Human Judgements and Automated Measures of Text Readability and Comprehensibility: A Textual Analysis of Three English Qur'ān Retranslations

Abdulrahman Fahad Albalawi

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

The University of Leeds
School of Languages, Cultures and Societies

March 2022

The candidate confirms that the work submitted is his/her own and that appropriate credit has been given where reference has been made to the work of others.

Dedication

I dedicate this work to:

my dearest mother and father, my beloved wife, my wonderful children, and my kind sisters and brothers

Acknowledgements

First and foremost, all praises and thanks go to the Almighty God for giving me the patience, endurance, ability, strength, guidance, and opportunity to undertake and complete this research. Throughout the journey of this academic study, I have become deeply indebted to many people who helped me along the way.

First, I would like to express my deepest gratitude and appreciation to my main supervisor, Professor James Dickins, for his support, guidance, patience, understanding, reading, scrutiny, and revisions of all drafts of this study. Professor Dickins is not only an esteemed scholar of Arabic and translation studies, but he is also an expert in supervising interdisciplinary PhD topics. He is known as a friend to people and especially to students. He is also an engaging, supportive, and welcoming person to his students when it comes to their health, financial, and family matters. His comments and expert input into the development of every section of this thesis were invaluable and appreciated. My words do not do justice to him; may God bless him.

I would like to extend my sincere thanks and a deep sense of gratitude to my co-supervisor, Dr Mustapha Sheikh, for his sincere support, encouragement, and invaluable advice. I would like to further express my deepest thanks to Mrs Karen Priestley, Doctoral College Officer, for facilitating the administrative work during the PhD journey. I am also thankful to Dr Mustapha Sheikh, Dr Fozia Bora, and Mrs Helen Costelloe for their cooperation in distributing my online questionnaires to all AIMES undergraduate students.

I owe sincere gratitude to my parents, Fahad Albalawi and Helah Aljohani, for their unconditional love, daily prayers, and kind guidance. My special thanks go to my lovely wife, Alaa Aljohani, and to my three wonderful children: Joud, Omar, and Siba, I am particularly thankful to my wife for her loving and caring for our children and her efforts to ease our stress and anxiety whilst we endured the challenging times of the Covid-19 pandemic and the closures of UK schools, university facilities, shops, and parks. I would like to offer my gratitude to my siblings, half-brothers, and half-sisters for their prayers, unwavering support, affection, and sense of brotherhood. I am extremely thankful to my mother Helah, my brother Abdulaziz, my brother Abdullah, and my sister Dr Shuruq for their daily phone calls, encouragement, and financial support.

Finally, I owe a special sense of gratitude to the government of Saudi Arabia and Taibah University for their sponsorship. I am grateful to my fellow postgraduates Dr Abdulrahman Albeladi, Dr Majdi Alzahrani, Dr Abdallah Al-Turki, and Husain Alajmi for their friendship and helpful advice.

Abstract

This study investigates the text readability and comprehensibility of three prominent English retranslations of the Qur'ān: Abdel-Haleem (2004), Arberry (1955, revised edition 1983), and Yusuf Ali (1934, revised edition 1989). The study uses three quantitative approaches for the measurements of text readability and comprehensibility: the classic approach (i.e., classic readability formulae); the computational approach (i.e., automated evaluation of text and discourse with Coh-Metrix); and the human judgement approach (i.e., human ratings of text readability and ease of comprehension). This multimethod research employs multilevel theoretical frameworks of automated measures and human judgements to analyse and estimate the readability of the Qur'ān translations, based on the following primary text factors and reader characteristics: *style*, *cohesion*, *literalness*, *genre*, *register*, *page layout*, *verse comprehensibility*, *comprehension aids*, *reader's prior knowledge,* and *retranslation.*

This research provides evidence that the earlier translations of the 20th century, in comparison with Abdel-Haleem's later version, are rendered using a more complex style of words and syntax, with higher rates of lexical repetition, semantic coherence, connectives, and literalness; lower rates of narrativity; greater use of linguistic features of register variation associated with text difficulty and discourse complexity; and greater source language-orientation in terms of their adherence to source language syntax and cohesion. The translation with low text comprehensibility (i.e., Arberry's) lacked comprehension aids (e.g., introductory information and footnotes in the Qur'ānic chapters). The human judgements differed significantly only in three elements of page layout: *chapter title*, *chapter length*, and *verse number*. Retranslation has been shown to reveal changes in text readability and comprehensibility when the timespan between the translations is long. Readers with low knowledge of the Qur'ān have been shown to better understand more from the more readable Qur'ān version (i.e., Abdel-Haleem's), whilst high-knowledge readers have been shown to comprehend more from the less readable Qur'ān version (i.e., Ali's). We argue that readers' background knowledge and levels of understanding are influenced not only by text properties (i.e., those versions with high or low text readability), but also by non-textual factors related to retranslation such as the background characteristics of the translator and the role of powerful institutions and publishers. These factors could contribute to the comprehensibility of a Qur'ān translation.

The findings converge with and validate other approaches. However, the classic formulae give an improper estimation of the reading grade levels of the Qur'ān versions and are less

predictive of verse comprehensibility than the computational indices used in Coh-Metrix, including lexical, syntactic, semantic, and discourse features.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

SL	Source Language
ST	Source Text
TL	Target Language
TT	Target Text
RH	Retranslation Hypothesis
CRFs	Classic Readability Formulae

Arabic Transliteration System

The automated transcriber Intellibe was used in this research to transliterate the Arabic words into English, whenever needed. The following table[1] explains the transliteration system used for the Arabic consonants and vowels:

Arabic	Transliteration	Arabic	Transliteration
أ	ʾ/ā	ط	t
ب	b	ظ	ḍ
ت	t	ع	ʿ
ث	ṯ	غ	ġ
ج	ǧ	ف	f
ح	ḥ	ق	q
خ	ḫ	ك	k
د	d	ل	l
ذ	ḏ	م	m
ر	r	ن	n
ز	Z	ه	h
س	S	و	ū\w
ش	š	ي	ī\y
ص	ṣ	ء	ʾ
ض	ḍ		

Arabic short – long vowels and case endings:

Arabic	Transliteration	Arabic	Transliteration
ا	ā	ـُ	o
ـو	ū	ـِ	i
ـي	ī	ـَ	a

[1] http://www.intellaren.com/. This relies on the above table to produce automatic Arabic transliterations.

Chapter 1: Introduction

1.1 Background

English Qur'ān translation research has long focused primarily on comparisons between the source text (ST) and target text (TT), exploring notions of (in)equivalence or (un)translatability of Qur'ānic discourse into the TT language (e.g., Abdul-Raof, 2001; 2004; 2005; 2007; 2018; 2019; Al-Amri, 2008; Boullata, 1988; Poonawala, 1990; Aldahesh, 2014; Dastjerdi and Jamshidian, 2011; El-Zeiny, 2011). The Qur'ān, in Islamic belief, is the word of God. Qur'ānic discourse has its own linguistic, stylistic, and pragmatic idiosyncrasies (Abdul-Raof, 2001; 2018; 2019). The Qur'ān is undoubtedly miraculous in its form and content because no other being could produce a discourse similar to that of God. This can be seen in the fact that God challenges the Qur'ān sceptics and Arab polytheists to produce a discourse similar to the Qur'ān, either in Arabic or another language, as the following verse informs us:

(قُل لَّئِنِ اجتمعت الإنس والجن على أَن يَأْتُواْ بِمِثْلِ هذا القرآن لاَ يَأْتُونَ بِمِثْلِهِ وَلَوْ كَانَ بَعْضُهُمْ لِبَعْضٍ ظَهِيراً)
[الإسراء: 88]

> [Say, 'If mankind and the jinn gathered in order to produce the like of this Qur'an, they could not produce the like of it, even if they were to each other assistants'. Q17:88] (Saheeh International, 1997).

Some of these previous studies of English translations of the Qur'ān have given special attention to Qur'ān-specific linguistic and cultural features and to the inherent characteristics of Qur'ān discourse that are difficult to reproduce in English (e.g., Abdul-Raof, 2001) or entirely absent from English. As Abdul-Raof (2018) observes, there are 'limits of translatability' of Qur'ān-specific language, which lead to translation problems. This is because Arabic and English have linguistic and cultural incongruities (Abdul-Raof, 2018). When a translator attempts to preserve Qur'an-specific stylistic and linguistic features in a given target language (TL), the rendering might result in a text that is less readable and less comprehensible for the target audience.

These previous studies on Qur'ān translations, however, have not attempted to explore multiple TTs (only) or what the textual factors of multiple TTs have brought to the receiving language system and reader (see Susam-Sarajeva, 2003; Venuti, 2004; Desmidt, 2009). This phenomenon of studying multiple TTs of the same ST in the same language is known in translation studies as 'retranslation' (Chesterman, 2000).

The retranslation of a translated work is a universal act in translation studies and occurs in all types of translation (Susam-Sarajeva, 2003). It has primarily been discussed in the context

of canonical and literary texts (e.g., Bensimon, 1990; Berman, 1990; Susam-Sarajeva, 2003; Venuti, 2004; Paloposki and Koskinen, 2004, 2010; Gürçağlar, 2009, Hanna, 2016). This is despite religious scriptures such as the Qur'ān and the Bible having been more frequently retranslated than literary works. Translations of the Qur'ān, the Bible (Susam-Sarajeva, 2003), and other sacred writings are rarely explored from the viewpoint of retranslation (including the practice's influences on text readability and comprehensibility).

Retranslation can be viewed as a theory that is not yet fully mature. It dates back to the latter end of the 20th century (Bensimon, 1990; Berman, 1990), when two French scholars generated several hypotheses that are now referred to as the 'retranslation hypothesis' (RH). A common sentiment in the RH is that the first translations tend to be more target culture-oriented or (TT-oriented) than later retranslations, which tend to be more source culture-oriented (Bensimon, 1990). This hypothesis seems to reflect the general strategies of translation and the context of literary translation. The validity of this claim has been tested and rejected by several researchers (see Chapter 2, section 2.5).

Retranslation offers an alternative understanding of the translation activity by describing its motives and reasons, rather than being restricted to the initial claim of the RH regarding translation strategies. This is because the motives of retranslation can be influenced by a complex mix of linguistic, cultural, social, political, ideological, educational, and commercial factors (e.g., Susam-Sarajeva, 2003; Venuti, 2004; Paloposki and Koskinen, 2004, 2010; Gürçağlar, 2009; Deane-Cox 2014; Taivalkoski-Shilov; 2015; for more details of these factors, see section 2.6 in Chapter 2). For example, retranslations of the Qur'ān into English produced by some Muslim or non-Muslim translators are heavily affected by ideological and dogmatic matters (for more discussion of this, see Robinson [1997] and Chapter 2), reflecting the needs, interpretations, and power of certain social institutions (Venuti, 2004).

Retranslation, as Susam-Sarajeva rightly indicates (2003, p.5), might have 'more to do with the needs and attitudes within the receiving system than any inherent characteristics of the ST which make it "prone to" retranslations'. There have been retranslations of the Qur'ān into English since the middle of the 17th century (see Chapter 2, section 2.3.4), but the reasons why these works have been reproduced hundreds of times have not yet been fully considered from the perspective of retranslation studies. This investigation demonstrates the role of retranslation in English Qur'ān translations from the viewpoint of the linguistic factor, which fits well with the main theoretical framework of this research used to assess text readability and comprehensibility.

The study of text readability – in the sense of language comprehensibility – began in education and has been expanding since the 1920s (Klare, 1963; Chall, 1988; for more discussion of the development of readability measurements, see Chapter 3). This field of readability research is primarily concerned with the text factors and reader characteristics that interact with each other and could hence influence text readability and comprehensibility (e.g., Gray and Leary, 1935; Klare, 1963, 1974, 1988; Gilliland, 1972; Harrison, 1980; Chall, 1988; 1996; Zakaluk and Samuels, 1988; DuBay, 2004, 2007; Bailin and Grafstein, 2016). Research in translation studies has recently begun to examine readability in terms of the comprehensibility of language in medical translations (Jensen, 2015; Hill-Madsen, 2015); legal translations (Wolfer et al., 2015); and scientific and technical translations (Iljinska et al., 2015).

Previous translation studies on readability are limited and have not provided insights into the text factors and reader characteristics that might contribute to the readability and comprehensibility of translated texts in general and Qur'ān translations in particular. Hence, this interdisciplinary thesis borrows measures and approaches from the fields of education, psychology, computational linguistics, and translation to present a new approach to the analysis of translations, the phenomenon of retranslation, and the readability and comprehensibility of Qur'ān translations. Using this approach, this research is able to identify the primary text factors and reader characteristics (i.e., readers' prior knowledge) that contribute to making one translation of the Qur'ān more readable and comprehensible than another. In addition, it sheds light on the ways in which retranslation may increase or reduce the readability and comprehensibility of a Qur'ān translation.

As this research project explores text readability and comprehensibility, these two concepts should be given concrete, objective, and measurable definitions. Readability and comprehensibility are used interchangeably by some researchers (see Chapter 3, section 3.3), but others differentiate between the concepts (e.g., Harrison, 1980). This research employs Harrison's general definitions of the two concepts. First, 'Readability is an attribute to text' (Harrison, 1980, p.30). Readability is viewed primarily from the perspective of the text itself. It is defined as the extent to which a translation can be easily read – regarding text properties, including *style, cohesion*, *literalness*, *genre*, *register*, *page layout*, *verse comprehensibility*, *title comprehensibility,* and *comprehension aids.* These primary text factors – in conjunction with the reader's prior knowledge, drawn from relevant literature – form the theoretical framework used in this research to analyse the data.

On the other hand, 'comprehension is an attribute of readers' (Harrison, 1980, p.30). Text comprehensibility is never static and can change. Reader characteristics – such as their prior knowledge of content – that they bring to the text will affect their understanding of it (e.g., Klare, 1976; 1988; Gilliland, 1972; Zakaluk and Samuels, 1988; McNamara and Kintsch, 1996; Pikulski, 2002; Ozuru et al., 2009; O'Reilly and McNamara, 2007). For example, English Qur'ān versions might be incomprehensible to readers without knowledge of the Qur'ān, whilst being entirely comprehensible to other readers with prior knowledge. Thus, this research considers whether readers with low or high levels of prior knowledge of the Qur'ān can more easily understand versions with lower or higher levels of readability.

As this study attempts to translate the aforementioned concepts into measurable and numerical entities, it exclusively adopts a quantitative approach, which fits well with the philosophical stance of testing and understanding hypotheses and theories related to the retranslation, readability, and comprehensibility of a text. This research strategy is closely associated with positivist and empiricist paradigms[2] (Bryman, 2012; Saldanha and O'Brien, 2014) and driven by a 'theory-testing approach' (Bryman, 2012, p.621). The reliability and validity of measures are crucial factors in quantitative research (Bryman, 2012, p.163, see Chapter 4, section 4.2.2). This study employs a quantitative approach, with the help of statistical techniques (see Chapter 4, section 4.7). Quantitative investigations are not only useful for testing hypotheses in theory, but also for uncovering abnormalities in data and theory when 'look[ing] for a new phenomenon' (Kuhn, 1961, p. 181).

Finally, the study uses a multimethod research design (see Chapter 4, section 4.2) to assess and explore the readability and comprehensibility of three prominent English Qur'ān versions: Abdel-Haleem (2004), Arberry (1955, revised edition 1983), and Yusuf Ali (1934, revised edition 1989). It employs three quantitative approaches to the measurement of text readability and comprehension: the classic approach (i.e., classic readability formulae [CRFs]); the computational approach (i.e., automated evaluation of text and discourse with Coh-Metrix); and the human judgement approach (i.e., human ratings of text readability and comprehensibility).

