reitsma and L. Verhoeven (Eds.), Problems and interventions in literacy development (pp. 147–161). Dordrecht: Kluwer Academic Publishers.

McBride-Chang, C. (1999). The ABCs of the ABCs: The development of letter-name and letter-sound knowledge. *Merrill-Palmer Quarterly,* 45, 285–308.

McDowell, K. D., Lonigan, C. J. and Goldstein, H. (2007). Relations among socioeconomic status, age, and predictors of phonological awareness. *Journal of Speech, Language, and Hearing Research,* 50 (4), 1079–1092. doi:10.1044/1092-4388 (2007/075).

McGuiness, D. (2004). *Early reading instruction: What science really tells us about how to teach reading* .Cambridge, MA: MIT Press.

McGuinness, D. (2005). *Language development and learning to read. The scientific study of how language development affects reading.* Cambridge, MA: MIT Press.

McIntosh, B., Crosbie, S., Holm, A., Dodd, B. and Thomas, S. (2007). Enhancing the phonological awareness and language skills of socially disadvantaged preschoolers: An interdisciplinary programme. *Child Language Teaching and Therapy*, *23*(3), 267-286.

Melby-Lervåg, M., (2012). The relative predictive contribution and causal role of phoneme awareness, rhyme awareness, and verbal short-term memory in reading skills: A review. *Scandinavian Journal of Educational Research*, *56*(1), 101-118.

Melby-Lervag, M., Lyster, S.A. and Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin, 138*, 322–353.

Menkes, J. H., Sarnat, H. B. and Maria, B. L. (2006). *Child neurology*. Philadelphia: Lippincott Williams and Wilkins.

Meyer, M. S., Wood, F. B., Hart, L. A. and Felton, R.H. (1998). Selected predictive value of rapid automatized naming in poor readers. *Journal of Learning Disabilities, 31*, 106-117.

Ministry of Education. (1998). Student Appraising List. The Higher Committee for Education Policy, Ministry of Education in the Kingdom of Saudi Arabia, Available online at: https://www.moe.gov.sa/Arabic/Pages/Default.aspx [accessed on 12/09/2015].

Mohamed, A.H. H., Hassan, A. S., Al-Qaryoutic, I., Al-Hashimi, A. and Al-Kalbanid, Z. (2019). The development of phonological awareness among pre-schoolers. *Early Child Development and Care.* https://doi.org/10.1080/03004430.2019.1607320.

Moll, K., Fussenegger, B., Willburger, E. and Landerl, K. (2009). RAN is not a measure of orthographic processing. Evidence from the asymmetric German orthography. *Scientific Studies of Reading, 13*, 1–25. https://doi.org/10.1080/10888430802631684.

Morais, J., Cay, L., Alegria, J., and Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? Cognition, 7, 323-331.

Motulsky, H.J. (1995). *Intuitive Biostatistics. Choosing a Statistical Test.* Oxford Oxford University Press.

Moura, S.R., Mezzommo, C.L. and Cielo, C.A. (2009).Phonemic awareness stimulation and its effects regarding the variable gender. *Pro Fono, 21*(1), 51–56.

Moyle, M. J., Heilmann, J. S. and Berman, S. S. (2013). Assessment of Early Developing Phonological Awareness Skills: A Comparison of the Preschool Individual Growth and Development Indicators and the Phonological Awareness and Literacy Screening–PreK, *Early Education and Development, 24*:5, 668-686.

Muter, V. and Snowling, M. (1998). Concurrent and longitudinal predictors of reading: The role of metalinguistic and short-term memory skills. *Reading Research Quarterly, 33*(3), 320-337.

Muter, V., C. Hulme, C., Snowling, M. J. and Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology, 40*(5): 665-681.

Muter, V., Hulme, C. and Snowling, M. (1997a). *Phonological Abilities Test (PAT)*. London: Psychological Corporation.

Muter, V., Hulme, C., Snowling, M., and Taylor, S. (1998). Segmentation, not rhyming, predicts early progress in learning to read. Journal of Experimental Child Psychology, 71, 3–2.

Nancollis, A., Lawrie, B. and Dodd, B. (2005). Phonological awareness intervention and the acquisition of literacy skills in children from deprived social backgrounds. *Language, Speech, and Hearing Services in Schools, 36*, 325–335.

Nation, N. and Hulme, C. (1997). Phonemic segmentation, not onset – rime segmentation, predicts early reading and spelling skills. *Reading Research Quarterly, 32*, 154-167.

National Institute for Literacy (2008). Developing early literacy: Report of the early literacy panel:a scientific synthesis of early literacy development and implications for intervention. Jessup, MD: National Institute for Literacy.

Nelson, K. (1996). *Language in cognitive development: The emergence of the mediated mind.* Cambridge: Cambridge University Press.

Nikolopoulos, D., Goulandris, N., Hulme, C. and Snowling, M. J. (2006). The cognitive bases of learning to read and spell in Greek: Evidence from a longitudinal study. *Journal of Experimental Child Psychology, 94*, 1–17.

Noble, K.G., Farah, M.J. and McCandliss, B.D. (2006). Socioeconomic background modulates cognition–achievement relationships in reading. *Cognitive Development, 21*, 349–368.

Norton, E.S. and Wolf, M. (2012). Rapid automatized naming (RAN) and reading fluency: Implications for understanding and treatment of reading disabilities. *Annual Review of Psychology, 63*, 427-52.

Olson, D. R. and Torrance, N. (2009). *The Cambridge Handbook of Literacy.* Cambridge: Cambridge University Press.

Öney, B. and Goldman, S. (1984). Decoding and comprehension skills in Turkish and English: effects of the regularity of grapheme-phoneme correspondences. *Journal of Educational Psychology, 76*: 557–568.

Pallant, J. (2010). *SPSS survival manual: A step guide to data analysis using spss,* 4th edn. Crows Nest, NSW: Allen and Unwin.

Paris, S. G. and Paris, A. H. (2007). Assessments of early reading. In Damon, D., Lerner, R. M., Renninger, K. A. and Sigel, I. E. *Handbook of Child Psychology, Vol. IV*. New York: John Wiley and Sons, pp.

Parrila, R., Kirby, J. R. and McQuarrie, L. (2004). Articulation rate, naming speed, verbal short-term memory, and phonological awareness: Longitudinal predictors of early reading development. *Scientific Studies of Reading, 8*, 3–26. https://doi.org/10.1207/s1532799xssr0801_2.

Passenger, T., Stuart, M. and Terrell, C. (2000). Phonological processing and early literacy. *Journal of Research in Reading, 23*(1), 55-66.

Patel, T. K., Snowling, M. J., De Jong, P.F. (2004) .A Cross-linguistic comparison of children learning to read in English and Dutch. *Journal of Educational Psychology*, 96 (4)785-797.

Pennington, B.F. and Lefly, D.L. 2001. Early reading development in children at family risk for dyslexia. *Child Development, 72*, 816–833.

Perfetti, C.A., Beck, I., Ball, L.C. and Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: a longitudinal study of first grade children. *Merrill-Palmer Quarterly, 33*, 283-319.

Plaut, D., McClelland, J. M., Seidenberg, M. S. and Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*: 56–115.

Porta, M. E. (2008). Hacia un modelo comprehensivo Del Aprendizaje lingüístico Inicial. Implementación del Modelo y Evaluación e sus Efectos sobre el Rendimiento Lector (Tesis doctoral no publicada). Universidad Nacional de Cuyo, Mendoza, Argentina.

Powell, D.,Stainthorp, R., Stuart, M., Garwood, H. and Quinlan, P. (2007). An experimental comparison between rival theories of rapid automatized naming performance and its relationship to reading. *Journal of Experimental Child Psychology, 98*, 46–68.

Prunet, J. F., Beland, R. and Idrissi, A. (2000). The mental representation of Semitic Words. *Linguistic inquiry, 31*, 609-638.

Puranik, C. and Apel, K. (2010). Effect of assessment task and letter writing ability on preschool children's spelling performance. *Assessment for Effective Intervention*, *36*(1), 46-56.

Ramirez, G., Chen, X., Geva, E. and Kiefer, H. (2010). Morphological awareness in Spanish-speaking English language learners: Within and cross-language effects on word reading. *Reading and Writing*, *23*(3-4), 337-358.

Rathvon, N. (2004). *Early reading assessment: A practitioner's handbook*. New York: Guilford Press.

Read, C. A., Zhang, Y., Nie, H., and Ding, B. (1986). The ability to manipulate speech sounds depends on knowing alphabetic writing. *Cognition, 24,* 31-44.

Ritchie, J. and Lewis, J. (2010). *Qualitative Research Practice*. London: Sage.

Riyadh City website Available online at: www.arriyadh.com [accessed: 15/05/2015].

Robertson, L.H. (2002). Parallel literacy classes and hidden strengths: Learning to read in English, Urdu and classical Arabic. *Reading Literacy and Language*, *36*(3), 119–126.

Rogoff, B. (1993). Aprendices del pensamiento. El desarrollo cognitivo en el contexto social. Buenos Aires: Paidós.

Rosner, J. (1999). *Phonological awareness skills program Test*. Austin, TX: PRO-ED.

Rosowsky, A. (2001). Decoding as a cultural practice and its effects on the reading process of bilingual pupils. *Language and Education, 15*(1), 56-70.

Rosowsky, A. (2010). "Writing it in English": script choices among young multilingual muslims in the UK. *Journal of Multilingual and Multicultural Development.*

Rubin, A. and Babbie, E. R. (2009). *Research methods for social work,* 2nd edn. London: Brooks/Cole.

Saffran, J. R., Aslin, R. N. and Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926-1928.

Saiegh-Haddad, E. (2003). Linguistic distance and initial reading acquisition: The case of Arabic diglossia, *Applied Psycholinguistics, 24*(3), 431-451.

Saiegh-Haddad, E. (2004). The Relevance of linguistic and sociolinguistic features of Arabic diglosssia to the acquisition of various reading skills in Arabic. British Dyslexia Association International Conference, (2004) Available online at: http://www.bdainternationalconference.org/2004/ presentations/sat_s6_c_1.shtml [accessed 21/03/15].

Saiegh-Haddad, E. (2005). Correlates of reading fluency in Arabic: Diglossic and orthographic factors. *Reading and Writing: An Interdisciplinary Journal, 18*, 559–582. https://doi.org/10.1007/ s11145-005-3180-4.