[2] Positivism is typically associated with a quantitative research approach and with empiricism, which suggests that 'ideas must be subjected to the rigours of testing before they can be considered knowledge' (Bryman, 2012, p. 23).

Each of the three approaches has multiple-indicator measures of text cohesion, readability, and comprehensibility, which estimate comprehension difficulty for different levels of text characteristics (see Chapter 4). The primary rationale for using multiple quantitative approaches and measures was to enable the measurement of multilevel theoretical frameworks of text readability and comprehensibility, as well as the converging and verifying of the findings of the automated measures (see Chapter 5) and human judgements (see Chapter 6). Thus, the final aim of this thesis is to understand the differences between objective measures and subjective judgements of text readability and comprehensibility in Qur'ān translations. This thesis uses alternative, rapid, and modern approaches to textual analysis and translation quality – unlike the traditional linguistic-textual approaches of translation quality assessments of equivalence (House, 1997) – to produce more tangible, objective, unbiased, and systematic results.

This introductory section has given a panoramic view of this study. It covered the three key concepts applied in this thesis: retranslation, readability, and comprehensibility. The following sections of this chapter present the thesis's motivations, aims, and questions, followed by a detailed outline of the remaining chapters.

1.2 Research Motivations

The original motivation for this study was the observation that no academic publication has provided insights into the contributors to text readability and comprehensibility in English Qur'ān translations. The Qur'ān has been retranslated into English many times, but the academic research into Qur'ān translations has tended to focus on direct comparisons between the original and the translations. Thus, this study explores an uncharted area of English Qur'ān translation, with its focus on the language of the receiving text and its accommodation of Qur'ānic discourse in more readable and understandable language.

Another motivation for this study was the consideration that a translation of the Qur'ān is intended to be 'readable' for a large proportion of the population, including adult readers who are less educated and who have little background knowledge of the text. A highly important text intended to be read by lay people should not contain textual features associated with text difficulty or discourse complexity. Texts comprised of more complex language will not facilitate text comprehensibility for many readers, especially those with little prior knowledge of the content (e.g., Klare, 1976; 1988, Gilliland, 1972; Zakaluk and Samuels, 1988; Chall, 1996; Kintsch and Miller, 1981; McNamara et al., 2011, 2014). Legal, technical, and medical translations contain language and terms that are not easy for the general public to comprehend

(e.g., Jensen, 2015; Hill-Madsen, 2015; Wolfer et al., 2015 Iljinska et al., 2015). However, it is extremely important that a version of the Qur'ān is comprehensible to most adult readers, especially low-knowledge readers. Therefore, this academic exploration of comprehensibility and readability of the selected versions is vital.

Another important and intriguing aspect of this study is its use of computational linguistic indices and statistics to explore textual changes and linguistic variation in English Qur'ān translations. Computational tools, such as Coh-Metrix, and statistical techniques are common approaches in non-translation studies to analyse and differentiate between linguistic characteristics and register variations of texts at many levels of language and discourse (e.g., Louwerse et al., 2004; Graesser and McNamara, 2011; McNamara et al., 2014; Crossley et al., 2014). This approach, with the help of advanced statistical techniques and the human judgement approach, opens a new horizon for the quantitative, objective, and subjective measurement of words, sentences, discourse, literalness, page layout, verse comprehensibility, and comprehension aids in Qur'ān translations.

1.3 Research Aims and Questions

❖ Research Aims

This study has three aims, which are summarised as follows:

- To identify the text factors and reader characteristics (i.e., readers' prior knowledge of the text) that contribute to making one translation of the Qur'ān more readable and comprehensible than another
- To explore the impact of retranslation on text readability and comprehensibility
- To explore the differences between automated measures and human judgements of readability and comprehensibility in Qur'ān translations

❖ Research Questions

This thesis poses the following research questions:

1. What are the primary textual factors that could affect the readability and comprehensibility of Qur'ān translations, as indicated by automated measures?
2. Does retranslation result in higher levels of readability and comprehensibility?

3. Which of the three translations under study is the most readable and comprehensible, and which is the least readable and comprehensible, as indicated by human judgement?
4. Do human judgements of readability and comprehensibility correlate with the automated measures and yield consistent findings?
5. Do readers with low (or high) levels of prior knowledge of the Qur'ān gain more understanding from lower (or higher) levels of readability?

1.4 Thesis Structure

This thesis is organised into seven chapters.

Chapter 1: Introduction provides a panoramic view of the study and presents the three key concepts used in the data analysis: retranslation, readability, and comprehensibility. This chapter also presents the thesis's motivations, aims, and questions.

Chapter 2: **Qur'ān Translation and Retranslation** begins with the nature of the Qur'ān and its importance to Muslims as their first source of reference. It presents a brief historical review of Qur'ān translation into Asian, African, and European languages, with a special focus on English and on the three versions used in this research – namely, Abdel-Haleem (2004), Arberry (1955, revised edition 1983), and Yusuf Ali (1934, revised edition 1989). The last part of this chapter is devoted to the concept of retranslation, its origin, and the motives for retranslation in general and for Qur'ān translations in particular. This chapter highlights the key reasons for retranslation and approaches Qur'ān retranslations in terms of their linguistic aspects, which are relevant to an assessment of readability and comprehensibility.

Chapter 3: Readability and Comprehensibility presents a review of the relevant literature in relation to the theoretical framework adopted in this research. This core chapter begins by defining readability and its relationship with comprehensibility, showing how these concepts are operationalised in the study. The chapter includes a brief overview of the development of readability formulae – primarily in English and then briefly in other languages (such as Arabic) – tracing the features used to probe content difficulty and comprehensibility in the early studies of readability and classic readability formulae. This is followed by criticisms of readability formulae and an overview of the computational software (Coh-Metrix) used in this study to scale content difficulty for multiple levels of text characteristics. The chapter concludes with a preliminary sketch of a text–reader model of text readability and comprehensibility, focusing on the textual factors (e.g., style, organisation, cohesion and coherence, and genre) and reader characteristics (prior knowledge, interest, and motivation) that can affect text readability and reader comprehension.

Chapter 4: Research Methodology describes the difference between 'multimethod' and 'mixed-method' research, the reasons for choosing a multimethod approach in this study, and the three quantitative approaches used here to measure comprehensibility and text difficulties (i.e., the classic approach, the computational approach, and the human judgement approach). These approaches allow for an analysis of multiple perspectives on the text and reader characteristics. This chapter includes a description of the five classic readability formulae used in the study for preliminary investigations of content difficulty. The chapter describes multiple measures of readability, cohesion, and language taken from Coh-Metrix and how the three corpora were collected and cleaned. The chapter provides a detailed explanation of how human judgements are implemented in the study, as well as the survey design, rating scale, pilot study, sampling, survey administration, survey type, and ethical considerations. Finally, the chapter presents a set of simple and advanced statistical techniques used for data analysis.

Chapter 5: Data Analysis of Automated Measures reports the findings of automated measures, derived from two approaches to readability and textual analysis: the classic formulae and the computational Coh-Metrix. Forty Qur'ānic chapters taken from each of the selected three corpora of Qur'ān translations are analysed, using the five classic readability formulae of reading difficulty, six variables of lexical difficulty, six variables of syntactic complexity, six variables of referential cohesion, four variables of semantic coherence, seven variables of connectives, and narrativity (a measure of the genre of a text). This multilevel theoretical framework for analysing readability, cohesion, and language used in this research describes and differentiates between the readability of the three Qur'ān versions according to the major textual factors: *style*, *cohesion*, *literalness*, *genre*, *register*, and *retranslation.*

Chapter 6: Data Analysis of Human Judgements presents the subjective judgements of the readability and comprehensibility of the two complete Qur'ānic chapters taken from the three selected Qur'ān versions. In this chapter, the readability of each translation is judged on a 5-point semantic differential scale, based on four textual factors: page layout, word usage, sentence usage, and cohesion usage. The two Qur'ān chapters are evaluated in relation to four aspects of translation comprehensibility: the comprehensibility of verses, the comprehensibility of titles, the influence of the introduction on text comprehensibility, and the influence of footnotes on text comprehensibility. This chapter is intended to examine readability and comprehensibility from differing perspectives and reader viewpoints, thus revealing the differences between automated measures and subjective judgements.

Chapter 7: Conclusion summarises the findings of the study. The implications and limitations of the study's findings and suggestions for future research are presented.

Chapter 2: Qur’ān Translation and Retranslation

2.1 Chapter Overview

In this chapter, section 2.2 discusses what the Qur'ān is; its importance to Muslims as the word of God and the first source of reference for Muslims; how it is understood and approached by other Islamic sciences of the Qur'ān, such as hadith and Qur'ānic exegesis; and finally, the arguments around the legitimacy (or otherwise) of translating the Qur'ān into other languages. Section 2.3 is divided into four sub-sections on the history of Qur'ān translations into Asian languages, African languages, and European languages, and finally, with a major focus on English. We review the best-known English Qur'ān translations and focus on the three that are used in this research: Abdel-Haleem (2004), Arberry (1955, revised edition 1983), and Ali (1934, revised edition 1989). Literalness of English Qur'ān translations is discussed after this section. Section 2.4 begins with a discussion of the concept of retranslation, exploring the various types and focusing on how English Qur'ān retranslations can be described or distinguished since the first publication of a Qur'ān translation. Section 2.5 reviews the literature relevant to the origins of retranslation in the context of canonical literary texts and its development under the so-called 'retranslation hypothesis' (RH). This is followed by section 2.6 on the motivations for retranslation in general and for Qur'ān retranslations in particular. Finally, there is an overall chapter summary, before the thesis moves onto a discussion of text readability and comprehensibility.

2.2 The Qur'ān

In Muslim belief, the words in the Qur'ān are not manmade, as claimed by the Qur'ān sceptics in the era of the Prophet Mohammed and even by some orientalists in the early Qur'ān translations (see section 2.3.4). The Qur'ān is the word of God, as the following verse informs us:

(وَإِنْ أَحَدٌ مِّنَ الْمُشْرِكِينَ اسْتَجَارَكَ فَأَجِرْهُ حَتَّىٰ يَسْمَعَ كَلَامَ اللَّهِ...) [التوبة: 6]

> [And if any one of the polytheists seeks your protection, then grant him protection so that he may hear the words of Allah. Q9:6] (Saheeh International, 1997).

The Qur'ān, revealed in Arabic to the Prophet Muhammad via the angel Gabriel, is considered the highest-authority legislation in Islam and the first authority for Muslims (Abdel-Haleem, 2004). The entire religious life of Muslims worldwide relies on the Qur'ān, whose prescriptions Muslims implement in their daily lives by believing in God, praying, almsgiving, fasting, and pilgrimage. The Qur'ān provides guidance to Muslims and covers numerous those related to monotheism (i.e., unity of God), eschatology (i.e., the Day of Judgement), reward and punishment, and prophethood and stories of previous Prophets. The Qur'ān covers other topics of Islamic principles, rules, ethics, teachings, and theological matters (see, Al-Attar, 2010; Abdul-Raof, 2010).

For Muslims and Arabic literary scholars, the Qur'ān is the primary text in terms of the beauty of its language and its stylistic elegance. According to Abdul-Raof (2018, p.78), 'the linguistic inimitability of Qur'anic discourse is its stylistic elegance, and the difficulty of translating Qur'anic discourse is its stylistic elegance'. The Qur'ān is neither poetry nor prose, but has its own unique discourse that is completely different to that of other books (Abdul-Raof, 2018), as is described in this verse:

(أَفَلَا يَتَدَبَّرُونَ الْقُرْآنَ وَلَوْ كَانَ مِنْ عِندِ غَيْرِ اللَّهِ لَوَجَدُوا فِيهِ اخْتِلَافًا كَثِيرًا) [النساء: 82]

> [do they not reflect upon the Qur'an? If it had been from [any] other than Allah, they would have found within it much contradiction. Q4:82] (Saheeh International, 1997).

The word 'Qur'ān' is an Arabic word, generally denoting 'reading' (Abdul-Raof, 2001), although some Muslim scholars and exegetes believe that there are several other meanings. Al-Zarkashi (1957, 1, pp. 276-7) writes:

أما القرآن فقد اختلفوا فيه؛ فقيل: هو اسم غير مشتق من شيء، بل هو اسم خاص بكلام الله... سمي القرآن قرآنا لأنه جمع السور بعضها إلى بعض... سمي قرآنا لكونه جمع ثمرات الكتب المنزلة السابقة.

As for the word 'Qur'ān,' they differed over its meanings. It was said that the Qur'ān is a name that is not a derived from anything, but it is a name specific to the word of God… It is called a 'Qur'ān' because it has collected the chapters together… It is named such because it has collected the fruits (or benefits) of all previously revealed scriptures [my translation].

The Qur'ān consists of 114 suras (سُوَر) of different size, typically referred to in English as 'chapters', which are traditionally arranged according to their length rather than their themes or the circumstances of the revelations (*ʾasbāb al-nuzūl*). Each Qur'ānic chapter has a set of *ʾāyāt* (آيات), typically referred to in English as 'verses'. Each Qur'ānic chapter has different themes and stories that are textually related to other parts of Qur'ānic chapters (see Abdul-Raof, 2018; 2019).

Qur'ānic chapters are organised from the longest chapters to the shortest. The longest is Chapter 2, and the shortest is Chapter 108. The first to be revealed was Chapter 96, and the last was Chapter 103, but some exegetes and the Prophet's companions do agree on the first and last chapters because they only narrated what they had heard of the last narration from the Prophet Muhammad (Al-Zarkashi, 1957). 'The first revelation of the Qur'ān came to him [the Prophet Muhammad] in 610 CE, when he was 40 years old' (Abdel-Haleem, 2004, xi). The Qur'ān consists of 77,439 words and 6,000 verses (Al-Zarkashi, 1957).

During the gradual revelation of the verses to the Prophet Muhammad, his companions began to learn and memorise the verses (see, Abdel-Haleem, 2004; Abdul-Raof, 2010). The Qur'ānic verses and the Prophet's comments and on the circumstances of the revelations were kept in the minds of his companions. The words were collected, written down by his companions, and then transmitted to subsequent successors. Through the companions of the Prophet and the successors of the Prophet's companions, the science of hadith was developed – 'hadith' being the Prophet's narrations or deeds relating to events and the meanings of verses. The hadith are normally considered the second source of the Islamic religion, acting as a bridge to understanding and elucidating the meanings of the Qur'ān (Abdul-Raof, 2010).

After the development of the hadith, exegesis (*tafsīr*) appeared as an independent science. Exegesis is integrally related to the hadith and acts as a link between the Qur'ān and hadith in explaining and facilitating the meanings of the Qur'ānic verses (Abdul-Raof, 2010). The schools of Qur'ānic exegesis differ from one another in the ways in which they explicate Qur'ānic verses. Abdul-Raof (2001) lists six key types of exegesis: linguistic, philosophical

and rationalistic, historical, intertextual, jurisprudence, and independent judgement-based. All these developments made the meanings of the Qur'ān more accessible and comprehensible to those who want to understand, study, and learn the Qur'ān: 'the Qur'ān was the starting point for all the Islamic sciences' (Abdel-Haleem, 2004, i).

Non-Arabic-speaking Muslims who live in Asia, Europe, Africa, and the Americas have no direct access to understanding of the Arabic Qur'ān, but the question of sharing the text through translation has been controversial among Muslim scholars, who question the legitimacy of translating the Qur'ān into other languages. Traditionally, some Islamic schools of thought hold that it is illegitimate to translate the Qur'ān, whilst others suggest that the Qur'ān may be translated if done by a Muslim translator (Mustapha, 2009). Abu Hanifa, a well-known Islamic scholar (c.700–67), believed that it was permissible to translate the Qur'ān into a foreign language if the Arabic texts were included in the translation (Mustapha, 2009). In addition, Mustapha (2009, p.228) argues that no translation of the Qur'ān should be termed a 'translation of the Qur'ān', but rather an 'interpretation of the Qur'ān in language X'. Mustapha (2009) affirms that no translation is a substitute for the Qur'ān and can only ever be a commentary or interpretation.

In conclusion, opinions on the illegitimacy of translating the Qur'ān have changed in the last two decades. The generally accepted view of many Muslim scholars is that (1) translation of the Qur'ān is essential if non-Arab communities are to understand the Qur'ān and Islamic teachings, (2) any translation of the Qur'ān is a form of exegesis or explanation and is more communicative than a literal translation, and (3) no translation of the Qur'ān can be recited in prayer in a foreign language (see Mustapha, 2009). The Qur'ān has today been rendered into most of the major languages of Asia, Europe, and Africa.

2.3 A Brief Historical Review of Qur'ān Translations

The Qur'ān has been re/translated into most of the major languages of the world. The following sub-sections will not cover the entire history of Qur'ān translations, but they will briefly cover the most important of the Asian, African, and European translations, with a focus on English.

2.3.1 Asian Translations

The first translation of the Qur'ān into an Asian language was done in Persian during the reign of the Abbasids (c.750–1258). This work was primarily considered a commentary using an interlinear translation, and it did not consider the Persian word order, according to Abdul-Raof

(2001). This Persian version included a short summary of the exegesis of Al-Ṭabari and was largely completed during the period of the Samanid King Abu Salih Mansur (c. 819-999; Al-Amri, 2005). The earliest Turkish translation of the Qur'ān appeared in the 11th century and was based on this medieval Persian version. Some authors claimed that Turkish scholars were among those who translated the Persian version, allowing them to make another version in Turkish at the same time (Al-Amri, 2005). Numerous interlinear translations of the Qur'ān then appeared in Turkish between the 15th and 19th centuries (for more on Turkish Qur'ān translations, see Wilson, 2009).

In South Asia, the first completed Urdu translation of the Qur'ān was published in 1826 by Shah Abdul Qadir Dehlvi. However, there had been earlier partial translations into Urdu. Other Urdu translations have since followed, including Bayan-ul-Quran by Israr Ahmed in 1998. The first Bengali translation of the Qur'ān was published in 1882 by Bhai Girish Chandra Sen. Bhai Girish Chandra Sen was the first person in South Asia to directly translate from the original Arabic (Dey, 2012).

Islam has had a long history in China, going back to the Tang dynasty (c. 618–907; Yang Huaizhong, 1996). The first translation of the Qur'ān into Chinese was done in 1927. This was a complete version by a non-Muslim. Other Chinese translations completed in the 20th century by non-Muslims were neglected by Muslims, though the reasons for this are unknown (Spira, 2005). Another modern Chinese translation of the Qur'ān was published in 1932 by a Muslim Chinese translator, Wang Jinzhai. Numerous Chinese translations have since been produced, based on well-known Qur'anic exegeses, with some based on English translations of the Qur'ān by Abdullah Ali (1934) and Muhammad Pickthall (1930; see Spira, 2005; Yang Huaizhong, 1996).

Other translations of the Qur'ān were produced from the 20th century onwards in other Asian languages, such as Japanese and Korean. The first contact of the Korean people with Arab Muslims was during the Unified Shilla dynasty (c. 668–935), and 'cultural dialogue has continued ever since, albeit sporadically, but cultural exchange and its influence became more salient in the twentieth century' (Choi and Kim, 2021, p.4). Two compete translations of the Qur'ān appeared in Korean in the late 20th century, published by two Muslim professors in Arabic studies. The first was completed by Kim Yong in 1970; and the second, produced by Choi Young in 1997, was commissioned by the King Fahd Complex for the Printing of the Holy Qur'ān. Other partial and complete translations of the Qur'ān have appeared in Korean in the 21st century (see Choi and Kim, 2021).

The history of Qur'ān translations into Japanese is not long. The first Japanese contact with Islam is more recent than that of other Asian languages (see Mateen and Ali, 2015; Obuse, 2019; Rahayu and Fauzan, 2020). Four complete Japanese translations of the Qur'ān appeared in the 20th century. The first was published in 1920 by a non-Muslim, Keuiche Sakamoto. This was based on English translations (Mateen and Ali, 2015). The second was produced in 1950 by a non-Muslim, Shumei Okawa. Another was published in 1957 by Toshihiko Izutsu, a non-Muslim professor and author of several books related to Islam. Izutsu's translation was the first to be done directly from the Arabic Qur'ān (Mateen and Ali, 2015). Another Japanese translation of the Qur'ān was published in 1972 by a Japanese Muslim, Haji Umar Mita.

Translations of the Qur'ān into other Asian languages, such as Indonesian and the closely related Malay (spoken in present-day Indonesia and Malaysia), appeared in the late 15th century. In the production of translations into Malay and Indonesian, there have been three waves: the first was from 1500 to 1920, the second from the 1920s to the 1960s, and the third from the 1960s to the 21st century (Riddell, 2014). These periods included partial and complete translations of the Qur'ān and commentaries into Malay and Indonesian.

To conclude, the translation of the Qur'ān into Asian languages such as Persian, Turkish, Urdu, and Indonesian has a long history. New translations and editions appear in these Asian languages almost every decade. Translation into other Asian languages such as Japanese, Korean, and Chinese has a shorter history, but numbers have been rapidly expanding since their appearance at the beginning of the 20th century. Asia was the starting point for Qur'ānic translation, appearing before any such African or European Qur'ān translations.

2.3.2 African Translations

The Qur'ān has been translated into most major African languages, including Yoruba, Hausa, Swahili, Wolof, and Kiswahili. The translations into Swahili and Yoruba were first done by Christians and then by African Muslims (Naudé and Miller-Naudé, 2011; Lacunza-Balda, 1997). Three complete translations of the Qur'ān in Swahili (spoken in Kenya, Tanzania, and Uganda) appeared in the 20th century. According to Lacunza-Balda (1997), Muslims in East Africa claimed that some of these works had been produced with the aim of refuting the Qur'ān and Islam. The first Swahili translation of the Qur'ān was published in 1923 by Godfrey Dale, who served as a bishop and worked for the Universities' Mission to Central Africa (Naudé and Miller-Naudé, 2011; Lacunza-Balda, 1997). Lacunza-Balda (1997, p.101) comments laconically that 'there is a widespread feeling among Muslims in East Africa that Dale's translation is polluted by the bias of the Christian approach to the Qur'ān'. Since Dale's

translation, two other complete translations have been published in Swahili by Muslim translators. The first was produced with commentaries in 1953 by Mubarek Ahmad Ahmadi, 'the head of the mission (Raiis-ut Tabliigh) [known in Arabic as *Jamāʿa Al-tablīġ* [3] جماعة التبليغ] and leader of the Ahmadiya community in East Africa' (Lacunza-Balda, 1997, p.102). Another complete translation into Swahili was also produced in 1969, with commentaries by Abdullah Saleh Al-Farsi. Al-Farsi's translation, largely taking the Sunni point of view, was intended to correct some of the errors and mistranslations in some verses in the previous Swahili translations, such as the Ahmadi translation (see Lacunza-Balda, 1997; Loimeier, 2005).

Three complete translations of the Qur'ān into Yoruba (spoken in Nigeria, Benin, and Togo) appeared in the 20th century. The first Yoruba translation was produced in 1906 by the Anglican Bishop Michael Samuel Cole, and it was republished in 1924 in Nigeria (Naudé and Miller-Naudé, 2011; Solihu, 2015). The intention of this Yoruba version was not far from that of Dale's translation into Swahili. Cole's view, stated in his text, was that 'it will help the cause of Christianity, and dispel the darkness of the ignorance that … prevails among Mohammedans in Yorubaland and they will be in a position to compare the Bible with the Koran and see which satisfies best the needs of humanity' (Solihu, 2015, p.16). Two other complete translations of the Qur'ān into Yoruba were produced by Muslim translators. The first was done in 1973 by a committee of Nigerian Muslims and the second in 1976 by Ahmadiya missionaries (see Solihu, 2015; Naudé and Miller-Naudé, 2011). Since then, 10 other Yoruba translations of the Qur'an have been published (see Solihu, 2015).

Translations of the Qur'ān into other African languages – such as Hausa Wolof, Kiswahili, and Luganda – have appeared since the beginning of the 20th century. Translations of the Qur'ān into these African languages can be found in Loimeier (2005), Naudé and Miller-Naudé (2011), Solihu (2015) and Lacunza-Balda (1997). Most African Qur'ān translations appeared in the 20th century. As a final point, in the initial days, Qur'ān translations into some African languages were either published by non-Muslims aiming to refute Islam or by Muslim translators hoping to introduce non-standard theological interpretations, mirroring the behaviour and attitude of some non-Muslim and Muslim translators of European translations.

[3]Jamāʿa Al-tablīġ was established in 1926 in India by Muhammad Ilyas Kandhlawi, an Indian Islamic scholar. Jamaat Tablighi has been criticised for their misinterpretations of some verses and for their practices of preaching, which do not reflect either the true approach of Islam or the Prophet's practices (see Al-Hilali, 1979).

2.3.3 European Translations

The first translation of the Qur'ān into a European language was into Latin in the 11th century. This Latin version was completed in 1143 by Robertus Retenensis and remained unpublished for four centuries, for unknown reasons, finally being published in 1543. Arberry (1996) considered this medieval version to be the foundation of many European Qur'ān translations. However, he criticised this Latin version for two reasons. First, it contains many mistakes and misinterpretations. Second, the attempt to translate the Qur'ān was done with 'hostile intentions'. However, since this Latin version was first published, there has been an interest in producing translations of the Qur'ān in other modem European languages; and these later translations have relied heavily on the Latin version of Robertus Retenensis which has polemical purposes (see Arberry, 1996, Abdul-Raof, 2001; Al-Amri, 2005; Fatani, 2006; Elimam and Fletcher, 2021).

The first translation of the Qur'ān into a modern European language was in Italian and published in 1547 by Andrea Arrivabene. Solomon Schweigger completed the first translation in German in 1616. A Dutch version followed in 1641, published anonymously (Fatani, 2006). A French version appeared in 1647, done by André du Ryer. Stubbe (2014, cited in Thomas and Chesworth, 2017, p.301) describes this translation it 'very corrupt, altering and omitting many passages'. Postnikov's Russian translation appeared in 1716. The titles of these early modern European Qur'ān translations reflect their bias against Islam and the Prophet Mohammad. These translated versions were often titled such as 'Alcoran of Mohammad', wrongly naming Mohammad as the author of the Qur'ān.

According to Fatani (2006), most of the early European Christian translators of the Qur'ān were not familiar with Arabic. Their works were not based on the Arabic original and sought to discredit the Qur'ān and Islam, instead of translating it. The roots of European Christians' negative perceptions of Islam are explained by Watt (1972, cited in Gokkır, 2006, p.475), who lists the primary views of medieval Christianity on the nature of Islam:

- It is a falsehood and a deliberate perversion of the truth.
- It is a religion of violence and the sword.
- It is a religion of self-indulgence.
- Muhammad is the Antichrist.

These claims reflect misinterpretations and a lack of understanding of the Qur'ān verses, the Prophet Muhammad's life, and the history of Islam. For example, the claim about the violence of Islam is based purely on the so-called 'sword verses' that are frequently taken out of context, even though the word 'sword' is not mentioned in the Qur'ān (Abdel-Haleem, 2004). Abdel-

Haleem (2004) comments on the misinterpretation and misunderstanding of these verses related to war, which some claim wrongly as evidence for violence in Islam.

In conclusion, most recent Qur'ān translations into European languages have been published by honest scholars attempting to make their renderings very close to the meanings of the Qur'ān (Fatani, 2006).

2.3.4 English Translations

Since the 17th century, Qur'ān translations into English have been amongst the most numerous of the European Qur'ān translations. There are more than 100 complete English Qur'ān translations and 890 editions (Kidwai, 2007). The following sub-sections will summarise some critical and popular Qur'ān translations in English, citing them chronologically and highlighting the prominent shared features in each stage. In our historical review below, the English Qur'ān translations are schematised into three stages: (1) initial stage, (2) second stage, and (3) third stage.

2.3.4.1 Initial Stage

This first stage of Qur'ān translations covers the early attempts, which started in the mid-17th century and concluded in the early 20th century. During this period, some of the English translations resemble some early attempts into European languages, including a lack of knowledge of Arabic, Islam, and the sciences of the Qur'ān.

The first English translation of the Qur'ān appeared in 1649 and was done by Alexander Ross. This English rendering was not direct from the Arabic original. Rather, it was based on the French version of André du Ryer (1647). For various reasons, Ross's translation has been considered unsuccessful by Muslims and non-Muslims alike. For example, Sale (1888) criticises Ross's French-based rendering, noting that Ross had poor knowledge and no proficiency in either Arabic or French, causing him to add a number of 'fresh mistakes' to those made in André du Ryer's French version. Ross's translation is commonly known as *The Alcoran of Mahomet*, which overtly reflects his polemical stance regarding the Qur'ān and Islam (Abdel-Haleem, 2004).

Ross's view of the Qur'ān was also negative. He explicitly states that his translation is 'From the Translator to the Christian Reader' and he harshly criticises the Qur'ān as being 'so rude, and incongruous a composure, so farced with contradictions, blasphemies, obscene speeches, and ridiculous fables' (Ross, 1649, p. A2). According to Abdel-Haleem (2011), Ross labelled

the Qur'ān 'the heresy of Mahomet' to explain to his intended readers why he had published this 'dangerous book'. Kidwai (2007) also notes that Ross's translation labels Islam a 'Turkish' religion.

A century later, in 1734, George Sale produced an English Qur'ān translation titled, *The Alcoran of Mohammad*. According to researchers such as Abdul-Raof (2001) and Kidwai (2007), Sale's translation was based on Marracci's Latin version, which aimed to discredit Islam. Sale's work was then used to translate the Qur'ān into other European languages, although it abounds in errors, omissions, mistranslations, and illegitimate additions (Kidwai, 2007). One of Sale's illegitimate additions was 'You O Arabians', or 'of Mecca', in place of 'O mankind', phrases which do not appear in the Qur'ān (Kidwai, 2007). His unlawful additions seem to indicate that the Qur'ān is intended only for Arabians. Moreover, the word *Al-ġayb* ('unseen') was mistranslated by Sale as 'mysteries of faith'. Despite Sale having committed these mistakes due to his insufficient knowledge of Islam, his translation was once considered one of the most esteemed in the United States and the United Kingdom, according to Qur'ān translators such as Arberry (1996) and Abdel-Haleem (2004).

In the late 19th century, two popular English Qur'ān translations appeared. The first was by the Revd. J.M. Rodwell in 1861 and the second by E.H. Palmer in 1880. These translators followed the approach of previous works in terms of their hostility to the Qur'ān and to Islam, according to Abdul-Raof (2001) and Kidwai (2007). Rodwell was also the first translator to change the traditional arrangement of the Qur'ānic chapters. This was unacceptable to Muslims, as the finished text was unlike the original arrangement of the Qur'ān. In addition, Rodwell's translation contains incorrect footnotes, which are offensive to Muslims (Abdel-Haleem, 2004; Kidwai, 2007).