Saiegh-Haddad, E. (2011). Phonological processing in diglossic Arabic: The role of linguistic distance. *Perspectives on Arabic Linguistics*, *22*, 269–28.
Saiegh-Haddad, E. and Geva, E. (2008). Morphological awareness, phonological awareness, and reading in English-Arabic bilingual children. *Reading and Writing: An Interdisciplinary Journal*, *21*, 481-504.

Saiegh-Haddad, E. and Joshi, R. (eds.) *Handbook of Arabic literacy: insights and perspective*. Heidelberg: Springer.

Saunders, K. J. (2011). Designing instructional programming for early reading skills. In Fisher, W.W., Piazza, C. C. and Roane, H. S. (eds) *Handbook of Applied Behaviour Analysis.* New York; London: Guilford Press, pp.92-112.

Sawyer, D.J. and Fox, B.J. (eds). (2012). *Phonological awareness in reading: The evolution of current perspectives,* New York: Springer.

Scanlon, D. M., Anderson, K.L. and Sweeney, J.M. (2010). Early intervention for reading difficulties: the interactive strategies approach. New York: Guilford Press.

Schaefer, B., Fricke, S., Szczerbinski, M., Fox-Boyer, A. V., Stackhouse, J. and Wells, B. (2009). Development of a test battery for assessing phonological awareness in German-speaking children. *Clinical Linguistics and Phonetics*, *23*, 404–430.

Schatschneider, C., Carlson, C. D., Francis, D. J., Foorman, B. F. and Fletcher, J. M. (2002). Relationship of rapid automatized naming and phonological awareness in early reading development: Implications for the double-deficit hypothesis. *Journal of Learning Disabilities, 35,* 245–256.

Schiff, R. (2012). Transition from shallow to deep orthography: The role of vowelization in reading development of unvowelized scripts. Journal of Psycholinguistics Research.

Schiff, R. and Lotem, E. (2010). Effects of phonological and morphological awareness on children's word reading development from two socioeconomic backgrounds. *First Language, 31*(2) 139–163.

Schiff, R. and Saiegh-Haddad, E. (2018). Development and relationships between phonological awareness, morphological awareness and word reading in spoken and standard Arabic. *Frontiers in psychology*, *9*, 356.

Schneider, W., Roth, E., and Ennemoser, M. (2000). Training Phonological Skills and Letter Knowledge in Children at Risk for Dyslexia: A Comparison of Three Kindergarten Intervention Programs. *Journal of Educational Psychology, 92*, 284-295.

Scribner, S. and Cole, M. (1981). *The psychology of literacy*. Cambridge, Mass.: Harvard University Press.

Seidenberg, M. S. (2005). Connectionist models of word reading, *Current Direction in Psychological Science, 14,* 238-42.

Seidenberg, M. S. and McClelland, J. M. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*: 523–568.

Seidenberg, N. E., Seidenberg, M. S. and Iverson, I. A. (1972). The effect of kindergarten instruction in alphabet and numbers on first grade reading. *Journal of Learning Disabilities, 5,* 54-261.

Seymour, P. H. K., Aro, M and Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*, 143–174.

Share, D. L. (2004). Knowing letter names and learning letter sounds: A causal connection. *Journal of Experimental Child Psychology, 88*, 213-233.

Share, D.L., Jorm, A.F., Maclean, R. and Matthews, R. (1984). Sources of individual differences in reading acquisition. *Journal of Educational Psychology*, *76*, 1309–1324.

Shatil, E. and Share, D. L. (2003). Cognitive antecedents of early reading ability: a test of the modularity hypothesis. *Journal of Experimental Child Psychology*, 86 (1), 1–31.

Shaw, C., Brady, L-M and Davey, C. (2011). *Guidelines for research with children and young people.* London: National Children's Bureau.

Shiotsu, T. (2010). *Linguistic and processing factors in the reading test performances of Japanese EFL learners.* Cambridge: Cambridge University Press.

Silberberg, N.E., Silberberg, M.C. and Iverson, I.A. (1972). The effects of kindergarten instruction in alphabet and numbers on first grade reading. *Journal of Learning Disabilities, 5*, 54-261.

Smythe, I., Everatt, J., Al-Menaye, N., He, X., Capellini, S., Gyarmathy, E. and Siegel, L. (2008). Predictors of word-level literacy amongst grade 3 children in five diverse languages. *Dyslexia, 14*(3), 170–187.

Smythe, I., Everatt, J., Al-Menaye, N., He, X., Capellini, S., Gyarmathy, E., and Siegel, L.S. (2008). Predictors of word-level literacy amongst grade 3 children in five diverse languages. *Dyslexia*, **14** (3), 170– 187. doi:10.1002/dys.369

Snow, C.E., Burns, M.S. and Griffin, P. (eds.) (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Snowling, C.E. (2000). Dyslexia. 2nd edn. Oxford: Blackwell.

Stackhouse, J. and Wells, B. (1997). *Children's speech and literacy difficulties: a psycholinguistic framework.* London: Whurr.

Stackhouse, J., Vance, M., Pascoe, M. and Wells, B. (2007). *A compendium of speech and auditory procedures. Children's Speech and Literacy Difficulties*, 4, Chichester: John Wiley and Sons.

Stahl, S.A. and Murray, B.A. (1994). Defining phonological awareness and its relationship to early reading. *Journal of Educational Psychology, 86*, 221-234.

Stanovich, K. E. (1992). Speculations on the causes and consequences of individual differences in early acquisition. In P. B. Gough, L. E. Ehri, and R. Treiman (Eds.), Reading acquisition (pp. 307–342). Hillsdale, NJ: Erlbaum.

Stuart, M. and Coltheart, M. (1988). Does reading develop in a sequence of stages? *Cognition*, *30*, 139-181.

Suggate, S., Reese, E., Lenhard, W. and Schneider, W. (2014). The relative contributions of vocabulary, decoding, and phonemic awareness to word reading In English versus German. *Reading and Writing: An Interdisciplinary Journal, 27* (8), 1395–1412.

Sunseth, K. and Bowers, P. G. (2002). Rapid naming and phonemic awareness: Contributionsto reading, spelling, and orthographic knowledge. *Scientific Studies of Reading, 6*, 401–429.

Swanson, H. L., Trainin, G., Necoechea, D. M. and Hamill D. D. (2003). Rapid naming, phonological awareness, and reading: A meta-analysis of the correlation evidence *Review of Educational Research,* 73(4) 407-440.

Taha, H. (2013). Reading and spelling in Arabic: Linguistic and orthographic complexity. *Theory and Practice in Language Studies, 3*(5), 721-727.

Taibah, N.J. (2006). *Cross-Linguistic Transfer of Phonological Processing Skills in Bilingual Arabic/English Speaking Children*. Doctoral Thesis: University of Washington.

Taibah, N.J. and Haynes, C.W. (2011). Contributions of phonological processing skills to reading skills in Arabic-speaking children. *Reading and Writing: An Interdisciplinary Journal*, *24*, 1019–1042.

Taub, G.E. and Szente, J. (2012). The impact of rapid automatized naming and phonological awareness on the reading fluency of a minority student population. *Journal of Research in Childhood Education, 26*(4), 359-370.

Termsarasab, P., Thammongkolchai, T., Rucker, J.C and Frucht, S.J. (2015).The diagnostic value of saccades in movement disorder patients: a practical guide and review. *Journal of Clinical Movement Disorders*, **2**, 14

The World Factbook. (2013) Available online at: https://www.cia.gov/library/publications/download [accessed: 15/05/2015].

Tibi, S. (2010). Developmental hierarchy of Arabic phonological awareness skills. *International Journal of Special Education*, *25*(1), 27-33.

Tibi, S. and Kirby, J. R. (2018). Investigating phonological awareness and naming speed as predictors of reading in Arabic. *Scientific Studies of Reading, 22,* 70-84.

Tibi, S. and Kirby, J. R. (2019). Reading in Arabic: How well does the standard model apply?. *Journal of Speech, Language, and Hearing Research*, *62*(4), 993-1014.

Tomasello, M. (2000). *The cultural origins of human cognition*. Cambridge, MA: Harvard

Torgesen, J. and Bryant, B. (1994). *Test of Phonological Awareness*. Austin, TX: PRO-Ed.

Torgesen, J.K., Wagner, R.K., Rashotte, C.A., Burgess, S.R. and Hecht, S. (1997). Contributions of phonological awareness and rapid automatic naming ability to the growth of word-reading skills in second- to fifth-grade children. *Scientific Studies of Reading,* 1(2), 161–185.

Treiman, R. (1993). *Beginning To Spell*. New York: Oxford University Press.

Treiman, R. and Kessler, B. (2003). The role of letter names in the acquisition of literacy. In R.V. Kail (ED, *Advances in child development and behavior* (vol.31). San Diego, CA: Academic Press.

Treiman, R. and Kessler, B. and Pollo, T. C. (2006). Learning about the letter name subset of the vocabulary: Evidence from US and Brazilian preschoolers, *Applied Psycholinguistics*, 27,211-227.

Treiman, R. and Rodriguez, K. (1999). Yong children use letter names in learning to read words, *Psycological Science, 10*, 334-338.

Treiman, R. and Zukowski, A. (1991). Levels of phonological awareness. *In* S. Brady and D. Shankweiler (Eds.), *Phonological Processes In Literacy*. Hillsdale: Lawrence ErlbaTum, Pp. 67-84.

Treiman, R. and Zukowski, A. (1996). Children's sensitivity to syllables, onsets, rimes, and phonemes. *Journal of Experimental Child Psychology,* 61, 193-215.

Treiman, R., Cassar, M. and Zukowski, A. (1994). What types of linguistic information do children use in spelling? the case of flap. *Child Development*, *65*, 1318-1337.

Troia, G. A., Roth, F. P. and Yeni-Komshian, G. H. (1996). Word frequency and age effects in normally developing children's phonological processing. *Journal of Speech, Language and Hearing Research*, *39*(5), 1099-1108.

United Nations Development Programme. (2003). *Arab human development report 2003: Building a knowledge society*. New York: UNDP, Regional Bureau for Arab States. Available online at: http://www.miftah.org/Doc/Reports/Englishcomplete2003.pdf [accessed 12/03/2015].

Vaessen A, Gerretsen P, Blomert, L. (2009). Naming problems do not reflect a second independent core deficit in dyslexia: Double deficits explored. *Journal of Experimental Child Psychology, 103*(2): 202–221.

Van den Bos, K. P. (1998). IQ, phonological awareness and continuous naming speed related to Dutch poor decoding children's performance on two word identification tests. *Dyslexia, 4*, 73–89.

Vellutino, F. R., Tunmer, W. E., Jaccard, J. J. and Chen, R. (2007). Components of reading ability: Multivariate evidence for a convergent skills model of reading development. *Scientific Studies of Reading, 11*(1), 3–32.