Palmer's translation in 1880 also takes a negative view of the Qur'ān. Palmer did not follow Rodwell's rearrangement but had two issues with the language and credibility of the Qur'ān. For the former, Palmer (1880, Ixxvii) wrote that the language of the Qur'ān 'is not elegant in the sense of literary refinement', and the ideas and thoughts of the Qur'ān are expressed in Bedouin language and metaphor. Regarding its credibility, he said that the Qur'ān was written by Mohammed (Palmer, 1880). Since the language of Qur'ān is 'rude and rugged', according to Palmer (1880), he chose to translate it into colloquial English.

The last work in this stage to be reviewed is *The Quran Translated with a Critical Re-Arrangement of the Surahs* by the Scott Richard Bell, who published it in 1937. Bell's view was that the Qur'ān was 'actually written by Muhammad himself' (Bell, 1937, vi). Bell's approach was in line with Rodwell's earlier rearranging of the Qur'ānic chapters. Bell argues

in his preface that the Qur'ānic chapters should not be ordered as they appear the original but should be rearranged to 'unravel the[it] composition' (Bell, 1937). Bell's peculiar rearranging led to his deleting and substituting any elements that he believed to cause confusion in the verses.

He believed that his rearrangement 'contributed to the solution of the larger problem' of understanding the Arabic original (Bell, 1937, vi). According to Kidwai (2007), Bell's translation does not contain major novel features but is only famous for its critical rearrangement. When a translator substitutes elements, deletes parts of sentences, or interferes with the structure of a sensitive text such as the Qur'ān, he leaves doubts as to the quality of his translation. The reader has the right to know what is written in the ST; thus, it should be allowed it to speak for itself. In other words, the translation must act as a mirror of the original, rather than as a mirror of what the translator believes (or does not believe).

To summarise, although translators in this first stage made great efforts, Muslim readers might find these early attempts unacceptable, as they were often intended to discredit their religion or cast doubts on the authenticity of the book in which Muslims believe. Since the publication of Bell's translation, negative views have noticeably reduced, and professional standards of translation were the primary aim of the translations in the second stage, as discussed below.

2.3.4.2 Second Stage

This section reviews the second stage of English Qur'ān translation, which began in the early 20th century and largely concluded in the 1950s. The texts that appeared during this second stage resemble the early attempts in terms of their use of archaisms and biblical expressions. Most of the second stage translations show much respect to the Qur'ān and its language, with a focus on changing the negative attitudes towards Islam and the Qur'ān that are found in some previous works of orientalists. Irving (1992, xviii), a Canadian-American author and translator of the Qur'ān who converted to Islam, states:

> The orientalists who worked for London, Paris, the Hague or Lisbon wanted to control Islam for their own purposes, and they decreed that the word Islam means 'submission'. How far they strayed is now apparent, as the Islamic lands throw off their yokes and try to regain their ancient principles.

In this stage, the most famous English translation is Marmaduke Pickthall's rendering, titled *The Meaning of the Glorious Koran*, published in 1930. Pickthall was the first British Muslim to render the Qur'ān in English. Pickthall was not convinced by some of the early translations. He writes that they 'included commentation offensive to Muslims and employed a style of language which Muslims at once recognise as unworthy' (Pickthall, 1930, vii). Pickthall's translation has been reprinted many times since its first publication.

Pickthall (1930) explains that his version was generally rendered literally and written in a befitting language. Whilst working on his translation, he consulted many well-known exegetes to convey the 'authenticity of traditions' (Pickthall, 1930). The overall form of his language is archaic, and the overall orientation of his translation is very close to the style of the ST, both of which make his version more difficult for present-day readers to understand. In the eyes of many Muslims and researchers, Pickthall's work is one of the most faithful and popular translations. As Kidwai (2007) observes, the text contains no dogmatic matter, allowing readers to obtain first-hand knowledge about the original message.

Four years later, the second-most popular translation in this stage was published by Abdullah Yusuf Ali in 1934. Ali's education and experience in the West and his service to Islam had inspired him to publish an English Qur'ān translation. The first edition originally appeared in Lahore. This first edition is titled *The Holy Qur'an, Text, Translation and Commentary*. The text places the Arabic in front of his English translation. Ali's general view was that his text was an 'English interpretation' of the Qur'ān. According to Ali, this version 'shall be, not a mere substitution of one word for another, but the best expression I can give to the fullest meaning which I can understand from the Arabic text' (1934, vii).

Ali explains his methodology in his preface, noting that his extensive commentary uses some of the early exegeses, such as Abu Jafar Muhammad ibn Jarir Al Tabari, Al Mufradat by Abu Qasim Husain Raghib, Al Kashshaf by Abu Qasim Mahmud Zamakhshari, Al Tafsir Al Kabir by Fakhr Al-Din Al-Razi, and Anwar Al Tanzil by Quadi Nasir Al Din. Ali used these early classical works to obtain 'the general sense of accepted commentaries', although he was not entirely convinced by some of these works, as they belong to different schools of thought, some of which express extreme views (1934, xii).

Another method that Ali used was to provide an introduction for each Qur'anic chapter to explain its subject matter, giving explanatory notes for each verse at the end of each page. His explanatory notes are very long and shed light on historical, geographical, and cultural matters. The purpose of his explanatory notes is to help those English readers who do not have a background in Islam and to explain some of the literal meanings.

Ali's work has gone through more than 200 reprints (Kidwai, 2007). King Fahd Complex for the Printing of the Holy Qur'ān and Amana Publication are reputable publishers that have adopted his translation. In 1980, the Custodian of the Two Holy Mosques, King Fahd ibn Abdul Aziz authorised the General Presidency of the Departments of Islamic Researches, Ifta, Call and Guidance to scrutinise the available English translations of the Qur'ān and to choose the most reliable. The authorised committee was mindful of the considerable criticisms that had been levelled against this translation (Ali, 1983). The committee revised Ali's translation and cleared up various misconceptions in his notes. The committee also revised some words such as '*zakat'* and '*tagut'*, arguing that these words could not be properly translated. They decided to give a transliteration of these words in English, with a brief explanatory note. In 1985, a royal decree was issued by the Custodian of the Two Holy Mosques, King Fahd ibn Abdul Aziz, to select Ali's translation for printing at the King Fahd Holy Qur'ān Printing Complex because of 'its distinguishing characteristics, such as a highly elegant style, a choice of words close to the meaning of the original text, accompanied by scholarly notes and commentaries' (Ali, 1985, vi). This version was given the title, *English Translation of the Meanings and Commentary*. This version has since been distributed all over the world, subsidised by the Saudi Government until 2002, according to Ahmed et al. (2015).

Some parts of Ali's translation have been criticised from various perspectives. One criticism came from Irving, a translator of the Qur'ān. According to Irving (1992), Ali's commentary is satisfactory, but Ali's language contains extra words that neither clarify the original nor embellish its true meaning: 'True embellishment is a simple telling word that does not detract, but carries the mind directly to the meaning' (Irving, 1992, xviii). In addition to the above criticisms, Mohammed (2005) points out problems in Ali's footnotes. He (2005, p. 60) claims that 'Ali constructed his oeuvre as a polemic against Jews'. He also indicates that Ali's translation has lost its influence because of its archaic language. However, some of Mohammed's claims are not grounded in substantial supportive or empirical evidence, and some researchers have contradicted this claim of lost influence. For example, Kidwai (2018) states that some previous and current English Qur'ān translations borrow from or heavily rely upon Ali's translation.

Despite the above criticisms, Ali's translation has received approval from several researchers and Qur'ān translators. For example, Abdel-Haleem (2004, xxviii) maintains that Ali's translation is one of the most significant works and 'an extremely useful work, especially his notes and indices, for those who want a fuller and more guided understanding of the background and text of the Qur'ān. His language [however] contains poetic features and archaic words that

make the style outdated'. Kidwai (1987), a bibliographer of English Qur'ān translations, states that Yusuf Ali is without doubt one of only a handful of Muslims to have enjoyed a masterful command of the English language, and this is wholly reflected in his translation. Kidwai (1987) also mentions that, although Ali's translation method is 'more of a paraphrase than a literal translation', it still faithfully conveys the original meanings.

The last work to be reviewed in this stage is the English Qur'ān translation produced in the mid-20th century by Arthur John Arberry, a non-Muslim British orientalist. Arberry was not convinced by the works and methodology used by his predecessors. His capacities in academic writing and his view that previous Western orientalists had taken the wrong approach to English Qur'ān translations led him to publish an honest and rather different translation. The first edition of Arberry's rendering is titled *The Koran: Interpreted*. It was published in 1955 in two volumes, the first having 20 chapters and the second 94. In 1964, his second version was published in a single volume by Oxford University Press. Arberry, like Ali, refers to his version as a 'mere interpretation' of the ST.

Arberry's primary motivation was his view that the previous Qur'ān translations had ignored the composition-evoked emotion and rhythmical effects of the Arabic Qur'ān. He specifies in his preface that 'the rhapsodic nature of its composition has been largely lost to ear and sight; by showing the text as here presented, some faint impression may be given of its dramatic impact and most moving beauty' (1964, p. xii). Abdel-Haleem, a highly prolific Qur'anic translator, observes that Arberry displays great respect for the language of the Qur'an, 'particularly its musical effects' (Abdel-Haleem, 2004, xxvi).

According to Arberry (1964), he writes in clear, unmannered English, in contrast with the Biblical style preferred by some of his predecessors. Arberry chose not to include footnotes so that the smooth flow of the Arabic Qur'an would not be interrupted (Arberry, 1964). His purpose here is clearly that of reading enjoyment, although some researchers have criticised this lack of footnotes. One of the criticisms of Arberry's rendering is that its language is unidiomatic and very close to the grammatical structure of the ST, and the lack of comments and footnotes make his rendering more challenging to understand for those unfamiliar with the ST (Abdel-Haleem, 2004). Similarly, Kidwai (2007) concludes that Arberry's lack of footnotes makes the text inappropriate for uninitiated readers who are keen to fully comprehend the message of the Qur'an. However, Arberry (1996, p.28) insists that these elements are unnecessary 'because notes in plenty are to be found in other versions, and the radiant beauty of the original is not clouded by such vexing interpolations'.

Islamic exegeses and recognised Arabic dictionaries are helpful resources for explanations of the nuances of Qur'ānic words and terms in different verses and contexts. For example, Arberry's translations of 'ذَٰلِكَ الْكِتَابُ لَا رَيْبَ ۛ فِيهِ ۛ هُدًى لِّلْمُتَّقِينَ' (Q2:2) and 'وَلَا تَعْزِمُوا عُقْدَةَ النِّكَاحِ حَتَّىٰ يَبْلُغَ الْكِتَابُ أَجَلَهُ' (Q2:235) are, respectively, 'That is the Book, wherein is no doubt, a guidance to the godfearing', and 'And do not resolve on the knot of marriage until the book has reached its term' (Arberry, 1955). The word 'book' in the first verse is rendered acceptably here. In contrast, in the second verse, 'book' is not a precise rendition of the intended meaning of the ST. The phrase يَبْلُغَ الْكِتَابُ أَجَلَهُ is translated by most Qur'ānic translators as, 'the prescribed period reaches its end' (Abdel-Haleem, 2004), or 'the term prescribed is fulfilled' (Ali, 1934). Some Qur'ānic words require different renditions for different contexts. Thus, Islamic exegeses and well-known Arabic dictionaries are good references for better renditions. Another issue in Arberry's rendering is that it does not include any general background about the topics of the chapters, and this would enhance understanding of his translation.

Arberry has also received other criticisms. According to Kidwai (2007, p.118), Arberry did not follow his predecessor orientalists such as Bell, Palmer, Ross, and Rodwell in terms of polemical leanings, but his translation 'should be approached with caution'. Kidwai (2007) notes that Arberry mistranslates or omits various small portions – and sometimes large sections – of Qur'anic verses. Arberry seems to omit some parts of verses for the sake of rhythm. He states in his preface, 'I have preferred to indicate these terminations and connections by rounding off each succession of loose rhythms with a much shorter line' (Arberry, 1996, p.25). In addition to making some of the points also noted above, Shah (2017) criticises Arberry's rendition from another point of view, stating that this version does not follow an appropriate system of transliteration and is lexically inconsistent across verses or chapters. For example, 'al-Rahman' is translated as 'the Merciful' in Bismillah and 'the All-Merciful' in Al-Fatihah (Shah, 2017).

However, despite these deficiencies, no one can deny that Arberry produced a genuine and remarkable work. Mohammed (2005, p.63) notes that Arberry translated the Qur'an into comprehensible English and 'separated text from tradition. The translation is without prejudice and is probably the best around. The Arberry version has earned the admiration of intellectuals worldwide'. Without doubt, Arberry's rendering goes beyond previous English translations by non-Muslims in its quality and approach (Kidwai, 1987).

Arberry's rendering has been widely circulated around the world. It has appeared in 30 reprints, but it has been reprinted less frequently in recent decades, according to Kidwai (2007). Arberry's rendering seems to have been the last significant translation in this stage to use

archaic English. Although some Qur'ān translations were produced in more-or-less archaic language after and before Arberry's version, these were less widely circulated than Arberry's version. Other translations, such as Asad (1981) and Ahmadi translations by Muhammad Ali (1917), Sher Ali (1955), and Farid (1969) have also been criticised by critics on dogmatic grounds. Some Muslim theologians disagree with Asad's 'rationalistic' approach because his translation contains theological problems, such as his denial of miracles in the Qur'ān (Abdel-Haleem, 2004; Kidwai, 2007; 2018). Muslim theologians also disagree with the heretical beliefs of Ahmadism. According to Kidwai (2007, p.194), Ahmadiya translations contain 'all trappings of Qadyanism [i.e. Ahmadiya belief] – projecting Mirza Ghulam Ahmad as the Promised Messiah' and claiming that the Prophet Jesus (peace and blessings be upon him) was buried in Kashmir. According to Kidwai (1987), Qadiyanis (Ahmadis) are considered to have abandoned Islam.

2.3.4.3 Third Stage

This final section on the history of English Qur'ān translations reviews contemporary English texts. This period began in the 1950s and continued into the present century, with the translations largely produced by Muslim translators. The aims at this stage have been to produce texts that employ contemporary English and are more readable, avoiding language that reduces the comprehensibility of the text (see Chapter 3).

The first attempt at an English Qur'ān translation that avoided any cryptic language was by the Iraqi Jewish N. J. Dawood in 1956. He consulted well-known exegetes such as Al-Zamakhshari, Al-Baiḍāwī, and Al-Jalalayn. Although Dawood's version is not free of errors and inaccuracies that are offensive to Muslims, according to Abdel-Haleem (2004) and Kidwai (2007), it is free of the archaic biblical words that had been a prominent feature of previous Qur'ān translations. One of Dawood's objectives in his version was 'to present the modern reader with an intelligible version of the Koran in contemporary English' (Dawood, 1990, x). Since the publication of Dawood's rendering, one aim of subsequent Qur'ān translations has been simplicity of language.

The most popular translation in this stage was written by two Muslims, M. T. Al-Hilali and M. M. Khan. This rendering was first published in 1977 in the United States, in 1978 in Turkey, and in 1985 in Saudi Arabia. It was subsidised by the King Fahd Complex for the Printing of the Holy Qur'ān, following the replacement of the Abdullah Yusuf Ali translation. Al-Hilali and Khan's translation was selected by the King Fahd Complex because its commentary is based on well-known exegetes and authentic hadiths. The language of Al-Hilali and Khan's

translation is free from archaisms but not of grammatical, lexical, stylistic, and discoursal errors, according to Jassem (2014). Kidwai (1987; 2007) does not rank Al-Hilali and Khan's translation – or some of the other translations by Muslims or non-Muslims – as 'significate ventures' in the history of Qur'ān translation.

Some researchers have criticised Al-Hilali and Khan's translation of some Qur'ānic verses as antisemitic and anti-Christian (see Mohammed, 2005; Wild, 2015). However, these researchers fail to criticise some of the earlier works by non-Muslims for their hatred of and bias against Islam and the Qur'ān. Wild (2015), in particular, claims that the King Fahd Complex replaced the Abdullah Yusuf Ali translation because the translator was a 'Bohra Shi'ite' and Shi'ites are vastly 'unwelcome' in Saudi Arabia. This latter accusation is not based on substantial evidence. It seems that some false allegations do not serve the translated Qur'ān as a product but increase sectarian conflicts within Islam.

Since the late 20th century, English Qur'ān translators have focused on simplicity, clarity, and accuracy. One translation was written by three American women who had converted to Islam: Emily Assami, Mary Kennedy, and Amatullah Bantley. This was the first Qur'ān rendering produced by women and was titled, *The Qur'an: Arabic Text with Corresponding English Meanings*. It was published in 1997 and is known as '*Saheeh International*'. The objectives of Saheeh International were to correct specific errors in previous works concerning the standard practices of Sunni beliefs; to produce simple and uncomplicated language for the benefit of all readers; and to avoid any Arabic transliteration, as found in previous works, because transliteration is not always beneficial for those readers with limited knowledge. Saheeh International stands out as one of the important efforts in the realm of English Qur'ān translation. It represents the Qur'ānic text in simple and clear English for those unfamiliar with the Qur'ān, and it generally represents a faithful rendition of the original. Saheeh International appears to be one of many Qur'ān translations that have not been fully critically examined.

One of the most distinguished modern English Qur'ān translations appeared at the beginning of the 21st century and was written by Muhammad Abdel-Haleem. Abdel-Haleem is a Professor of Islamic Studies at the School of Oriental and African Studies, University of London. He has published many articles and several books related to Qur'ānic studies, the Arabic language, dictionaries in Arabic and Qur'ānic usage, and other comparative studies related to Qur'ānic exegesis, Islam, and the Bible.

Unlike some of his predecessors, Abdel-Haleem has a high level of academic knowledge and substantial experience of Qur'ānic studies. He published his Qur'ān translation following seven years of work on the text. According to Abdel-Haleem (2004), some of his predecessors

translated the Qur'ān into archaic language, using incoherent and alien structures that are very difficult to comprehend. Abdel-Haleem's intention was to produce a coherent and flowing translation in an easy, readable style, using modern English, free from any archaic elements (Abdel-Haleem, 2004). The first edition of his translation was published in 2004 in London, titled, *The Qur'an: A new translation*. He published a revised version in 2005 and another in 2010. The latest version is titled, *The Qur'an: English Translation and Parallel Arabic Text*. Abdel-Haleem invited his graduates and MA students in Arabic and English, who were English native speakers, to read his translation prior to publication (Abdel-Haleem, 2004).

His new title includes the word 'translation', unlike those of his Muslim predecessors, who deliberately avoided referring to their works as translations, which is a hotly debated issue among Muslim scholars (see Mustapha, 2009). Most of his Muslim predecessors described their works as 'meanings' or 'interpretations' of the Qur'ān. However, Abdel-Haleem (2004, vi) explained why he named his new version as translation:

> It is vital to stress that it is only the Arabic text that is recognized by Muslims as the Qur'an and no translation can substitute for it. Any translation is no more than an interpretation or form of exegesis to attempt to explain, in the target language.

Abdel-Haleem's translation is seen as a point of departure in terms of explaining his professional translation skills compared to some previous translations. Abdel-Haleem clarifies his translation commission (or 'translation brief'), explaining that the methodology he used to enhance the accuracy and clarity of his work focuses on nine translational issues: general style, intertextuality, context, identifying aspects of meaning, Arabic structure and idiom, pronouns, classical usage, paragraphing and punctuation, and footnotes and explanatory introductions. These issues are briefly explained in turn below.

According to Abdel-Haleem (2004), the general style of his 'new' translation goes one step further than previous works. He says that this new version is translated into a readable style, a contemporary form of language, avoiding the use of archaisms that tend towards vague connotations (Abdel-Haleem, 2004). Abdel-Haleem's purpose in using this style is to make his version accessible to anyone who speaks English, including all modern English readers.

Intertextuality is generally defined as 'significant external relationships to other texts' (Dickins et al., 2017, p. 293); but in the Qur'ān, the term – as used by Abdel-Haleem – is about the relationship of a verse in a chapter to another verse in another chapter. Intertextuality is a significant factor in translation studies, but it received little attention from Abdel-Haleem's

Qur'ān predecessors. Abdel-Haleem explains the meaning of intertextuality in the Qur'ān as follows: there are some verses that explain others, and this is considered by Ibn Taymiyya (d. 1328 CE) to be the most accurate means of understanding and elucidating the meanings of the Qur'ān. This method is known in Arabic as تفسير القران بالقران '*tafsīr al-Qur'ān bi-l-Qur'ān*', 'interpreting the Qur'an through the Qur'an'. It is an exegetical method for meaning extraction. Abdel-Haleem (2004, xxix) uses footnotes in his translation to explain intertextuality across verses to clarify 'the meaning of ambiguous passages of the Qur'an'.

One of the issues discussed in his version is the context of certain verses. Abdel-Haleem (2004, xxx) describes context as 'crucial in interpreting the meaning of any discourse, Qur'anic or otherwise'. He suggests that understanding the occasion or revelation of a verse or chapter is important for producing an accurate rendition.

Another issue that Abdel-Haleem addresses in his version is the need to identify aspects of meaning – crucial for distinguishing finer shades of lexical-based context. Abdel-Haleem (2004, xxxi) states that some key Qur'ānic terms are repeatedly used 'in the Qur'an with different meanings for different contexts'. This feature in Arabic is called الوجوه والنظائر '*al-wujūh w al-naẓā'ir*' and is known in English as polysemy. '*Amr'*, '*jihād'*, '*awliyā'*, and '*taqwā'* are examples of key Qur'ānic terms. For the sake of consistency, these are usually translated in all verses as if they conveyed a single meaning, which results in a misrepresentation of the intended meaning (Abdel-Haleem, 2004). Abdel-Haleem (2004, xxxi) notes that, 'it is important for the translator to recognize when it is appropriate to be consistent in the translation of a repeated term, and when to reflect the context'.

One of the difficulties that Abdel-Haleem faced is that literal translation of Arabic structure and idioms into English can be problematic. Abdel-Haleem (2004) indicates that forms which are unnecessarily close to the Arabic structure and idioms are avoided in his version, since they sound mannered in English.

Pronouns are another issue addressed in his version, regarding the shifts from plural to singular (or vice versa) between verses of the Qur'ān. This shift is a stylistic feature known in Arabic as التفات '*iltifāt'* (for more see Abdul-Raof, 2018; 2019). Abdel-Haleem (2004) says that ambiguity and distortion of meaning may occur if the proper reference of the pronoun is not correctly identified. According to Abdel-Haleem (2004, xxxiii), one solution for translating pronoun shifts in the Qur'ān is to add a word which makes plain the reference – 'Prophet', for example – 'where it is clear that it is he who is being addressed, to make the passages as clear in English as they are in Arabic. This is particularly important in passages where, within the same verse, there is a shift between plural and singular address'.

The classical usage of some Arabic terms is another important point, as the meaning and use of some classical Arabic words in the Qur'ān have been changed in modern Arabic. For example, '*walad'* in classical (Qur'ānic) Arabic means the non-gender-specific 'child' or 'children', whilst in modern Arabic it can only mean 'boy' or 'son' (Abdel-Haleem, 2004, xxxiii). Abdel-Haleem's version relies on classical Arabic dictionaries (such as the Lisān al-ʿarab by Ibn Manḍūr, al-Qāmūs al-Muḥīṭ by al-Fayrūz ʾābādī, and al-Muʿjam al-Wasīṭ by the Arabic Language Academy in Cairo) to address the semantic changes of key terms.

Paragraphing and punctuation in Abdel-Haleem's version are unlike those in previous works. He (2004) notes that a very long, continuous paragraph is not normal in English and is overwhelming for those who are not familiar with the Arabic paragraphing conventions. Thus, he (2004, xxxiv) adopts the prose paragraph, which he believes 'is extremely important for the referencing and cross-referencing which contributes so much to understanding the meaning of the text'. Such combining or separating of verses might present the text in a flowing and coherent manner, but his new paragraphing may be subject to criticism because it appears in the style of prose writing. He also uses English punctuation marks (such as commas, full stops, colons and semicolons, question marks, dashes, and quotation marks), which do not exist in the original Qur'ān. Abdel-Haleem (2004, xxxiv) says that, 'the Qur'ān has its own system of marking pauses'. He uses English punctuation marks in his version to 'present dialogue and direct speech', 'to make the sentence structure and the flow of ideas clearer', and 'to break what might appear to be a single sentence into smaller units' (Abdel-Haleem, 2004, xxxiv).

Finally, footnotes and explanatory introductions are also discussed in Abdel-Haleem's version. He (2004) notes that extensive comments are avoided, as these are overwhelming for readers; instead, short footnotes are adopted to explain allusions, references, and the cultural background. He also provides an explanatory introduction for each chapter to help readers, explaining the title of each chapter and giving information about the main themes.

In comparison with other works, Abdel-Haleem's rendering stands out as a major achievement because it reflects his expertise in Qur'ānic studies and in translation studies. His rendering also contains a very valuable introduction to Islam and the Prophet's life and a brief review of previous Qur'ānic translations – although the information in this introduction is criticised harshly by Mohammed (2005, p.69), who questions the reliability and authenticity of Muslim tradition and writes that Abdel-Haleem's introduction 'reflects the age-old Muslim tradition, and therefore, simply reports the Muslim stories without any question as to their reliability'. Mohammed (2005, p.69-70) also adds that Abdel-Haleem must be aware of 'the haziness of early Islamic history', concluding that 'Abdel-Haleem has incorporated his

doctrinal bias into his translation'. However, Mohammed declares these older Muslim stories to be unreliable, without giving any examples. Thus, it seems that Mohammed's assessment of the Qur'ānic translators is purely subjective and may be due to his own doctrinal bias.

Abdel-Haleem's work was not the final translation in this stage to employ modern language. There are many translations such as *The Qur'an: the first American Version*, published by T. B. Irving in 1985; *The Qur'an: A Modern English Version*, produced by Majid Fakhry in 1996; *The Qur'ān: A New Interpretation*, by Colin Turner in 1997; *The Noble Qur'an, a New Rendering of its Meaning in English*, by Abdulhaqq and Aisha Bewley in 1999; *The Qur'an: A new Translation*, by Thomas Cleary in 2004; and *Towards Understanding the Ever-Glorious Quran*, by Mohammad Ghali in 2008. Some of these works include the translators' own interpretations of Islamic beliefs, making the translated Qur'ān a never-ending circle. This is not to say that exegetical work should not be included in a translated Qur'ān. Rather, the translated Qur'ān should possibly offer a rendition close to the pure meaning or truth message of the Qur'ān. Any exegetical interpretations should be given in footnotes.

To summarise, the first stage of English Qur'ān translations was inspired by a desire to refute Islam and to declare the Qur'ān as having been written by Mohammed himself. This does not devalue the great contributions of some Christian translators in translating the Qur'ān for those who are not speakers of Arabic. Some of the early English Qur'ān translations were heavily criticised by Muslim and non-Muslim translators for their omissions, mistakes, unacceptable inferences, and misinterpretations. Some of these translations were not based on the Arabic original. Nonetheless, any damage done – knowingly or unknowingly – to the Qur'ān or to Islam ultimately encouraged many Muslim and non-Muslim translators in the second stage to produce faithful translations that showed much respect for the language and for believers of the Qur'ān. Most Qur'ān translations by Muslim and non-Muslim translators in the second and third stages were derived directly from the Arabic Qur'ān and attempted to present as faithfully as possible the religion and Islamic perspectives of the Qur'ān, which gave the texts a better position within Muslim communities and academic circles. Some of these Qur'ān translations lack annotations, whilst others are annotated with footnotes and commentaries that are helpful for those who are unfamiliar with the Arabic Qur'ān. Some of the Qur'ān translations from the second and third stages can be found in every part of the world, accessible to Muslim and non-Muslim readers.

2.3.4.4 Literalness of English Qur'ān translation

The English Qur'ān translations reviewed so far can be generally classified into two types, according to their overall translation strategy: semantic translation and communicative translation (Abdul-Raof, 2001). These two concepts are presented by Newmark (1981) to distinguish the overall strategy of a translation product. Semantic translations are defined as 'attempts to render, as closely as the semantic and syntactic structures of the second language allow, the exact contextual meaning of the original' (Newmark 1981: 39). This type of translation is similar to others, such as literal translation, Nida's formal equivalence (1964), and Nord's documentary translation (1988), all of which are considered to be source-text-oriented. In semantic translation, the translator preserves the 'local colour' or linguistic aesthetic values of the ST in the TT (Newmark, 1981; Nord, 1988). An example of a semantic translation can be found in the following rendition:

(وَلَا تَجْعَلْ يَدَكَ مَغْلُولَةً إِلَىٰ عُنُقِكَ وَلَا تَبْسُطْهَا كُلَّ الْبَسْطِ فَتَقْعُدَ مَلُومًا مَّحْسُورً) [الإسراء: 29]

1- [And keep not thy hand chained to thy neck, nor outspread it widespread altogether, or thou wilt sit reproached and denuded. Q17:29] (Arberry, 1955).

2- [Make not thy hand tied (like a niggard's) to thy neck, nor stretch it forth to its utmost reach. Q17:29] (Ali, 1934).

These two translators, Arberry and Ali, employ a semantic rendering by maintaining the general word order of the ST in the TT. Although this type of ST-oriented translation, semantic translation, is faithful to the ST, it sometimes leads to peculiarity and meaningless in the TT – for example, in the fogginess of 'not thy hand chained to thy neck' in these two translations. This type of translation has been criticised as 'characterised by dogged adherence to source language syntax' (Abdul-Raof, 2001, p.22). This adherence always sounds 'unnatural' and is often 'meaningless' in English (Abdel-Haleem, 2004). Akbar (1978) and Abdul-Raof (2001; 2018) state that most English Qur'ān translations suffer from problems of literalness. Semantic translation can be found, in particular, in English Qur'ān translations by Rodwell (1861), Palmer (1880), Ali (1934), Bell (1937), Arberry (1955), Pickthall (1969), and Hilali and Khan (1977).

In contrast, a communicative translation, as described by Newmark (1981, p.39), 'attempts to produce on its readers an effect as close as possible to that obtained on the readers of the original'. Communicative translation is similar to Nida's dynamic equivalence (1964) and Nord's instrumental translation (1988), all of which involve TT-oriented translation. In this

method, the form of the source language (SL) is respected, but the text is more oriented to the norms of the target language. In communicative translation, according to Newmark (1988, p.41), 'both content and language are readily acceptable and comprehensible to the readership'. We can use the Arabic verse cited in the previous example to provide an example of a communicative translation:

> [Do not be tight-fisted, nor so open-handed that you end up blamed and overwhelmed with regret. Q17:29] (Abdel-Haleem, 2004).