Versteegh, K. (2014). The Arabic language. Edinburgh: Edinburgh University Press.

Wagner, D.A. (1994). *Literacy, Culture and development: becoming literate in Morocco*. New York, NY: Cambridge University Press.

Wagner, R. K., Torgesen, J. K. and Rashotte, C. A. (1994). Development of reading related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology*, *30*, 73–87.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., Donahue, J. and Garon, T. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: a 5-year longitudinal study. *Developmental Psychology, 33*, 468-479.

Wagner, R.K. and Torgesen, J.K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*, 192–212.

Wagner, R.K., Torgesen, J.K. and Rashotte, C. (1999). *Comprehensive Test of Phonological Processing (CTOPP)*. Austin, TX: PRO-ED.

Wallach, L., Wallach, M., Dozier, M. and Kaplan, N. (1977). Poor children learning to read do not have trouble with auditory discrimination but do have trouble with phoneme recognition. *Journal of Educational Psychology*, *69*, 36–39.

Walonick, D. (2004). *Survival Statistics*, Bloomington: Statpac.

Walsh, D.J., Price, G.G. and Gillingham, M.G, (1988). The critical but transitory importance of letter naming. *Reading Quarterly*, *23*,108-122.

Warmington, M. and Hulme, C. (2012). Phoneme awareness, visual-verbal paired-associate learning, and rapid automatized naming as predictors of individual differences in reading ability. *Scientific Studies of Reading, 16*(1), 45-62.

Wesseling, R. and Reitsma, P. (2000). The transient role of explicit phonological recoding for reading acquisition. *Reading and Writing: An Interdisciplinary Journal, 13*, 313–336.

Wild, S. (2006). Reading. In Leaman, O. *The Qur'an: An encyclopaedia* London; New York: Routledge, pp.635-638.

Wimmer, H., Landerl, C. and Frith, U. (1999). Learning to read German: Normal and impaired acquisition. In M. Harris and G. Hatano (Eds.), Learning to read and write: A cross-linguistic perspective (pp. 34–50). Cambridge, UK: Cambridge University Press.

Wimmer, H., Landerl, K., Linortner, R. and Hummer, P. (1991). The relationship of phonemic awareness to reading acquisition: More a consequence than a precondition. *Cognition*, *40*, 219-249.

Wimmer, H., Mayringer, H. and Landerl, K. (2000). The double deficit hypothesis and difficulties in learning to read a regular orthography. *Journal of Educational Psychology, 92*, 668–680.

Wium, A.M., Louw, B. and Eloff, I. (2011). Evaluation of a programme to support foundation-phase teachers to facilitate literacy. *South African Journal of Communication Disorders*, *58*(2), 72-78.

Wolf, M. and Bowers, P. (1999a). The question of naming - speed deficits in developmental reading disabilities: an introduction to the double-deficit hypothesis. *Journal of Educational Psychology, 19,* 1-24.

Wolf, M. and Bowers, P. G. (1999b). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology*, *91*, 415–438.

Xhafaj, D.C.P. (2011). L1 literacy and L2 learning: Connecting the dots. Doctoral thesis, Federal University of Santa Catarina.

Yafi, W. S. (2012). *Inevitable democracy in the Arab world*: *New realities in an ancient land*. New York: Palgrave Macmillan.

Yopp, H. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly, 23*, 159-177.

Yopp, H. (1992). Developing phonemic awareness in young children. *The Reading Teacher, 45*(9), 696-703.

Zaid, M. A. (2011). Language acquisition, linguistic creativity and achievement: Insight from the Qur'an. *Kemanusiaan*, 1*8*(2), 75-100. Available online at: web.usm.my/…ol%2018.2.2011/KAJH%2018.2.2011%204.pdf [accessed 12/10/2017]/

Zayed, A.M., Roehrig, A.D., Arrastia-Lloyd, M.C. and Gilgil, N.M. (2013). Phonological awareness and working memory in Arabic-speaking Egyptian pre-school children at risk for dyslexia. *International Journal for Psychological Studies, 5*(1), 139-149.

Ziegler, J. C. and Goswami, U. (2005).Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131* (1), 3-29.

Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L. and Blomert, L. (2010). Orthographic depth and its impact on universal predictors of reading: A cross-language investigation. *Psychological Science, 21*, 551–559.

Zoccolotti, P., De Luca, M., Di Pace, E., Gasperini, F., Judica, A. and Spinelli, D. (2005). Word length effect in early reading and in developmental dyslexia. *Brain and Language, 93*, 369–373.

APPENDICES

Appendix A: Approval letter from Saudi Ministry of Education

[redacted]

Appendix B: Letter of consent for schools (Study One) (translation)

Re: An investigation of socioeconomic effects on phonological awareness and lexical retrieval in Saudi children

Dear Head of School

I am a qualified speech therapist with experience in assessing and working with young children. Currently, I am completing a PhD research project in the Department of Language and Communication Science in London. As part of my studies, I am carrying out a research project to investigate the performance of typically developing Saudi children from different socioeconomic backgrounds in completing various tasks (Phonological Awareness, Rapid Naming and Letter Knowledge) which may provide us with information about children's reading abilities. These tasks have been developed to assess pre-school children in order to identify those at risk of literacy problems. They have been used with a number of children in a previous research project at the King Fahad Academy, London and children found them fun. Before professionals can use them to identify children at risk of literacy and communication problems, we need to know how typically developing Saudi children from different social and economic backgrounds perform on these tests.

We are inviting you to take part in this study which will involve my meeting four- to seven-year-old typically developing children in your school. If you wish to take part in this study, we would appreciate your help in selecting children to participate in the study, and sending information sheets and consent forms to parents.

Children taking part in this study must meet the following criteria:

- Have Saudi nationality with both parents also Saudi.
- Have Arabic as the first language.

The assessment tasks would need to be carried out in a quiet room at your school. Each child would be seen individually in a session lasting up to 40 minutes. The child's mother or teacher may attend the session. Parents will also be asked to fill in a short questionnaire about their child and their family, to provide information about children's health and their language, social and economic background.

After establishing rapport with the child, the researcher will introduce the assessment tasks which take the form of a game. I will ask the child to say which word does not rhyme with a set of other words, whether or not two words have the same initial, and to say the name or sound of letters. Pictures are incorporated in the tasks to stimulate children's interest. If you have any queries or are interested in seeing the full version of the different tasks, please feel free to contact me.

I will only include children who are willing to join the researcher and participate in the activities. Children have been found to enjoy these activities; however if the child does not want to go on, we will stop the session and return the child to an appropriate member of staff.

All information provided will be confidential and will be kept in a locked and secure location until the end of the research. Each child participating in this study will be given a number and the list of names and numbers will be kept in a secure file separately from any information about the child.

If the assessments raise any concerns regarding the child's development, we will discuss these concerns with the child's teacher/parents.

At the end of the project, a short report on the research will be made available to you if you are interested.

You have the right to withdraw at any time from the study without having to give a reason.

Thank you for giving this your consideration.

If you have any further questions or would like to discuss any part of the study, please do not hesitate to contact us.

Investigator

Ghada Najmaldeen

Project supervisors:

[redacted]

Appendix C: Parental information and consent form (Study One) (translation)

If there is an aspect of the study which concerns you, you may make a complaint. The City University has established a complaints procedure via the Secretary to the Research Ethics Committee. To complain about the study, you need to phone [redacted] You can then ask to speak to the Secretary of the Ethics Committee and inform them about the name of the project.

You can also write to the Secretary at:

[redacted]

Secretary to Senate Research Ethics Committee
CRIDO
City University
Northampton Square
London
EC1V 0HB

[redacted]

Alternatively, you can contact:

[redacted]

Department of Rehabilitation Sciences
College of Applied Medical Sciences
King Saud University

[redacted]

Further Information

If you have any further concerns or questions, please do not hesitate to contact us.

Project Supervisor:

[redacted]

City University

Northampton Square

London

EC1V 0HB

Investigator:
Ghada Najmaldeen

[redacted]

CITY UNIVERSITY LONDON

An invitation for your child to take part in a research project

An investigation of socioeconomic effects on phonological awareness and lexical retrieval in Saudi children

About the research project

I'm a speech language pathologist interested in finding ways of identifying children at risk of literacy problems.

I'm interested in running assessments that have been developed in the UK to assess pre-school children's abilities, in order to provide preliminary information about the performance of Arabic speaking children on these assessments.

The tests have been used in the UK with a large number of children and the children found them fun.

We need to know how typically developing Saudi children from different social and economic backgrounds perform on these tests before professionals can use them to identify children at risk of literacy and communication problems.

What is involved

I aim to see 280 children whose parents agree to their participation, and who are themselves willing to participate. A researcher will see each child in the school in a session lasting up to 40 minutes. The child's mother or teacher may attend the session.

After a brief warm-up, the researcher will introduce each of the tests which take the form of games. During these games the researcher will score the child's performance.

In addition, we are asking parents to compete a short questionnaire about their child and their family, to provide information about children's health and their language, social and economic background.

If you are willing for your child to participate in the project, we would be grateful if you would fill in the attached consent form, and return this to the head of the school.

I will only include children if they agree to join the researcher and participate in the activities. Children have been found to enjoy these activities, but if your child does not want to go on, we will stop the session and take the child back to an appropriate member of nursery staff.

Each child participating in this study will be given a number and the list of names and numbers will be kept in a secure file separately from any information about the child. All information provided will be confidential and will be kept in a locked and secure location until the research is done.

At the end of the project, a short report on the research will be made available to you if you are interested.

If the assessments raise any concerns regarding your child's development, we will discuss these concerns with your child's teacher.

Your child does not have to take part in this study if you do not want them to, and even if you agree to your child taking part, you may withdraw them at any time without having to give a reason.

Many thanks for giving this your consideration.

CITY UNIVERSITY
LONDON

Informed Consent Form

Project title: An investigation of socioeconomic effects on phonological awareness and lexical retrieval in Saudi children

Supervisor: [redacted]

Investigators: Ghada Najmaldeen

I agree that my child…………………………………………….(full name of child) may take part in the above City University research project. I have read the Explanatory Statement, had the opportunity to ask questions, discuss the study, and received enough information about the study.

I understand that any information provided on my child is confidential, and that no information that could lead to the identification of any individual will be disclosed in any reports on the project, or to any other party. No identifiable personal data will be published. The identifiable data will not be shared with any other organization.

I also understand that my child's participation in the project is voluntary that I can withdraw my child's from the study without being penalized or disadvantaged in any way.

I do/do not wish you to give feedback on my child's performance to my child's teacher (please delete as appropriate)

Parent's Name:……………………………………………………………………………… Date:………………………………

Signature:……….

(cc. researcher)

Appendix D: Student information sheet and consent form (translation)

Information Sheet

An investigation of socioeconomic effects on phonological awareness and lexical retrieval in Saudi children

What is the project?

We are going to play some games with words, numbers and letter. We will use coloured pictures

What do I have to do?

You will be asked if you want to take part in the project. Your parents will also talk to you about it.

If I say yes what will happen?

You will be asked some questions about words, numbers and letters. Using pictures you will sign a form about this if you say **YES**.

What will I have to do?

You will be asked to:

Take part by answering some question about words, letters and numbers, using games.

Where will I have to go?

You will be seen at your school.

Is there anything else I have to do?

If you want to take part in the project then we will ask you if you want to see us again.

If you don't want to take part you can say no at any time.

If I have understood what will happen I can colour a picture to show you my choice:

Yes: I do want to take part

No: I do not want to take part in the project

Would you like to take part in our project?

Yes: I want to take part in the project 

No: I do not want to take part in the project 

Please colour the Yes picture  or the No picture  to show us your choice.

Appendix E: Parental questionnaire (pilot) (translation)

PARENTAL QUESTIONNAIRE

Phonological awareness and lexical retrieval in Saudi children

Child's number (to be filled in by researcher): ……………………………………

Date: ……………………………………

Date of Birth: ……………………………………

Phone number: ……………………………………

Original area your child belongs to:

- Central
- Western
- Eastern
- Northern
- Southern

Name of the Quarter: ……………………………………

- North
- East
- South
- West

Name of the school: ……………………………………

Type of school (private/public): ……………………………………

Please choose one answer to each of the following questions:

1. Gender:
   - ☐ Boy
   - ☐ Girl

2. What is your child's age?
   - ☐ 4- 4.11
   - ☐ 5 – 5.11
   - ☐ 6 – 6.11
   - ☐ 7 years old

3. Who currently lives at home with your child?
   - ☐ Mother
   - ☐ Father
   - ☐ Both of them
   - ☐ Parents and grandparents or one of them
   - ☐ Parents and whole family (grandparents, uncles…)
   - ☐ other (specify),……………………

4. Who is your child's primary carer?
   - ☐ Mother
   - ☐ Father
   - ☐ Live-in nanny/domestic help
   - ☐ Other.

5. Do you have a live-in nanny/ domestic help?
   - ☐ Yes
   - ☐ No

6. If there is a live-in nanny, how much time does your child spend with her?
   - ☐ None at all
   - ☐ 2 to 4 hours daily
   - ☐ 5 to 7 hours daily
   - ☐ 8 to 10 hours daily
   - ☐ More than 10 hours daily

7. What language does your live-in nanny/domestic help speak with your child?
   - ☐ English
   - ☐ Arabic
   - ☐ Arabic and English
   - ☐ Other languages (e.g. Maid's native language)

8. What language does your family speak with your child?
   - o English
   - o Arabic
   - o Arabic and English
   - o Other languages

9. How many of your child's siblings currently live at home?
   - ☐ none
   - ☐ 1
   - ☐ 2
   - ☐ 3
   - ☐ 4

- ☐ 5
- ☐ more than 5

10. What is your child's order in your family?
    - ☐ 1
    - ☐ 2
    - ☐ 3
    - ☐ 4
    - ☐ 5
    - ☐ other

11. What is your child's language background?
    - ☐ Arabic only
    - ☐ Arabic and other language

    Please indicate languages used:……………………………………………………………………………

12. Does your child learn another language?
    - ☐ Yes
    - ☐ No

    If yes,
    - ☐ At school
    - ☐ At home with parents
    - ☐ Private classes
    - ☐ At school and home
    - ☐ While travelling
    - ☐ All the above

13. What is the highest level of education completed by the child's mother?
    - ☐ No formal qualifications
    - ☐ Elementary
    - ☐ Intermediate
    - ☐ High school
    - ☐ Bachelor
    - ☐ Postgraduate studies

14. Does the child's mother work?
    - ☐ Yes
    - ☐ No
    - ☐ Sometimes
    - ☐ Retired

    If the child's mother works, what is her occupation?……………………………………………………

    What is the field?
    - ☐ Education
    - ☐ Medical/Health
    - ☐ Administration
    - ☐ Retail/commerce
    - ☐ Engineering
    - ☐ Other:……………………………………………………………………………………………

15. What is the highest level of education completed by the child's father?
    - ☐ No formal qualifications
    - ☐ Elementary
    - ☐ Intermediate
    - ☐ High school
    - ☐ Bachelor
    - ☐ Postgraduate studies

16. Does the child's father work?
    - ☐ Yes
    - ☐ No
    - ☐ Sometimes
    - ☐ Retired

    If the child's father works, what is his occupation?……………………………………………………

    What is the field?
    - ☐ Education
    - ☐ Medical/Health
    - ☐ Administration
    - ☐ Military
    - ☐ Retail/commerce
    - ☐ Engineering
    - ☐ Other:……………………………………………………………………………………………

17. How much is the family income in Saudi Riyals?

    The range of the family income:
    - ☐ Less than 3 000 SR
    - ☐ Between 3 000 – 5 999 SR
    - ☐ Between 6 000 – 8 999 SR
    - ☐ Between 9 000 – 11 999 SR
    - ☐ Between 12 000 – 14 999 SR
    - ☐ Between 15 000 – 17 999 SR
    - ☐ Between 18 000 – 20 999 SR
    - ☐ Over 21 000 SR

18. What type of property do you live in?
- ☐ Flat
- ☐ Villa
- ☐ Palace
- ☐ Other……………………………………………………………………………………………

19. Is your property
- ☐ Rented?
- ☐ Owned
- ☐ Provided as part of your employment entitlements?

20. Does your child have access to books at home?
- ☐ Yes
- ☐ No

21. Does your child watch TV?
- ☐ No
- ☐ Yes, please specify:
  - ☐ Sometimes
  - ☐ Once a week
  - ☐ Once a day
  - ☐ Most of the day

22. Does your child play video games?
- ☐ Yes
- ☐ No

23. Does your child use the Internet?
- ☐ Yes
- ☐ No

24 What other types of activities does your family use to promote your child's language and reading skills?

Helping your child with homework
- ☐ Yes
- ☐ No

Providing private tutoring
- ☐ Yes
- ☐ No

Computer software
- ☐ Yes
- ☐ No
- ☐ Other……………………………………………………………………………………………

25. Does your child receive extra tuition in Qur'anic studies?
- ☐ Yes
- ☐ No

The following section will be detached from the questionnaire as soon as we receive this:

Child's name:……………………………………………………………………………………………

Your name:……………………………………………………………………………………………

Signed:…………………………………………………… Date……………………………

THANK YOU VERY MUCH FOR COMPLETING THIS QUESTIONNAIRE

Appendix F: Parental questionnaire (final) (translation)

Child's identifier (to be filled in by researcher): ....

Child's name: .... Gender:.................. Date of Birth:...............

Phone number:………………………… Residential District:…………………………...

Type of school (private/public):……. School Area:……………………………………..

Please choose one answer:

1 What is your child's language background?
- ☐ Arabic only
- ☐ Arabic and other language (please indicate languages used):…………………………………………

2 Does your child learn any other languages?
☐ No ☐ Yes (please specify)
…………………………………………….

3 What is the highest level of education completed by your child's mother?
- ☐ No qualifications
- ☐ Elementary
- ☐ Intermediate
- ☐ High school
- ☐ Diploma
- ☐ Bachelor
- ☐ Postgraduate studies

4 Is the child's mother in paid employment?
☐ No ☐ Yes

5 What is the highest level of education completed by your child's father?
- ☐ No qualifications
- ☐ Elementary
- ☐ Intermediate
- ☐ High school
- ☐ Diploma
- ☐ Bachelor
- ☐ Postgraduate studies

6 Is the child's father in paid employment?
☐ No ☐ Yes

7 How much is the family's income in Saudi Riyals per month? Please specify a range:

- ☐ Less than 3000 SR
- ☐ Between 3000 – 5999 SR
- ☐ Between 6000 – 8999 SR
- ☐ Between 9000 – 11999 SR
- ☐ Between 12000 – 14999 SR
- ☐ Between 15000 – 17999 SR
- ☐ Between 18000 – 20999 SR
- ☐ Between 21000- 24999 SR
- ☐ Between 25000-29000 SR
- ☐ Between 30000-34000 SR
- ☐ Above 35000 SR

8 What is type of property do you live in?

- ☐ Flat
- ☐ Detached house
- ☐ Palace
- ☐ Part of detached house
- ☐ Other (please specify) ....

9 Is the property

- ☐ Rented?
- ☐ Owned?
- ☐ Provided free as part of your job entitlement?

10. Does your child have access to books at home?
☐ No ☐ Yes

11. Does your child watch children's programmes on TV?
☐ No ☐ Yes

12. Does your child play video games?
☐ No ☐ Yes

13. Does your child use the Internet?
☐ No ☐ Yes

14. Do you help your child with homework?
☐ No ☐ Yes

15. Does your child receive private tutoring
☐ No ☐ Yes

16. Does your child use educational computer software?
☐ No ☐ Yes

17. Does your child receive extra Qur'an sessions?
☐ No ☐ Yes

Appendix G: Test battery scoring form

PHONOLOGICAL AWARENESS TEST BATTERY

SCORING FORM

Name	
Child Number	
Phone Number	
School	
Gender	

	Year	Month	Day
Date Tested			
Date of Birth			
Age			

Phonological awareness

Rhyme Awareness

Demonstration

فاس	حوت – فول - ناس
حار	فيل – نار - كيك

	Stimuli Words	Response Items	Score(1 or 0)
1	لوز	تور – خيط – موز	
2	دم	فم – دب - ر ز	
3	طوق	توم– فوق - توب	
4	خط	جد – ز ر – بط	
5	فيش	عين – ريش – ديك	
6	خد	يد - بر – سين -	
7	راس	بيت – شال – كاس	
8	خيل	ذيل – بيض – صوف	
		Total	

Syllable Segmentation

Demonstration

صدى	2
شمام	2

	Stimulus Word	No. of syllables	Response	Score(1 or 0)
1	دنيا	2		
2	ابهام	2		
3	دلو	2		
4	شماعة	3		
5	مجرفة	3		
6	مروحة	3		
7	جاذبية	4		
8	مدفعية	4		
			Total	