Abdel-Haleem uses a communicative rendering in the above verse. Such a rendition maintains the form of the ST, but it is used to reproduce the SL in clearer and smother English, according to the norms of the target language. His rendering does not stay as close to the SL structure as the previous two examples of semantic translation. Communicative translation can also be found in particular in English Qur'ān translations by Dawood (1956), Akbar (1978), Irving (1985), Sarwar (1982), and Turner (1997), Saheeh International (1997), and Abdel-Haleem (2004).

2.4 Retranslation

As previously shown, the Qur'ān has been 'retranslated' several times into most of the major world languages. There is some consensus between translation scholars as to what this concept of 'retranslation' means (Bensimon, 1990; Berman, 1990; Susam-Sarajeva, 2003; Venuti, 2004; Paloposki and Koskinen, 2004, 2010; Gürçağlar, 2009). The notion is typically defined as either the act of translating again the same text into the same language or the same text that was retranslated (Gürçağlar, 2009). In simple words, any translation completed after the first translation of the same text is considered a 'retranslation'.

However, the term has also been used to describe a 'relay' or 'indirect' translation (Gürçağlar, 2009; Paloposki and Koskinen, 2010), which denotes the action of translating a text not from the original language of a text but from the language of other translations (Shuttleworth and Cowie, 1997). Paloposki and Koskinen (2010) note that some scholars consider indirect translations to be a kind of retranslation, but this usage is more confusing than helpful. They affirm that even when indirect translations are excluded, the classification of retranslation is extremely complicated. For example, Jianzhong (2003) differentiates between two categories of retranslations:

- '**Direct retranslation**': translation derived directly from the original, more than twice

- ‘**Indirect re/translation**’: translation not derived directly from the original but from other translations

There is no straightforward way of distinguishing a translation as either an indirect or a direct retranslation, unless the translator indicates the direction of the transfer from the original. This is particularly so for the hundreds of translations of the Bible and the Qur’ān.

In the case of English Qur’ān translations, it is difficult for a retranslation researcher to classify which of the English Qur’ān translations is a direct or indirect translation or retranslation. However, given that the first attempt at an English Qur’ān translation – by Alexander Ross in 1649 – was based on du Ryer’s French translation, Ross’s work can be considered an indirect translation. George Sale made the second attempt at an English Qur’ān translation in 1734, and this was based on Maracci’s Latin translation; thus Sale’s work is also an indirect retranslation. All English Qur’ān translations after the works of Ross and Sale might be regarded as direct retranslations; certainly, most subsequent Qur’ān translators have explicitly claimed that their renderings are based on the original Arabic. Conducting Qur’ān translation research from the perspective of retranslation, the chosen English Qur’ān translations used in the analysis can be classified into two types: two earlier translations and a recent retranslation.

Paloposki and Koskinen (2010, p. 294) propose a broader definition of ‘retranslation’, which they consider from two perspectives:

> Retranslation (as a product) denotes a second or later translation of a single source text into the same target language. Retranslation (as a process) is thus prototypically a phenomenon that occurs over a period of time, but in practice, simultaneous or near-simultaneous translations also exist, making it sometimes hard or impossible to classify one as a first translation and the other as a second translation.

Paloposki and Koskinen propose two processes of retranslation: synchronic and diachronic. They also suggest another categorisation of ‘adaptations’, denoting a retranslation completed for a new readership (e.g., a translation for children).

Pym (1998) similarly proposes three types of retranslation: periodical, passive, and active. ‘Periodical retranslations’ occur over a period of time, such as retranslations of the Bible (Pym, 1998). ‘Passive retranslation’ is when a retranslation is completed without any awareness of previous translations, having been separated by time or geographical boundaries (Pym, 1998). ‘Active retranslation’, in contrast, is when a retranslation is done with prior knowledge of

previous translations, sharing the same culture and time, but where there are disagreements about translation strategies (Pym, 1998).

Brisset (2004, cited in Gürçağlar, 2009, p.236) highlights the importance of probing retranslations from diachronic and synchronic points of view, as retranslations can highlight the 'the cognitive and creative aspects of translation'. Paloposki and Koskinen (2010, p.298) note that although research on retranslation has been ongoing since the beginning of the 21st century, research on retranslation is providing new insights into a variety of issues in translation studies, 'ranging from the ethical to the aesthetic'. Venuti (2004) shows that retranslations can highlight differences and changes related to commercial publishers, the translator's intentionality, new interpretations of the ST, and new norms in the target culture. Many scholars have paid special attention to retranslation, which has been primarily studied in the context of canonical literary texts (Bensimon, 1990; Berman, 1990; Susam-Sarajeva, 2003; Venuti, 2004; Paloposki and Koskinen, 2004, 2010; Gürçağlar, 2009). However, the practice is not limited to canonical literary texts and also concerns religious texts.

According to Paloposki and Koskinen (2010), retranslation is rarely discussed outside of the context of literary translated texts, and even the numerous retranslations of the Bible have been overlooked, leaving a significant gap in the research. English Qur'ān translations are also under-researched in the context of retranslation. Most published studies on English translations of the Qur'ān have focused on the (in)equivalence or (un)translatability of the linguistic factors, including syntactic, phonetic/prosodic, rhetorical, sematic, and pragmatic features, with comparative analyses of the Arabic Qur'ān and English translations or of the discourse features of the Arabic Qur'ān (see Abdul-Raof, 2001; 2004; 2005; 2007; 2018; 2019; Al-Amri, 2008; Boullata, 1988; Aldahesh, 2014; Dastjerdi and Jamshidian, 2011; El-Zeiny, 2011). However, the role of retranslation in changing the readability and comprehensibility of English Qur'ān translations has not been explored. This research aims to identify the textual features that make some English retranslations of the Qur'ān more readable and comprehensible than others.

There has been slightly more attention given to Qur'ān retranslations in languages such as Korean and Japanese. El Damanhoury (2015) discusses the motivations for choosing certain techniques for three Japanese retranslations of the Qur'ān published in 1957, 1972, and 2014. El Damanhoury found that all three translations adopted foreignising strategies. In contrast, the original translation adopted more domesticating strategies. This result, in El Damanhoury's view, is compatible with the retranslation hypothesis of Bensimon (1990) and Berman (1990), which is discussed in the following section.

However, Choi and Kim (2021) examined how core concepts and terms in the Qur'ān were (re)translated in two Korean Qur'ān texts: Kim Yong-sun's first translation (1970) and Choi Young-gil's retranslation (1997). They found that the first translation was more oriented towards the ST culture, whilst 'Choi's retranslation strategy, which was to adapt the translation more in line with the receiving culture and social context, is also identified in the translation of the men-women relationship' (Choi and Kim, 2021, p.19). Their results reveal that Choi's retranslation was more domesticating – a result which is incompatible with the retranslation hypothesis (RH).

2.5 Retranslation Hypothesis and its Origin

Retranslation, as a phenomenon in translation studies, goes back to the 1990s, when two French scholars, Paul Bensimon and Antoine Berman, published two articles in the journal *Palimpsestes*. These articles express the theoretical assumptions about retranslation that are now referred to as the RH (Gürçağlar, 2009; Paloposki and Koskinen, 2004). These works paved the way for studying the motivations behind this phenomenon in translation studies and understanding the differences between initial translations and subsequent retranslations from different perspectives and identifying their proximity to – and distance from – their STs.

In his preface to a volume of *Palimpsestes*, Bensimon (1990) claims that there are necessary differences between initial translations and subsequent retranslations. Initial translations, as Bensimon (1990) argues, act as 'introductions', which naturalise 'foreign works' and introduce them into a given culture to ensure better reception of the translated work in the target culture. By contrast, subsequent retranslations, in Bensimon's view, do not seek to naturalise 'foreign works'; instead, they 'pay more attention to the letter and style of the source text and maintain a cultural distance between the translation and its source, reflecting the singularity of the latter' (Gürçağlar, 2009, p.233).

In the same volume of *Palimpsestes*, Berman (1990) writes that translation is not always a complete act; it usually seeks completion through retranslation. This completion, in Berman's view, is connected 'with the success of a translation in getting closer to the source text and in representing the encounter between the translator and the language of the original' (cited in Gürçağlar, 2009, p.233). According to Berman (1990), all translations face the issue of 'failure'. This failure, in his view, is 'at its peak' in initial translations.

Berman (1990) adds that the failure can remain present, but from the multiplicity of new retranslations sometimes emerges a great translation. Berman suggests that retranslations tend to progress through repetition and the passage of time. Susam-Sarajeva (2003, p.2) views the

phenomenon of retranslation in terms of ‘a history-as-progress model’. Antoine Berman, mirroring Goethe (1990, cited in Deane-Cox, 2014, p.3), observes that ‘the accomplishment of any human action demands repetition’. However, the act of retranslating existing translations might be redundant – for scientific or technical texts, for example (Gürçağlar, 2009) – unless it is motivated by a logical reason (Deane-Cox, 2014). Jianzhon (2003, p. 194) explains the purpose of retranslation as a movement towards a better translation:

> The significance of retranslation lies in surpassing. If the retranslation is not better than the former one(s), the retranslation will not be worth a penny, and it will not be encouraged but criticized.

Some of Bensimon and Berman’s claims in relation to the RH have been tested and challenged by several authors, with some opposing and some supporting the ‘domesticating’ claim regarding initial translations. For example, Gürçağlar (2009) states that some case studies have found no evidence to support the RH’s claim of a domesticating strategy in initial translations, noting that not all initial translations adopt domesticating strategies and not all later retranslations are foreignising. Partially supporting and validating the claim of a domesticating strategy, Paloposki and Koskinen (2004) showed that several initial literary translations of fiction into Finnish were ‘more or less domesticated’. In the context of literary translation, they note that the ‘domesticating phase’ might occur during the first stage as a feature in the development of a literary work, whilst some first literary translations may not be domesticating. In the case of English Qur’ānic translations, there are more than 110 complete translations and 890 editions (Kidwai, 2007; 2019), which makes it difficult to either reject or confirm the claim. This is because these texts vary substantially in their strategies in terms of wording, meaning, and form.

In a further argument for RH, Berman (1990) suggests that we retranslate because translations become old over time. In Berman’s words, the originals remain ‘young’, whilst translations ‘age’, making it possible to publish a new retranslation. The ‘ageing’ of existing translations is often linked with translations of canonical literary texts (Susam-Sarajeva, 2003). Since language never remains static, the lexical choices of earlier translations sometimes need rewording and updating (Hanna, 2006) to better meet the expectations and needs of modern readers (Susam-Sarajeva, 2003).

Gürçağlar (2009, p.234) points out that not all translations are influenced ‘by the passage of time, and not all translations will necessarily age’. This is partially true, when translations of

the same original are completed in a short period of time. Susam-Sarajeva (2003) notes that all subsequent retranslations of Barthes's works into Turkish were published within a very short time span, from 1975 to 1990, and they were largely uninfluenced by factors such as ageing and inadequacy of the existing translations. However, they were influenced by the needs and expectations of Turkish literary criticism. Similar cases can be identified in English Qur'ān translations produced in a short span of time. Kidwai (2018) observes that 40 complete new translations (albeit originally retranslations) of the Qur'ān into English were published in the period from 2000 to 2017, which amounts to more than two translations every year; and this was the consequence of ideological missionary work and sectarianism within Islam. In summary, the RH argument concerning 'ageing' could not be the single influencing factor in the act of retranslation.

Translation scholars have also proposed various other explanations for this phenomenon. Chesterman (2000) applies RH as an exemplary hypothesis to discuss three models of translation: comparative, process, and causal. Chesterman (2000, p.24-25) offers several explanations for retranslation:

- Later translators take a critical stance on the earlier translation, seek to improve on it.
- The existence of the earlier translation in the target culture affects the potential reception of the new one.
- The target language has developed and allows the translator more freedom of movement.
- [The] target culture translation norms have become relaxed, allowing a closer link to the source text.

Understanding of the act of retranslation is 'often based on a linear idea of progress' (Susam-Sarajeva, 2003, p.3). This line of progress leads either towards the receiving system (Susam-Sarajeva, 2003; Paloposki and Koskinen, 2003) or towards the ST (Susam-Sarajeva, 2003). In the latter case, several aspects of the ST can be lost in the initial translations, and retranslations reduce these losses. In the former case, as Paloposki and Koskinen (2003, p.23) note, retranslation is 'supplementary', which 'suggests a positive attitude towards difference', with different translated versions for different groups of readers in the receiving system and the ST categorised for different discourse, such as children's literature.

To conclude, there is more research into retranslations of literary works than of religious texts, despite the long history of the latter. Religious and literary retranslations have specific discourse characteristics, but they are often motivated by (dis)similar factors (linguistic, social,

ideological, and political). Translation scholars have discussed these factors, citing them as the primary motivations for the retranslation of specific texts. The reasons for retranslation in general and Qur'ān retranslation in particular are discussed in the following section.

2.6 Motives for Retranslation

Motives for retranslation are diverse. Some are positive and legitimate. According to Paloposki and Koskinen (2010, p.296), retranslations can be inspired by a need for 'increased knowledge of the source text, author and culture'. This is evident in the case of English translations of the Qur'ān (see section 2.3.4). Some of the initial English translations were derived from other European translations, rather than the Arabic original. These initial translations were strongly criticised by researchers and other translators, with the texts reflecting their translators' lack of familiarity with Arabic, Islam, and the science of the Qur'ān. Later translators attempted to reduce the deficiencies of these initial Qur'ān translations by introducing greater familiarity with the history, content, and context of the Qur'ān; the multi-layered meanings of the Qur'ān; idiomatic differences between Arabic and English; and the exegetical works of Muslim scholars. This greater understanding of the Qur'ān led to higher quality translations, in the eyes of many researchers and readers – although some of these later translations by Muslim translators were motivated by ideological and sectarian differences (see Robinson, 1997).

Other researchers have cited other factors that motivate a translator to produce a new retranslation. A primary factor in the decision to retranslate a text may be 'changing social contexts and the evolution of translation norms' (Gürçağlar, 2009, p.234). For example, Du-Nour (1995, cited in Gürçağlar, 2009) examined retranslations of children's literature into Hebrew to explore the 'linguistic and translational norms' of retranslations published in different periods of time. Du-Nour found that there were linguistic and stylistic differences between earlier translations and subsequent retranslations. In her study, readability was found to be a central concern of subsequent retranslations, whilst earlier translations were identified as being less readable and having a more biblical style.

In their exploration of social and cultural norms, Choi and Kim (2021) looked at two Korean translations of the Qur'ān, taking retranslations as data to reflect on how certain sensitive verses (related to women's status) had been retranslated to bring them into line with the social and cultural norms of Korean culture. They found that a verse related to women's status in the Qur'ān was translated in the first Korean translation as 'a man is superior to a woman', whilst the later retranslation presented this verse as 'a man is a guardian of a woman', which brought 'the phrase more in line with the Korean culture in the 1990s that increasingly appreciated

gender equality' (Choi and Kim, 2021, p.20). They state that certain sensitive verses are softened in Choi's retranslation and argue that retranslation is 'motivated by new ideologies and religious factors in a different social context' (Choi and Kim, 2021, p.3).

Retranslations, as Gürçağlar (2009) claims, also emerge as a result of ideological and political factors, for literary translations in particular. Gürçağlar gives the example of Monteiro Lobato, a Brazilian author, translator, and publisher who presented his own political leanings in his retranslations of children's classics. However, the retranslation of sacred scriptures is the practice most obviously subject to political, ideological, and theological influences. With regard to Bible retranslations, Koskinen and Paloposki (2003) examined various Finnish retranslations in the context of the RH. One of their cases was the retranslation of the Gospel of St. Matthew into Finnish in 1969. They note that this retranslation is more 'domesticated' and more easily readable than earlier versions 'but is also impinged with the translator's (then) Marxist ideology', concluding that 'first translations are found lacking, but the perceived need for a supplement may take different forms. […] Retranslation is a result of shifting needs and changing perceptions' (Koskinen and Paloposki, 2003, p.22-23). Similarly, Venuti (2004) mentions that the King James Bible was retranslated during the 17th century to strengthen the authority of the Anglican Church in England, in opposition to the Protestant versions of the Bible published by translators such as William Tyndale and Richard Taverner. According to Venuti, Tyndale's translations, in particular, were deemed heretical by the Roman Catholic Church because of the influence of Protestantism.

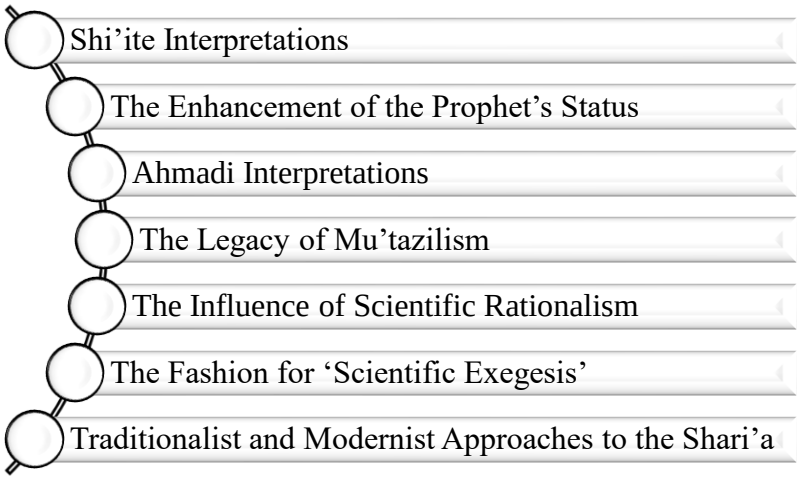


Figure 2.1: Sectarian and Ideological Differences in Qur'ān Translations by Muslims

In the case of English Qur'ān translations, Robinson (1997) scrutinised ideological and sectarian bias in English translations of the Qur'ān produced by Muslim translators. His study

did not treat Qur'ān translations as retranslation data, although all Qur'ān translations considered in his study are retranslations, nor did Robinson consider English Qur'ān translations by Christian or Jewish translators, as he felt that these needed to be assessed in a separate study. Robinson's study identified seven types of sectarian and ideological bias in these translations by Muslim translators. These are listed in Figure 2.1 and summarised below.

Shi'ite interpretations: Some Shi'ite translators such as Ahmed Ali and M. H. Shakir retain Arabic words such as 'imam' in their renderings, capitalising it to strengthen their Shi'ite beliefs in the *Imamate* (Robinson, 1997). Yusuf Ali also had the same rendering, but removes any Shi'ite connotations from his translation. Some Shi'ite translators such as Ahmed Ali and M. H. Shakir also display Mu'tazilite leanings in their translations of several verses (Robinson, 1997).

The enhancement of the Prophet's status: Some translators express their extreme devotion to the Prophet Muhammad in their renderings. Some insert additional words such as 'O beloved' or 'O beloved prophet' that show 'an exaggerated reverence for the Prophet', although these additional words do not exist in the original Arabic text (Robinson, 1997, p.276). According to Robinson, some of these leanings are derived from classical Sufi sources (for more on Sufism see Sheikh, 2020).

Ahmadi interpretations: This concerns another group of Muslims, similar to the previous in terms of their radical beliefs, but these are followers of Mirza Ghulam Ahmad and are generally considered non-Muslim, according to Robinson (1997). Some Ahmadi translators (such as Zafrulla Khan) insert additional phrases such as 'like one crucified to death' (Q.4:157) to show their Ahmadi leanings, although these additional words are not found in the original Arabic text (Robinson, 1997).

The legacy of Mu'tazilism: This was a theological movement known for its rationalism. Mu'tazilites believed that terms in the Qur'an relating to God's 'throne', 'face', and 'hands' are figures of speech (Robinson, 1997). For example, the word عرش *ʿarš* in some verses is usually rendered as 'throne'. Some Mu'tazilite-influenced translators remove this and replace it with words such as 'power', as metaphorical terms (Robinson, 1997). Their belief is seen by Sunni Muslims as heretical. However, some Shi'ite and Sunni translators – such as Muhammad Asad and Yusuf Ali – also have Mu'tazilite leanings (Robinson, 1997). Some of Yusuf Ali's Mu'tazilite-type comments in his footnotes were removed when his translation was sponsored by King Fahd Complex for the Printing of the Holy Qur'ān.

The influence of scientific rationalism: This group is also influenced by the Mu'tazilites. They do not take as fact the stories of miracles in the Qur'ān, nor do they believe in *jinns*,

which they believe are against the laws of nature (Robinson, 1997). Translators such as Muhammad Ali, Asad, Zafrulla Khan, and Ahmed Ali add footnotes to their renderings of the Qur'ānic miracles that they consider to be metaphorical and allegorical (Robinson, 1997).

The fashion for 'scientific exegesis': This is a group of Muslims who link Qur'ānic notions to those of the modern sciences. They believe that 'the Qur'an was scientific in the sense that it mentioned things unknown to Muhammad and his contemporaries—things which were only discovered by European scientists many centuries later […]. Modern science, far from posing a threat to Islam, actually strengthened the case of the Muslims by furnishing an additional proof of the Qur'an's supernatural origin' (Robinson, 1997, p.271). Some translators insert footnotes to explain scientific facts in some verses. However, those who use scientific exegesis to link certain verses to scientific facts should be approached with caution. This is because they frequently use the notion of the inimitability (*iʿjāz*) of the Qur'ān to exaggerate or wrongly interpret some Qur'ānic scientific-type statements or claim support for scientific facts that are not actually mentioned in the Qur'ān. Yacoub (2005, p.862-3), in a discussion of mistakes made by those who use scientific explanations to interpret the verses, states:

> some Muslim intellectuals make a mistake when they try to relate some of the Qur'ān's signs or aspects of its inimitability to modern discoveries and theories, thinking that they raise the status of the Qur'ān in this regard. However, their scientific claims and explanations leave the Qur'ān prone to contradiction and criticism [my translation].

Yacoub clarifies that the Qur'ān is a book of guidance. It is not intended to detail scientific miracles and is not a book of astronomical, medical, or natural sciences. This is not to deny that we can benefit from what modern science has revealed to clarify the scientific miracles in the Qur'ān (Yacoub, 2005).

Traditionalist and modernist approaches to the Shari'a: The most accurate and reliable procedures for interpreting the Qur'ān are proposed by Ibn Taymiyya (d. 1328), as follows: a Qur'ānic passage should first be interpreted according to other Qur'ānic passages; authentic hadiths are then used to explain the Qur'ānic passage. When there is no explanation either in the Qur'an nor in the Sunna:

> one should have recourse to interpretations attributed to the Companions; failing that, the interpretations of the Successors should be taken as authoritative provided that they are in

> agreement; finally, if the Successors disagree, or are silent on the matter, the meaning of the text should be determined on the basis of the Arabic language (Robinson, 1997, p.274).

Most Sunni traditionalists follow this approach. Among the translators who follow this approach are Al-Hilali and Khan (1977) and Saheeh International (1997). In contrast, modernists reject this because they question 'the authority of the Companions and Successors' (Robinson, 1997, 274). Examples of modernist translators include Ahmed Ali, Zafrulla Khan, Asad, and Irving.

Although Robinson's study did not consider English Qur'ān translations in the context of retranslation, the sectarian and ideological tendencies identified in this important study are seen as the primary motivations for the repeated Qur'ān retranslations into English. Readers of the Qur'ān can be dissimilar in terms of their beliefs, interests, and values. Thus, Qur'ān translators have frequently retranslated the Qur'ān and given it various interpretations, according to their readership's values and beliefs. As Venuti (2004) notes in his article discussing the nature of retranslation, the practice may aim to inscribe a set of values (linguistic and literary, religious and political, commercial, and educational) in the retranslated work.

Venuti (2004) discusses the motivations of retranslation from different angles that shape its nature. Exploring the Venuti (2004) article on retranslation, Deane-Cox (2014, p.13) writes that Venuti goes further than Pym into the extratextual causes of retranslation, 'situating motivating factors on the levels of canonicity, ideology, economics, and the subjectivity of translator'. With regards to ideology, retranslations are purposefully created with the aim of forming 'particular identities' to strengthen the authority of certain social institutions (Venuti, 2004). Retranslations also help to increase the canonicity of specific foreign texts in the target culture. As Venuti (2004, p.27) puts it, 'a foreign text that is positioned on the margin of literary canons in the translating language may be retranslated in a bid to achieve canonicity through the inscription of a different interpretation'. There is also a relationship between the selection of some translated texts by publishers and their potential commercial success. Venuti (2004, p.30) writes that:

> A commercially oriented publisher may decide to issue retranslations of foreign canonical texts that have fallen into the public domain simply because their canonicity ensures a market demand … Hence, an ideology of commercialism will govern the selection of a foreign text for retranslation and dictate a discursive strategy that enhances the readability of the translation to ensure sales. A publisher driven by a profit motive

> may in fact wish to save the expense of commissioning a retranslation by reprinting a previous translation that has proven itself in the marketplace, even if in a revised version.

With regards to the subjectivity of the translator, Venuti (2004, p.30) points out that retranslation might be merely motivated by 'the retranslator's personal appreciation and understanding' of the ST. For instance, the beauty of the language of the Qur'ān has inspired some Qur'ān retranslators to seek to reflect the aesthetic values of the Qur'ān in the target culture. Arberry (1964, x-xii) explains that the purpose of his rendering of the Qur'ān was to 'improve on the performance of my predecessors […]; to produce something which might be accepted as echoing however faintly the sublime rhetoric of the Arabic Koran', and to avoid the Biblical style that was the fashion of his predecessors. One prolific Qur'ānic translator observes that Arberry displays great respect for the language of the Qur'ān, 'particularly its musical effects' (Abdel-Haleem, 2004, xxvi).

Taivalkoski-Shilov (2015, p.62), summarising the various purposes of retranslation, explains that 'the causes might be linguistic, aesthetic, cultural, ideological, economic, practical, idiosyncratic, and so forth'. Gürçağlar (2009) states that updating or modernising the language of a translated text or correcting mistakes or misinterpretations in initial translations are 'legitimate justifications' for the production of a retranslation. Most of these motivations are exemplified in all types of text, including Qur'ān retranslations. The primary motivations for English Qur'ān translations can be categorised, as outlined previously, according to three stages of their history: (1) initial stage, (2) second stage, and (3) third stage.

Initial stage: This began with the Christian orientalists who first translated the Qur'ān into English. In this stage, translators attempted to understand Islam and the Qur'ān in light of the Bible. However, some of the translations in this stage were driven by a desire to discredit the Islam (see section 2.3.4.1). Some of the renderings were not based on the original Arabic, reflecting the translator's lack of understanding of Arabic, Islam, and the Qur'ān. The translations in this stage were strongly criticised by non-Muslim and Muslim translators and researchers for their omissions, misinterpretations, and altering of the traditional arrangement of the Qur'ānic chapters and verses.

Second stage: The major mistakes made in the first stage – and associated negative reviews – were motivations for the retranslations in the second stage. The translations in this stage were done to change the image of Islam and to reflect the true meanings and Muslim perspective of the Qur'ān. For instance, when Pickthall published his English translation in 1930, he criticised

some previous Qur'ānic translations because they 'included commentation offensive to Muslims and employed a style of language which Muslims at once recognise as unworthy' (Pickthall, 1930, vii). The translations in this stage attempted to strike a balance between defending Islam and the Qur'ān and presenting as faithful a rendering as possible of the ST. However, the translations in this stage were not free of errors and inaccuracies (for more see Kidwai, 1987; 2007; 2018).

Third stage: The translations in the third stage were motivated by a multitude of factors, but mainly by a desire for readability and comprehensibility. According to Irving (1992, xxxvi), the purpose of his translation was to 'make its clear message available for the English-speaking world at the end of the twentieth Christian century' and to present the Qur'ān's message in clear, comprehensible, contemporary English, and using very simple words 'so the Muslim child can understand it easily'. Abdel-Haleem's primary motivations for his recent translation were to surpass previous works in terms of 'accuracy, clarity, flow, and currency of language' and readability, avoiding 'archaisms that tend to obscure meaning', and producing a version that can be read by anyone who speaks English (Abdel-Haleem, 2004, xxix).

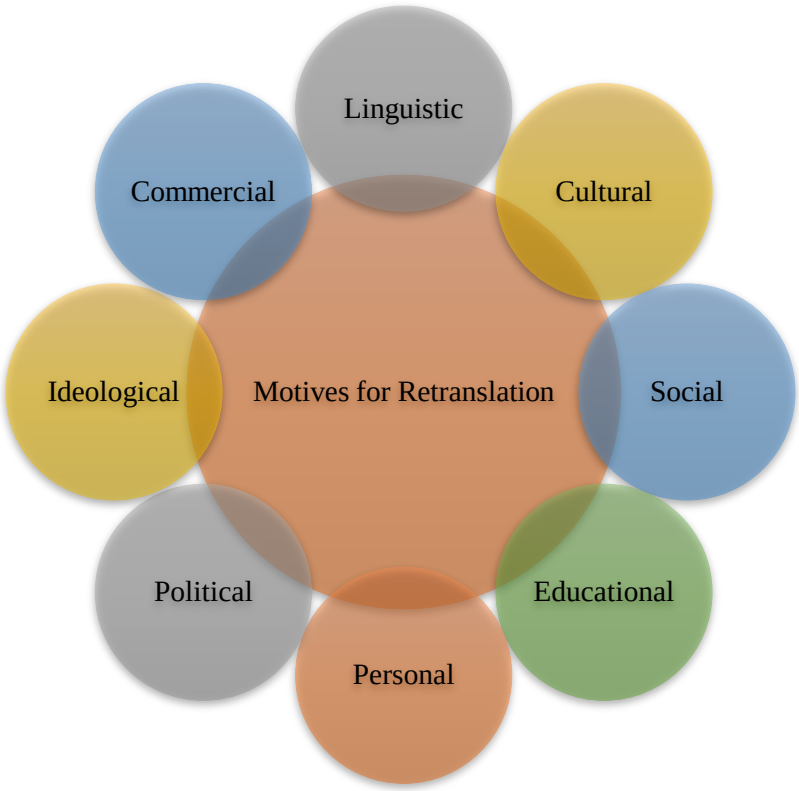


Figure 2.2: Multiple Motivations for Retranslation

To conclude, considering the motivations for retranslation in general and for that of the Qur'ān in particular, it is clear that retranslation can be a process of improvement and change in the receiving language, culture, and norms, rather than simply serving the original text. As Deane-Cox (2014, p.14) argues, 'retranslation as reinterpretation is no longer restricted solely to the service of the ST'. The changes involved in retranslation are driven by a multitude of factors ranging from the linguistic and cultural to the ideological. There is rarely a single reason for creating or choosing to publish a new retranslation. The multiple motivations for retranslation described by previous researchers are worth summarising, and Figure 2.2 provides a visual description of the most common motivations as a preliminary sketch or framework of textual analysis for future research in terms of how retranslation is approached. This research will then focus on the linguistic motivations. This research focuses on retranslation intended to improve the readability and comprehensibility of English Qur'ān translations, with the goal of identifying the linguistic and non-linguistic differences between translations that influence text readability and comprehensibility. The following chapter explores the factors that contribute to text readability and comprehensibility.