Alliteration Awareness

Demonstration

باب	بط	بيت	**سيف**
فم	**لوز**	فار	فيل

	Stimuli Words				Score (1 or 0)
1	ديك	دب	بيض	دم	
2	نسر	نحل	نخل	جرس	
3	مكتب	فستان	مفتاح	مسجد	
4	عسل	علم	عنب	سمك	
5	شوكه	شراب	حصان	شعر	
6	درج	جبن	جمل	جزر	
7	مجلة	مخده	شجرة	مثلث	
8	برتقال	مسطرة	منشفه	مطرقة	
				Total	

Phoneme Isolation Task

Demonstration

خس
شمس

	Stimuli Words	Response	Score (1 or 0)
1	فراش		
2	مقص		
3	أرنب		
4	شنطة		
5	لمبة		
6	كتاب		
7	عصفور		
8	لوحة		
		Total	

Blending Task

Demonstration

ثعبـــان	ثع___بان
هـــاشـــم	ها___شم
ر أ س	ر___أس
بـــيـــت	ب___يت
فـــي	ف___ي
د ا ر	د_ا_ر

	Stimuli Words	Response	Score(1 or 0)
1	زيـ___تون		
2	فسـ_تان		
3	أر___سـل		
4	أ___م		
5	طـ__يـر		
6	سـ___ور		
7	شـ_اة		
8	نـ___ار		
9	قـ___ط		
10	ب_ي_ع		
11	نـ_و_ر		
12	ـ_يـ_نـ_ـة سـ_		
13	قـ_ر_د		
14	ـ_ر عـ_صـ_ا_فـ_		
15	و_ر_ سـ_بـ		
16	أ_لـ_و_ا_ن		
17	عـ_نـ_كـ_بـ_و_ت		
18	تـ_تـ_طـ_ا_بـ_ق		
19	ا_سـ_تـ_هـ_تـ_ا_ر		
20	قـ_سـ_طـ_نـ_طـ_يـ_نـ_يـ_ـة		
		Total	

Letter Knowledge Test

	Letter	Response	Score (1 or 0)
1	أ		
2	ب		
3	ت		
4	ث		
5	ج		
6	ح		
7	خ		
8	د		
9	ذ		
10	ر		
11	ز		
12	س		
13	ش		
14	ص		
15	ض		
16	ط		
17	ظ		
18	ع		
19	غ		
20	ف		
21	ق		
22	ك		
23	ل		
24	م		
25	ن		
26	ه		
27	و		
28	ي		
			Total:

Rapid Naming Task

Rapid Colour Naming Test

<u>Form A</u>

أزرق أحمر أخضر أسود بني أصفر أحمر أسود أزرق

أصفر أخضر بني أزرق أحمر أخضر أسود أصفر بني

أخضر أصفر أسود أحمر بني أزرق أخضر أحمر أزرق

أسود بني أصفر بني أخضر أحمر أصفر أزرق أسود

Time:.................. Errors:.................

<u>Form B</u>

اسود أزرق أصفر أحمر أخضر بني أصفر بني أسود

أزرق أحمر أخضر أزرق بني أحمر أسود اصفر أخضر

بني أصفر أسود أخضر أحمر أزرق بني أخضر أصفر

أزرق أسود احمر أصفر بني أسود أخضر أحمر ازرق

Time:.................. Errors:.................

Rapid Object Naming Test

<u>Form A</u>

وردة كورة أرنب سمكة يـد بطة باب مفتاح بيت

أرنب يـد سمكة بيت شجرة باب قلم كرسي مفتاح

باب شجرة بيت وردة أرنب قلم كورة بطة كرسي

شجرة قلم مفتاح كرسي بطة سمكة وردة كورة يـد

Time:.................. Errors:.................

<u>Form B</u>

يـد كورة مفتاح شجرة بطة كرسي وردة قلم سمكة

كرسي بطة كورة قلم أرنب مفتاح بيت سمكة بيل

وردة كرسي قلم باب سمكة بيت شجرة يـد أرنب

بيت وردة باب بطة يـد شجرة ارنب كورة مفتاح

Time:.................. Errors:.................

Rapid Letter Naming

Form A

ك	ن	ل	ك	ز	م	ي	ل	ن
ن	ي	م	ل	ز	ك	ي	م	ز
ي	ك	ل	م	ي	ن	ك	ز	ل
م	ل	ن	ز	ك	ي	ن	م	ز

Time:……………… Errors:………………

Form B

ز	م	ن	ي	ك	ز	ن	ل	م
ل	ز	ك	ن	ي	م	ل	ك	ي
ز	م	ي	ك	ز	ل	م	ي	ن
ن	ل	ي	م	ز	ك	ل	ن	ك

Time:……………… Errors:………………

Rapid Digit Naming

Form A

1	4	7	1	6	5	9	7	4
4	9	6	7	5	1	5	6	9
9	1	7	6	9	4	1	5	7
6	7	4	5	1	9	4	6	5

Time:……………… Errors:………………

Form B

5	6	4	9	1	5	4	7	6
7	5	1	4	9	6	7	1	9
9	6	5	1	5	7	6	9	4
1	4	7	9	5	6	1	7	4

Time:……………… Errors:………………

Appendix H: Table 8.1: Mean, standard deviation and ANOVA results of scores over all measures, using six-month age bands

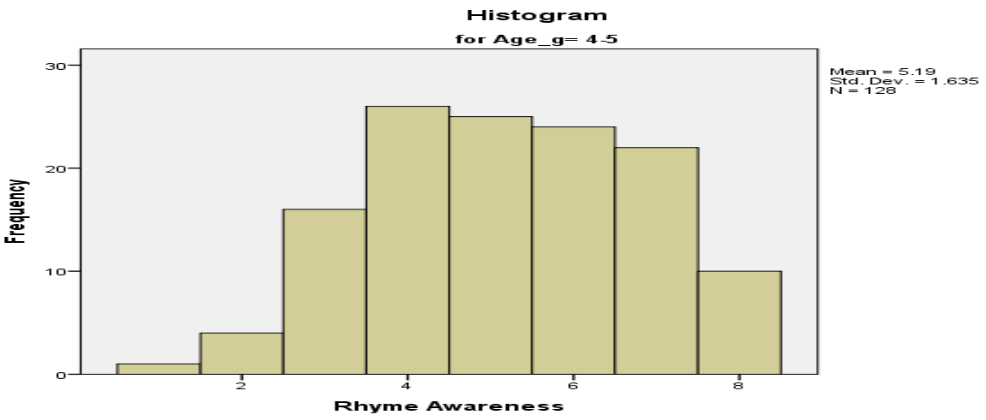
Measure / Age	4;0-4;5 (*n*=55)	4;6-4;11 (*n*=68)	5;0- 5;5 (*n*=66)	5;6-5;11 (*n*=62)	6;0-6;5 (*n*=79)	6;6- 7;0 (*n*=54)	
Rhyme Awareness	M=5.20 (1.59)	M=5.15 (1.70)	M=5.86 (1.63)	M=6.00 (1.34)	M=6.10 (1.18)	M=5.98 (1.37)	F (5.39) *** p (0.00)
Syllable Segmentation	M= 4.75 (1.71)	M=4.85 (1.53)	M=5.32 (1.49)	M=5.56 (2.15)	M=6.03 (1.67)	M=5.78 (1.72)	F (6.62) *** p (0.00)
Alliteration Awareness	M=2.80 (1.27)	M=2.62 (1.04)	M=3.27 (1.61)	M=3.35 (1.67)	M=4.18 (1.92)	M=3.78 (1.88)	F (9.12) *** p (0.00)
Phoneme Isolation	M=1.80 (1.88)	M=1.69 (1.92)	M=2.94 (2.83)	M=3.44 (2.84)	M=4.95 (2.23)	M=4.35 (2.69)	F (19.62) *** p (0.00)
Blending	M=5.35 (1.92)	M=6.00 (2.52)	M=5.65 (2.34)	M=6.84 (2.18)	M= 8.18 (3.01)	M=8.13 (7.32)	F(15.69) *** p(0.00)
LK	M=3.98 (6.23)	M=3.94 (6.08)	M=7.20 (8.55)	M=10.37 (9.62)	M=18.00 (8.97)	M=15.50 (10.35)	F (32.93) *** P (0.00)

***F-value is significant at the 0.001 level (p < 0.001)

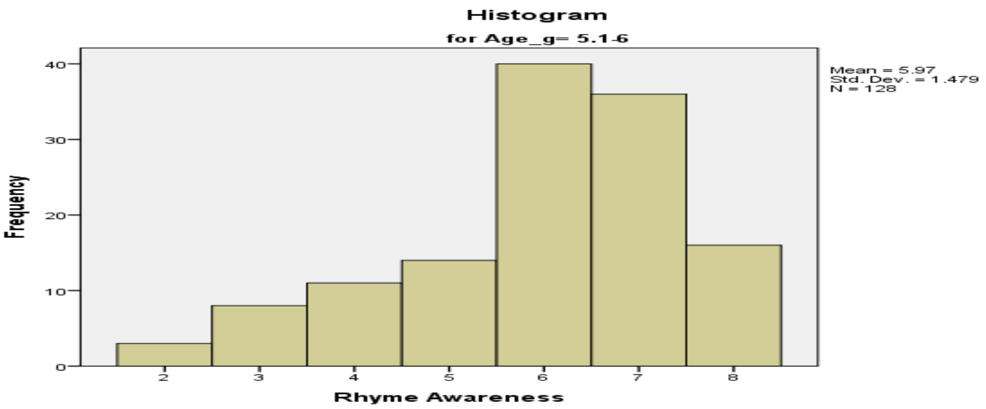
Appendix I: Graphical means for testing normality

Figure 8.1: Graphical means for testing normality: Rhyme Awareness

Age Group 4;0-4;11



Age Group 5;0-5;11



Age Group 6;0-7;0

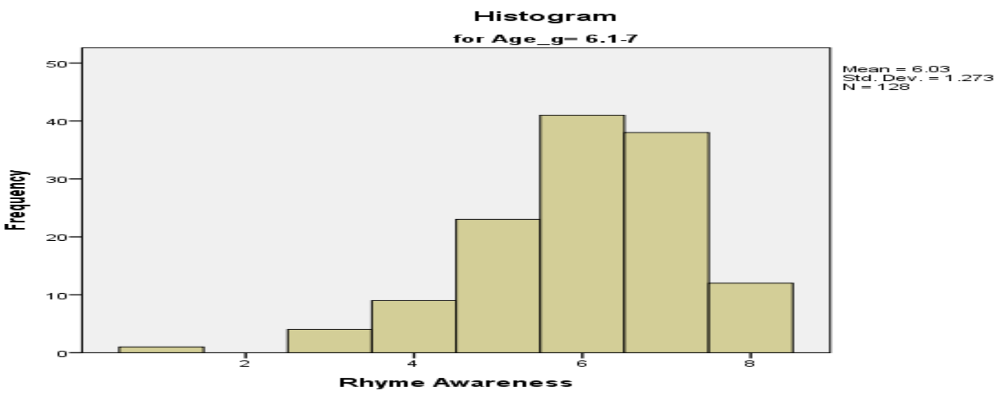
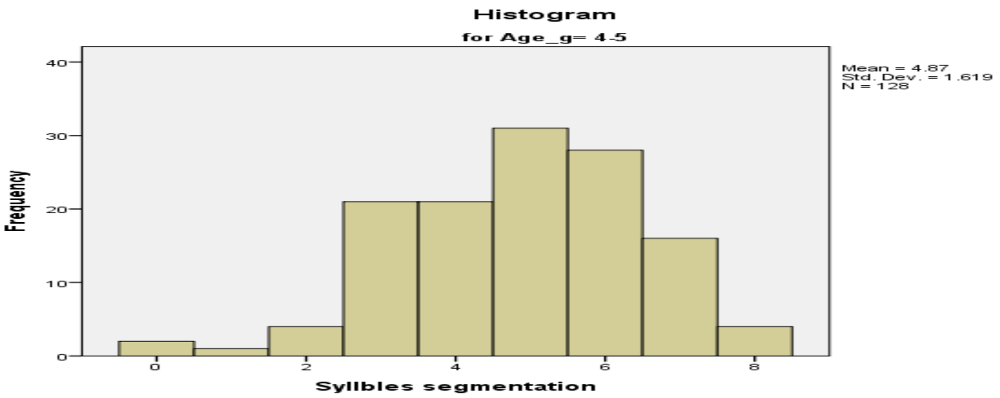
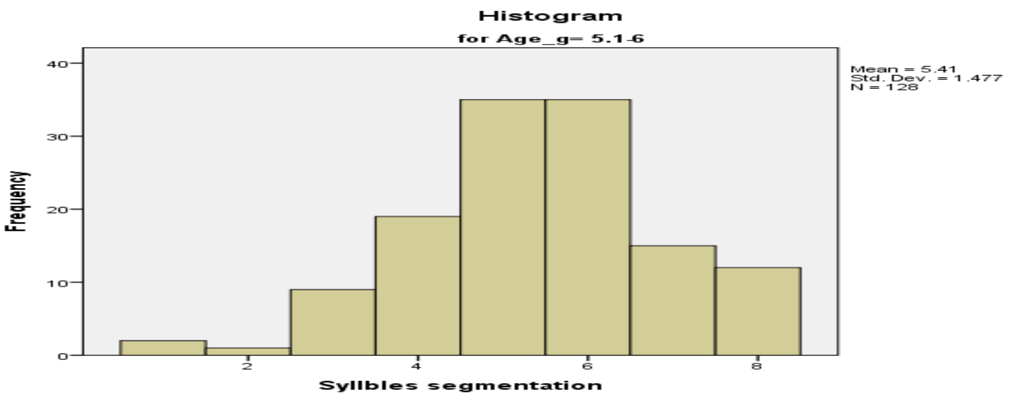


Figure 8.2: Graphical means for testing normality: Syllable Segmentation

Age Group 4;0-4;11



Age Group 5;0-5;11



Age Group 6;00-7;0

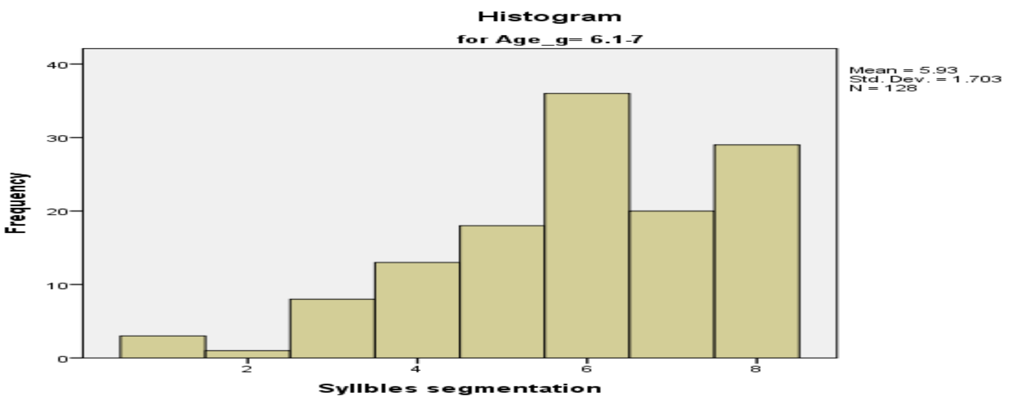
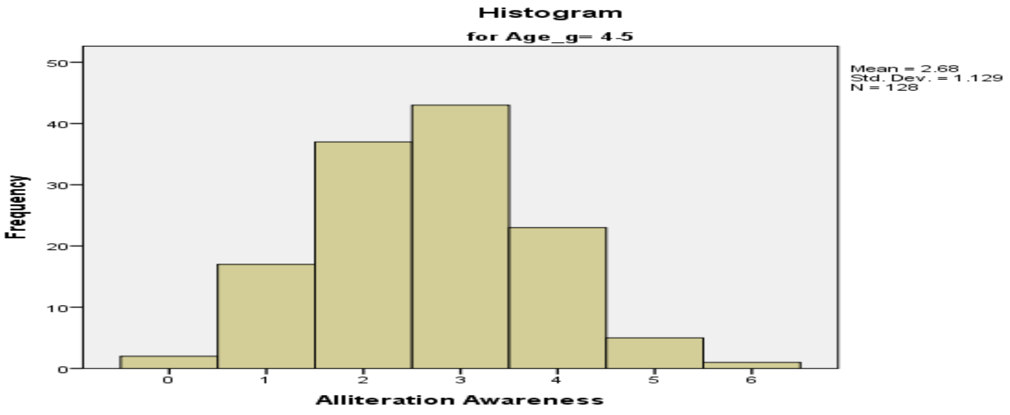
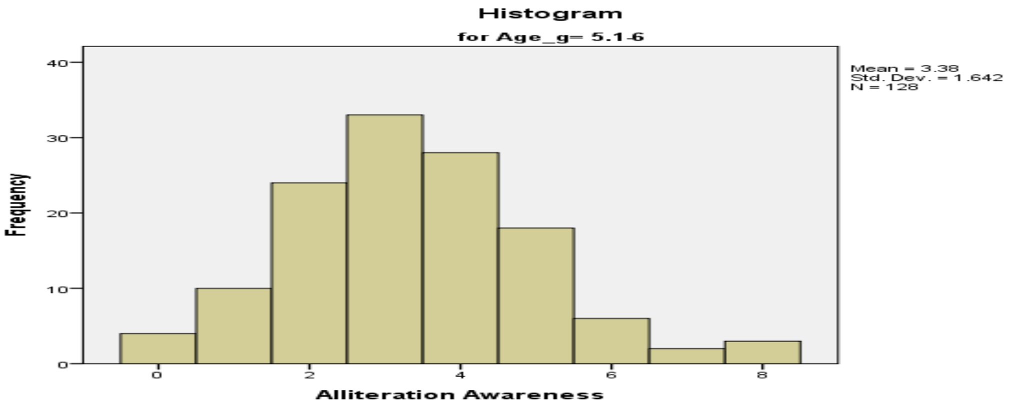


Figure 8.3: Graphical means for testing normality: Alliteration Awareness

Age Group 4;0-4;11



Age Group 5;00-5;11



Age Group 6;0-7;0

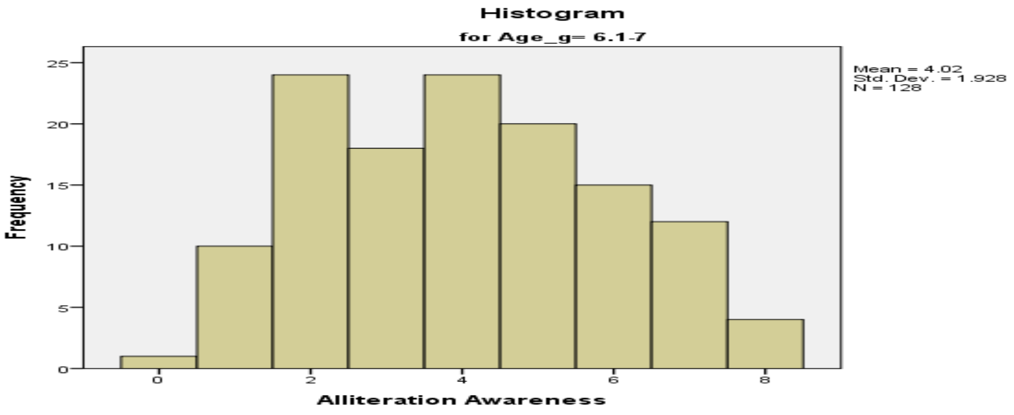
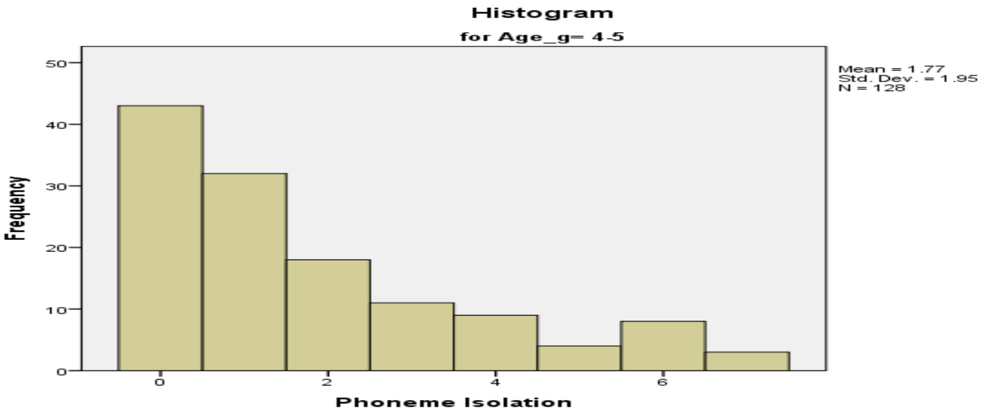
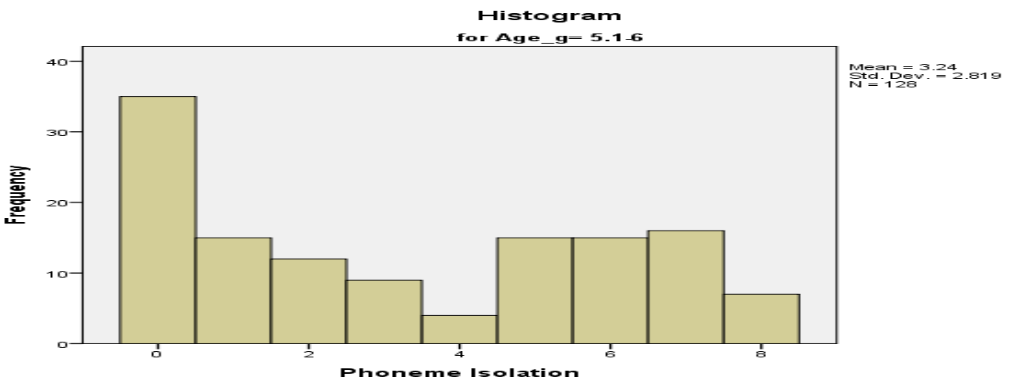


Figure 8.4: Graphical means for testing normality: Phoneme Isolation

Age Group 4;0-4;11



Age Group 5;0-5;11



Age Group 6;0-7;0

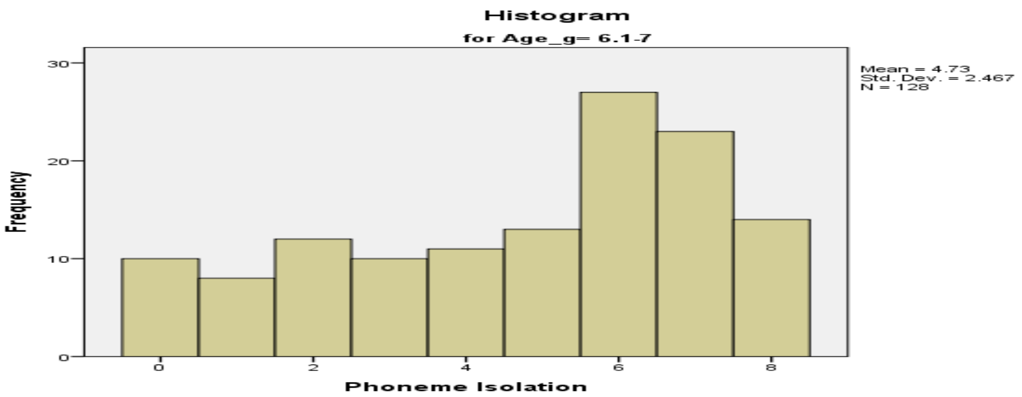
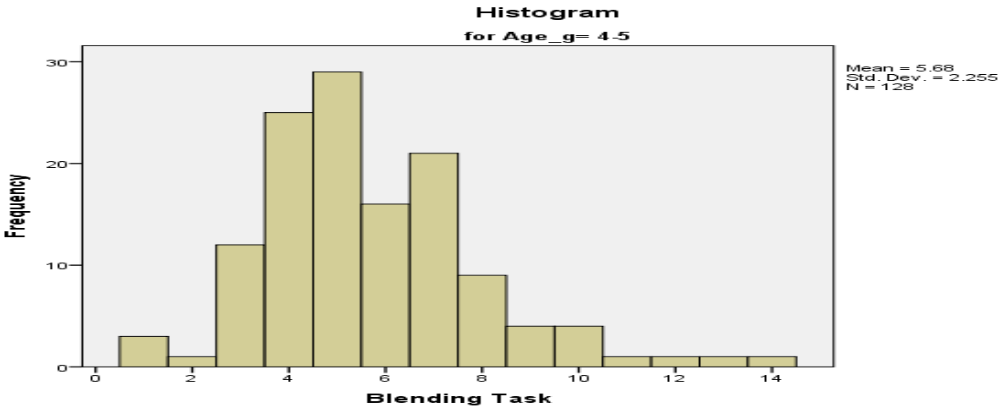
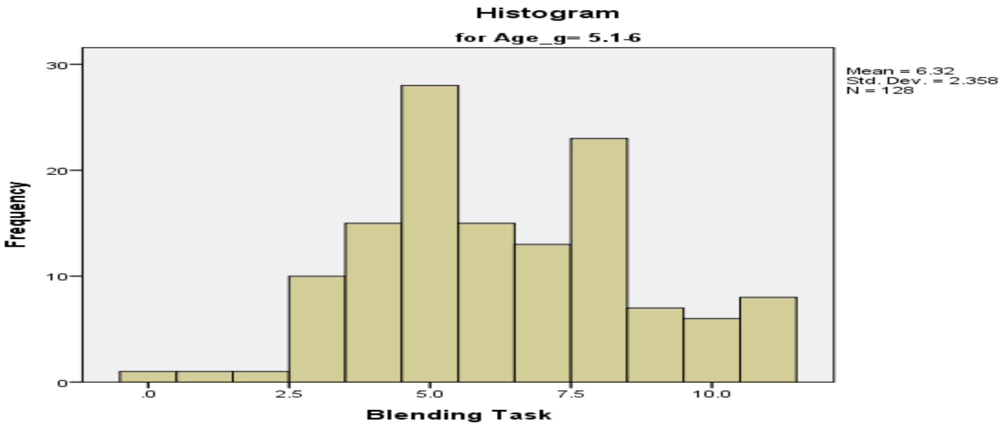


Figure 8.5: Graphical means for testing normality: Blending

Age Group 4;0-4;11



Age Group 5;0-5;11



Age Group 6;0-7;0

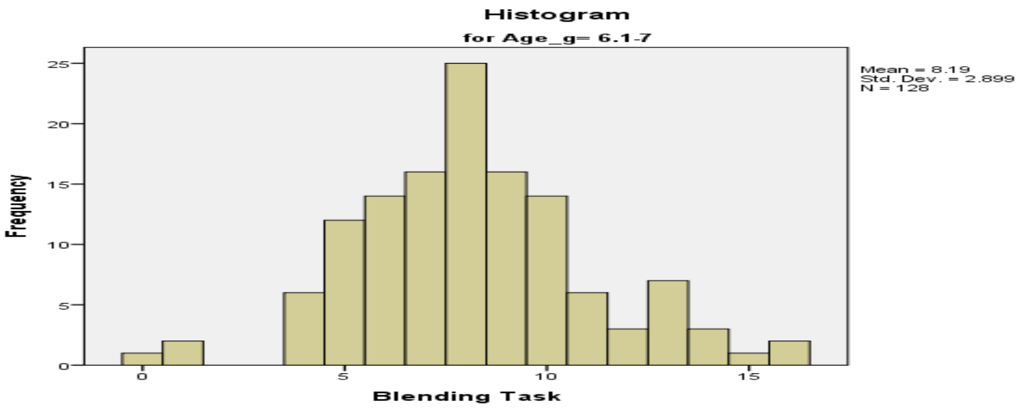
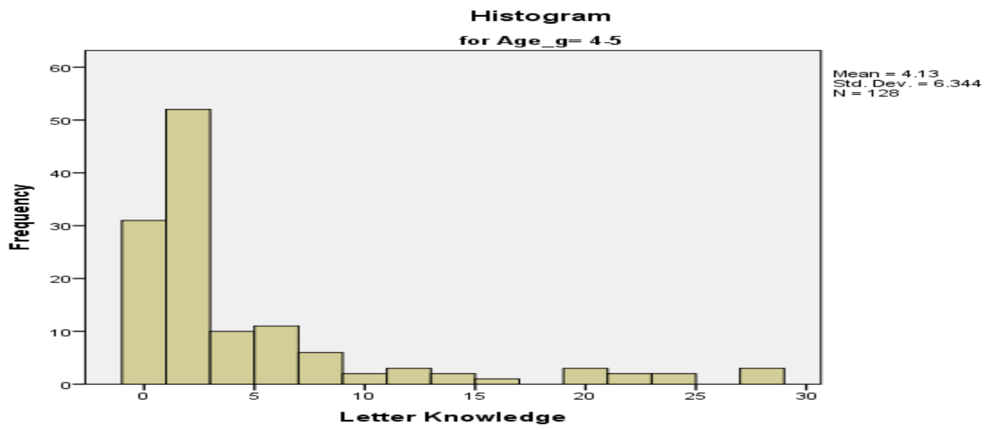
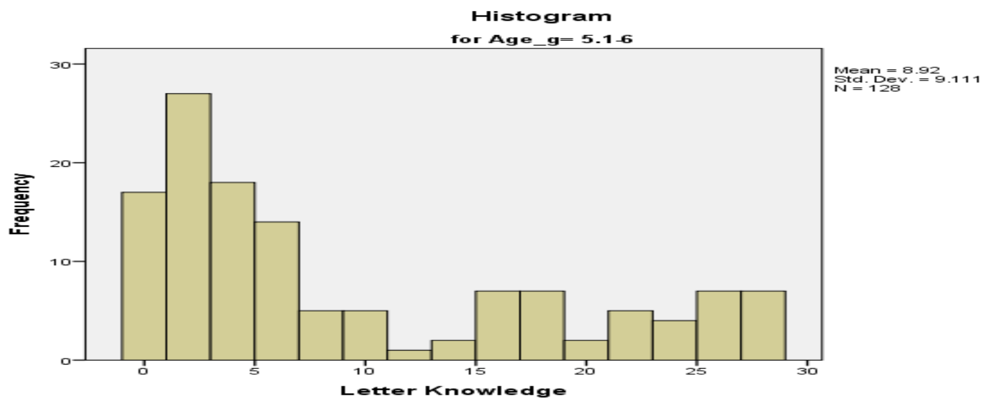


Figure 8.6: Graphical means for testing normality: LK

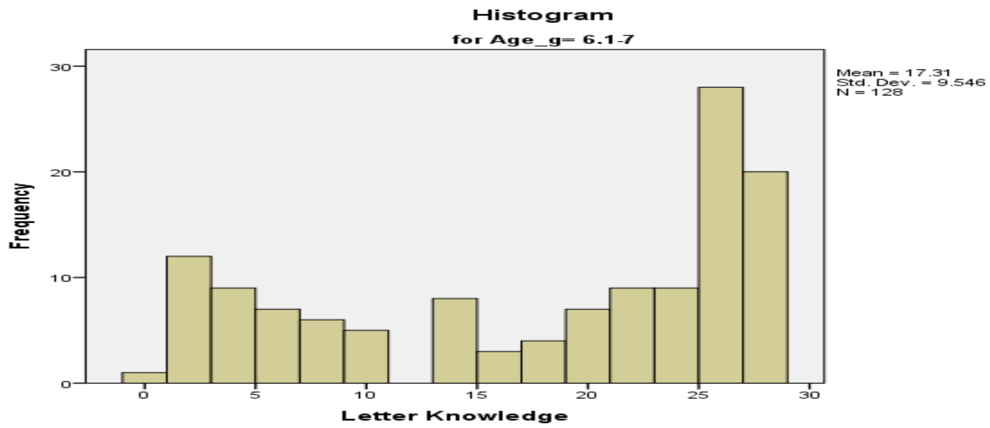
Age Group 4;0-4;11



Age Group 5;0-5;11



Age Group 6;0-7;0



Appendix J: Table 8.2: Skewness and Kurtosis values for PA Tasks and LK according to age group

	AGE GROUPS											
	4;0-4;11				5;0-5;11				6;0-7;0			
	Skewness		Kurtosis		Skewness		Kurtosis		Skewness		Kurtosis	
	Statistic	Std Error	Statistic	Std Error	Statistic	Std Error	Statistic	Std Error	Statistic	Std Error	Statistic	Std Error
Rhyme Awareness	-.088	.214	-.748	.425	-.791	.214	.129	.425	-.827	.214	1.236	.425
Syllable Segmentation	-.404	.214	.110	.425	-.298	.214	.247	.425	-.724	.214	.217	.425
Alliteration Awareness	-.126	.214	-.109	.425	.406	.214	.394	.425	.172	.214	-.852	.425
Phoneme Isolation	-1.125	.214	.319	.425	.220	.214	-1.496	.425	-.517	.214	-.930	.425
Blending	.897	.214	1.623	.425	.161	.214	-.410	.425	.242	.214	.575	.425
LK	2.234	.214	4.414	.425	.831	.214	-.797	.425	-.511	.214	-1.341	.425

*

Appendix K: Table 8.3: Kolmogorov-Smirnov and Shapiro-Wilk normality tests

TASK	Age (1 year age bands)	KOLMOGOROV-SMIRNOV[A]			SHAPIRO-WILK		
		Statistic	df	Sig.	Statistic	df	Sig
Rhyme Awareness	4;0-4;11	.135	123	.000	.951	123	.000
	5;0-5;11	.214	128	.000	.909	128	.000
	6;0-7;0	.205	133	.000	.904	133	.000
Syllable Segmentation	4;0-4;11	.150	123	.000	.950	123	.000
	5;0-5;11	.149	128	.000	.944	128	.000
	6;0-7;0	.179	133	.000	.910	133	.000
Alliteration Awareness	4;0-4;11	.181	123	.000	.934	123	.000
	5;0-5;11	.146	128	.000	.950	128	.000
	6;0-7;0	.124	133	.000	.956	133	.000
Phoneme Isolation	4;0-4;11	.237	123	.000	.831	123	.000
	5;0-5;11	.177	128	.000	.863	128	.000
	6;0-7;0	.196	133	.000	.909	133	.000
Blending Task	4;0-4;11	.158	123	.000	.940	123	.000
	5;0-5;11	.162	128	.000	.960	128	.001
	6;0-7;0	.116	133	.000	.972	133	.008
LK	4;0-4;11	.292	123	.000	.657	123	.000
	5;0-5;11	.233	128	.000	.822	128	.000
	6;0-7;0	.181	133	.000	.859	133	.000

Appendix L: Nonparametric Tests

Figure 8.7: Kruskal-Wallis Tests for PA tasks and LK (all age group)

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Rhyme Awareness is the same across categories of Age in 1 year age bands.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
2	The distribution of Syllbles segmentation is the same across categories of Age in 1 year age bands.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
3	The distribution of Alliteration Awareness is the same across categories of Age in 1 year age bands.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
4	The distribution of Phoneme Isolation is the same across categories of Age in 1 year age bands.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
5	The distribution of Blending Task is the same across categories of Age in 1 year age bands.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
6	The distribution of Letter Knowledge is the same across categories of Age in 1 year age bands.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Figure 8.8: Mann-Whitney U Test for PA tasks and LK

4;0-4;11 and 5;0-5;11 age groups

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Rhyme Awareness is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
2	The distribution of Syllbles segmentation is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.002	Reject the null hypothesis.
3	The distribution of Alliteration Awareness is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.002	Reject the null hypothesis.
4	The distribution of Phoneme Isolation is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
5	The distribution of Blending Task is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.046	Reject the null hypothesis.
6	The distribution of Letter Knowledge is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

5;0-5;11 and 6;0-7;0 age groups

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Rhyme Awareness is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.747	Retain the null hypothesis.
2	The distribution of Syllbles segmentation is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.004	Reject the null hypothesis.
3	The distribution of Alliteration Awareness is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.003	Reject the null hypothesis.
4	The distribution of Phoneme Isolation is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
5	The distribution of Blending Task is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
6	The distribution of Letter Knowledge is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Figure 8.9: Kruskal-Wallis Tests for RAN tasks (all age group)

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Colour naming time average valid is the same across categories of Age in 1 year age bands.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
2	The distribution of Object naming time average valid is the same across categories of Age in 1 year age bands.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
3	The distribution of Letter naming: time: average valid is the same across categories of Age in 1 year age bands.	Independent-Samples Kruskal-Wallis Test	.017	Reject the null hypothesis.
4	The distribution of Digit naming time: average valid is the same across categories of Age in 1 year age bands.	Independent-Samples Kruskal-Wallis Test	.002	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Figure 8.10: Mann-Whitney U Test for RAN tasks

4;0-4;11 and 5;0-5;11 age groups

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Colour naming time average valid is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.018	Reject the null hypothesis.
2	The distribution of Object naming time average valid is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
3	The distribution of Letter naming: time: average valid is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.017	Reject the null hypothesis.
4	The distribution of Digit naming time: average valid is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.026	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

5;0-5;11 and 6;0-7;0 age groups

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Colour naming time average valid is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
2	The distribution of Object naming time average valid is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
3	The distribution of Letter naming: time: average valid is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.	Unable to compute.
4	The distribution of Digit naming time: average valid is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.002	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

4;0-4;11 and 6;0-7;0 age group

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Colour naming time average valid is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.032	Reject the null hypothesis.
2	The distribution of Object naming time average valid is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.820	Retain the null hypothesis.
3	The distribution of Letter naming: time: average valid is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.	Unable to compute.
4	The distribution of Digit naming time: average valid is the same across categories of Age in 1 year age bands.	Independent-Samples Mann-Whitney U Test	.325	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Appendix M: Table 8.4:Confidence Intervals for PA and LK tasks (one-year age groups)

	Age Groups								
	4.00-4.11			5.00-5.11			6.00-7.00		
		95% Confidence Interval			95% Confidence Interval			95% Confidence Interval	
TASKS	Mean	Lower Bound	Upper bound	Mean	Lower Bound	Upper bound	Mean	Lower Bound	Upper bound
RA	5.17	4.91	5.43	5.93	5.67	6.19	6.05	5.80	6.30
SS	4.81	4.52	5.09	5.44	5.16	5.72	5.93	5.65	6.20
AA	2.70	2.42	2.98	3.31	3.03	3.59	4.02	3.74	4.29
PI	1.74	1.31	2.17	3.18	2.76	3.60	4.71	4.29	5.12
B	5.71	5.26	6.15	6.23	5.79	6.67	8.16	7.73	8.59
LK	3.96	2.45	5.47	8.91	7.43	10.38	16.99	15.54	18.43

RA= Rhyme Awareness SS = Syllable segmentation B = Blending

AA= Alliteration awareness PI =Phoneme Isolation

Table 8.5: Confidence Intervals for RAN tasks means by age group

	Age Groups								
	4.00-4.11			5.00-5.11			6.00-7.00		
		95% Confidence Interval			95% Confidence Interval			95% Confidence Interval	
RAN TASKS	Mean	Lower Bound	Upper bound	Mean	Lower Bound	Upper bound	Mean	Lower Bound	Upper bound
Colour Naming Time	46.68	43.47	49.89	41.60	39.33	43.86	38.22	36.75	39.70
Object Naming Time	48.96	46.03	51.88	48.60	46.59	50.62	42.75	41.12	44.38
Letter Naming Time	_	_	_	41.15	36.80	45.50	35.23	33.23	37.24
Digit Naming Time	47.45	41.87	52.97	43.33	39.40	47.25	36.82	34.99	38.65

Appendix N: Table 8.6: Means, standard deviations and confidence intervals for PA tasks and LK, by age group, with/without extra Qur'an sessions

		Age Groups								
		4;00-4;11			5;00-5;11			6;00-7;00		
			95% Confidence Interval			95% Confidence Interval			95% Confidence Interval	
Without Extra Tuition	PA tasks	M	LB	UB	M	LB	UB	M	LB	UB
	RA	5.09	4.81	5.36	5.86	5.58	6.14	5.98	5.71	6.25
	SS	4.72	4.41	5.02	5.39	5.09	5.69	6.00	5.71	6.29
	AA	2.69	2.38	2.99	3.27	2.96	3.57	4.00	3.71	4.30
	PI	1.53	1.07	1.99	3.02	2.56	3.48	4.73	4.29	5.17
	B	5.71	5.23	6.19	6.30	5.89	6.78	8.18	7.72	8.65
	LK	3.87	2.25	5.49	8.16	6.56	9.76	16.7	15.2	18.3
With Extra Tuition	PA tasks	M	LB	UB	M	LB	UB	M	LB	UB
	RA	5.71	5.01	6.40	6.30	5.66	6.94	6.50	5.82	7.18
	SS	5.35	4.59	6.11	5.70	5.00	6.40	5.44	4.71	6.18
	AA	2.77	1.10	3.53	3.55	2.84	4.26	4.11	3.37	4.86
	PI	3.06	.905	4.21	4.05	2.99	5.11	4.56	3.43	5.68
	B	5.71	4.50	6.91	5.85	4.73	6.96	8.00	6.83	9.17
	LK	4.53	.493	8.57	12.9	9.23	16.6	18.2	14.3	22.1

RA= Rhyme Awareness SS = Syllable segmentation

AA= Alliteration awareness PI =Phoneme Isolation

B= Blending LK = Letter Knowledge

M = Mean LB= Lower Bound UB =Upper bound

Appendix O: Examples of pictures used in testing battery