2.7 Chapter Summary

This chapter began by exploring the nature of the Qur'ān, how Muslims can understand and approach it through other Islamic sciences, and the opinions of Muslim scholars on the legitimacy (or otherwise) of translating the Qur'ān into other languages. This was followed by a brief historical review of Qur'ānic translation into Asian, African, and European languages. The chapter also discussed the history of English Qur'ān translations, organised into three stages on the basis of their shared features. A detailed background was given on the work of Abdullah Yusuf Ali (1934-1989), Arthur John Arberry (1955-1983), and Muhammad A. S. Abdel-Haleem (2004), who are the primary sources of data in this research (see Chapter 4).

This chapter also presented a discussion of the concept of 'retranslation', with a focus on the different types of retranslation and the classification of English Qur'ān translations into 'initial' translations and 'recent' retranslations, which will be useful in the discussion of the results later in this research. The chapter then moved onto the origin of retranslation as a concept and its development under the so-called 'retranslation hypothesis' (RH), as initiated by French scholars who proposed various assumptions on the nature of retranslation. Following this, the motivations for retranslation in general and for English Qur'ān retranslations in particular were discussed. The motivations for retranslation include commercial, political, ideological, educational, cultural, social, and linguistic factors. The current study focuses on the linguistic, emphasising the role of retranslation in improving the readability and comprehensibility of English Qur'ān translations. Thus, the two notions of 'readability' and 'comprehensibility' are used to present retranslation in a new light. The following chapter will present the factors that influence text readability and understanding, as identified in the field of readability research.

Chapter 3: Readability and Comprehensibility

3.1 Chapter Overview

In this chapter, a review of the literature relevant to this research framework is presented. The chapter begins with a general overview of the definition of readability, with a review of researchers' thoughts on the factors that influence it (Section 3.2) and then explains the relationship between readability and comprehensibility (Section 3.3). Sections 3.4 to 3.7 provide a brief historical overview of the development of readability formulae – primarily in English and then briefly in other languages (such as Arabic) – tracing the features used to probe content difficulty and comprehension in the early studies of readability and classic readability formulae (CRFs). In Section 3.8, the criticisms of readability formulae and their limitations, explaining why they are used in this research to estimate the 'difficulty' of a text, is highlighted. In Section 3.9, a general overview of the computational software Coh-Metrix, which scales content difficulty on multiple levels of text characteristics, is presented. This section details the software's examination of in-depth language features – rather than just surface features, as in CRFs – and explains the use of Coh-Metrix in this research. Section 3.10 describes a text–reader model of text readability and comprehensibility, focusing on text factors (such as style, organisation, cohesion and coherence, and genre) and reader factors (prior knowledge, interest, and motivation) that increase or reduce text readability and comprehension. Finally, there is a summary of the main points discussed in this chapter.

3.2 The Concept of Readability

The Oxford and Cambridge dictionaries define 'readability' as the quality of being 'easy', 'interesting' and 'enjoyable' to read (Cambridge Dictionary, 2013; Oxford Dictionary, 2020). These definitions do not indicate whether this quality is on the side of the reader or the text, nor do they explain the features that can make a text easier or more enjoyable to read.

The most critical early definition comes from Dale and Chall (1949, p. 23), who described readability as follows:

> In the broadest sense, readability is the sum total (including all the interactions) of all those elements within a given piece of printed material that affect the success a group of readers have with it. The success is the extent to which they understand it, read it at an optimal speed, and find it interesting.

Dale and Chall's (1949) definition of readability is wide-ranging and considers three factors that influence the ease of text comprehension: understanding, fluency and interest. Chall (1974, p. 7) adds that the concept of 'readability' has no specific meaning but incorporates three major aspects: 'ease of understanding', 'legibility' and 'interest'.

Readability as *ease of understanding*, according to Chall (1974, p. 7), means that a text's readability level is determined by whether the text contains a large proportion of 'hard words', 'long sentences' and other crucial elements, making it more difficult to understand than one with a smaller number of these features. The ability to read at optimal speed is influenced by the *legibility* of print, which is ensured by 'good paragraphing, introductory phrases in bold type and short chapters' (Chall 1974, p. 7). These factors permit readers to read texts at optimal speed. The last factor, *readability as interest*, reflects that the topic of a text is interesting and that the reader finds it interesting owing to individual preferences.

Klare (1963), reviewing the concept of readability in research, found that it has been used in the field to refer to three major issues:

- the legibility of either handwriting or typography
- the ease of reading due to either the interest value or the pleasantness of the writing
- the ease of understanding or comprehension due to the writing style

(Klare, 1963, p. 1).

The first of these – 'legibility' – concerns typographic issues and the physical appearance of the print, such as font size and type, which some use interchangeably with readability (see Harrison, 1980). Klare (1963) stated that the readability concept of many tended to focus on ease of understanding due to writing style. This suggests that what makes a text easy or hard to comprehend is the text style. This view also reflects the central measurement of most classic readability formulae (CRFs) that predict only 'style difficulty' and 'reading ease' of content (Klare, 1988). In 1935, Grey and Leary surveyed groups of publishers, librarians and teachers on 'what makes a book readable'. They found that style was the second-most crucial factor, after content. Format and organisation were deemed the least important factors influencing readability.

In the late 1960s, readability began to shift towards matching reader factors with text factors. McLaughlin (1968, p. 188), the developer of the 'SMOG' formula, describes readability as 'the degree to which a given class of people find certain reading matter compelling and comprehensible'. McLaughlin's definition focuses on the reader's own characteristics alongside the level of a text that is 'compelling'. McLaughlin also argues that readers only continue to read a text when they understand it and match their interests; thus, a reader's choice of reading material is indicative of the readability of the text. Klare (1963) found that preferences for or choices of reading materials were strongly affected by 'ease'. According to DuBay (2007, p. 70), 'easier readability of a text has more benefits for those of less knowledge and interest than those of more'.

In the early 1970s, Gilliland (1972) stated that readability requires matching readers' interests and background knowledge with the content's difficulty. Gilliland (1972, p. 13) explains the importance of matching the reader and the text as follows:

> In a scientific article, complex technical terms may be necessary to describe certain concepts. A knowledge of the subject will make it easier for a reader to cope with these terms and they, in turn, may help him to sort out his ideas, thus making the text more readable. This interaction between vocabulary and content will affect the extent to which some people can read the text with ease. Similarly, some uncommon words and ideas may be familiar to some readers because of their experience and interests. If the text itself is difficult because of the ways in which the ideas are expressed, then interaction between vocabulary and reader's knowledge will affect readability.

Readability, in Gilliland's view, is that some factors from the reader's side and others from the text's side interact with one another to influence text readability, suggesting that complex passages and a lack of familiarity with text topics may make a text difficult to understand. Therefore, Gilliland (1972, p. 12) defined readability as 'the study of this problem of matching the reader and the text has come to be called readability'.

Finally, Pikulski (2002, p. 1) views the concept of readability as 'the level of ease or difficulty with which text material can be understood by a particular reader who is reading that text for a specific purpose'. Pikulski's definition is not significantly different from McLaughlin's, but it is concerned with the *purpose* of the reading. Pikulski argues further that the concept of readability must be based on the characteristics related to both text and reader when assessing a text's readability.

In summary, some researchers have related the readability concept to text factors alone, as was particularly common at the beginning of readability research in the 1920s (see section 3.4). The difficulty of writing style is a central issue in readability research. In contrast, other researchers have linked the readability concept to both text and reader factors, marking the beginning of a 'text-reader model of readability' (Singer, 1988). In conclusion, the readability concept seems to involve an interaction between the characteristics of the text and the reader, with both playing a significant role in determining the comprehensibility of a text.

3.3 Readability and Comprehensibility

Readability and comprehensibility are slippery concepts. They are sometimes blurred and used interchangeably in the literature, inducing an unclear distinction between them (see Harrison, 1980; Jones, 1997; Smith and Taffler, 1992). Some researchers do not differentiate between readability and comprehensibility, treating them as synonyms. Adelberg (1893), for example, stated that the terms refer to the same concept. However, Smith and Taffler (1992) disagreed with Adelberg, arguing that the concepts are measurable in different ways. Harrison (1980, p. 33) differentiates between the concepts as follows:

> Readability is an attribute of text; comprehension is an attribute of readers. There is therefore a fundamental difference between the two concepts. Having made that distinction though, a moment's reflection makes it clear that the concepts are intimately related, in that very often when we use the term readability we mean in effect the comprehensibility of a text.

Harrison's view is preferable, as the lack of a clear distinction between the two might conceal the text and reader factors that make texts difficult or easy to understand. Jones (1997, p. 105) similarly argues that 'readability and understandability measure two closely associated, but intrinsically different, attributes'. Jones (1997) proposes that comprehensibility is a broader notion that measures the interaction between reader and text and is contingent on the reader's background, prior knowledge, purpose, interest and general reading ability, while readability is contingent on textual features and measures the difficulty of textual features.

In any case, readability and comprehensibility influence one another. Kolahi et al. (2012) and Jones (1997) emphasised that readability is a pre-requisite of comprehensibility, but there is no guarantee of individual comprehensibility. This means that the level of text readability reflects its comprehensibility. Readability comes before comprehensibility; thus, the former usually helps the latter, though some reader characteristics can influence text comprehensibility – if a reader is uninterested or unfamiliar with a text topic. Klare (1963, p. 17) explains this further:

> More readable written material is likely to produce greater comprehension, learning and retention than less readable only when one or more of the following factors are present: the less readable is much harder than the more readable, and clearly beyond the reader's usual level; reading time is limited; the reader does not have a large amount of background or experience with the topic being covered; and, the reader has a relatively strong set to learn.

Comprehensibility and readability have primarily been topics of concern in the field of readability research. However, some recent translation studies have shown an interest in the readability and comprehensibility of translated texts. In 2015, a special issue of *Translation and Comprehensibility* was dedicated to the topic, suggesting possible new directions for assessing the comprehensibility of translated texts (Maksymski et al., 2015). Wolfer (2015) introduced three models, taken from educational psychology, which can be used to optimise or assess the comprehensibility of a text or translation.

In the same volume, Hansen-Schirra and Gutermuth (2015) explore a range of multi-methods for measuring the comprehensibility of a translated text, noting the close relationship between comprehensibility and some concepts in translation studies such as 'simplification' and 'explicitation' that can be used to increase the readability and comprehensibility of translations. However, the concept of comprehensibility as the focus of this volume is examined from various perspectives. Jensen (2015), for example, notes the usefulness of using

intralingual translation to improve comprehensibility in translations from English into Danish. She argues that to maximise comprehensibility in translations, there is a need for intralingual translation and plain English to achieve plainer and more comprehensible language in the target text.

To summarise, Harrison (1980) distinguishes, significantly, between readability and comprehensibility. For this research, the readability of a translation is primarily viewed from the perspective of the text itself and is defined as the extent to which a translation can be easily read – regarding text properties, including the ease of the lexical and syntactic elements, the cohesion of linguistic elements and non-linguistic elements such as the clarity and organisation of the page layout elements and the genre and register of the text.

These text properties fall into four broad categories: style, organisation, cohesion and coherence, and genre (discussed in detail in section 3.10). These are important for obtaining a clearer view of Qur'ān translation readability from the text side and determining its impact on the comprehensibility of a Qur'ān translation. This research employs Harrison's definition (1980) of the concept of comprehensibility, which concerns reader characteristics or 'an attribute of readers'. Thus, comprehensibility is never static and can change. The chosen Qur'ānic versions might be incomprehensible to readers lacking or having low knowledge of the Qur'ān, while being entirely comprehensible to other readers with prior knowledge and familiarity with it. Thus, difficulties in text comprehension might arise from either the reader or the text itself (see Section 3.10). The following sections elucidate the most important textual aspects that contribute to text readability and comprehension.

3.4 Developmental History of Readability

Research into readability in education began in the 1920s (Klare, 1963; Chall, 1988). However, readability research in general has a long history. Chall (1988, p. 15) provides an overview:

> The study of readability, in the sense of language comprehensibility, has a long history. It has deep roots in the classical rhetoric of Plato and Aristotle and in the vocabulary analyses of the Bible by ancient Hebrew scholars.

The field developed from two research areas: 'studies of vocabulary control' and 'studies of readability measurement' (Chall, 1988). The two areas had a similar purpose and emerged at the beginning of the 1920s, while 'both sought objective means of measuring the difficulty of

printed materials for learning and for comprehension,' though the first area was 'more prevalent during the earlier years' (Chall, 1988, p. 15).

The first research area was 'concerned with the vocabularies that would be most effective for learning to read from reading textbooks' (Chall, 1988, p. 15). Thus, readability research in this area focuses on how the word difficulty of primary-level textbooks is processed and understood by children. The second research area, 'studies of readability measurement', was concerned with 'the comprehension difficulty of content area textbooks and other materials written for students of middle and upper elementary grades, high school, college and adults' (Chall, 1988, p. 15). In this area, readability measures' focus is on textual characteristics and reader variables that affect text comprehension, as discussed previously regarding defining the concept of readability.

Since the 1920s, readability research in English has proposed several hundred formulae that 'reliably and validly distinguish easier from more difficult texts' (Chall, 1988, p. 15). A hundred variables have been used to predict text readability and difficulty objectively (see Klare, 1963, 1974, 1988; Gilliland, 1972; Harrison, 1980; Chall, 1988). DuBay (2004, p. 4) notes that 'by the 1980s, there were 200 formulas and over a thousand studies published on the readability formulas attesting to their strong theoretical and statistical validity'. These English readability formulae are detailed in Section 3.7.

The growth of readability research in other languages began in the 1950s (Klare, 1974, 1988), with the first Spanish formula published in 1951 (Klare, 1988). According to Klare (1988), researchers have since written eight books, developing formulae for languages such as Afrikaans, Swedish, Russian, Hebrew, Hindi Chinese, Spanish, French, German, Korean, Finnish, Vietnamese and Danish. These readability formulae have been developed for various purposes: (1) for the benefit of learners in the United States studying other languages, (2) for 'the evaluation of second language texts for language study and for use with recent immigrants', and (3) to assess the reading levels of schoolbooks and other written materials (Rabin, 1988, p. 46). The history of readability research in these foreign languages is detailed in Klare (1974; 1988), Davison (1986), and Rabin (1988).

Research in Arabic readability formulae or even in Arabic readability itself remains rare, as many researchers have observed (see Dawood, 1977; Al-Heeti, 1984; Al-Ajlan et al., 2008; Al-Tamimi et al., 2014; Saddiki et al., 2015; El-Haj and Rayson, 2016). The developmental history of formulae for measuring Arabic text readability dates back to a master's thesis submitted by Dawood in 1977. Dawood's formula uses variables similar to those found in other formulae (including the English), with five as follows: average word length, average sentence

length, average word frequency, percentage of nominal clauses and percentage of definite nouns. In 1988, Al-Heeti proposed a formula in his PhD thesis that comprises just one variable: average word length.

In 2014, Al-Tamimi et al. developed an Arabic formula known as the 'automatic Arabic readability index'. This has three variables: the number of characters, the average number of characters in a word and the average number of words in a sentence.

In 2016, El-Haj and Rayson proposed the 'open source metric for measuring Arabic narratives' (OSMAN). OSMAN is unlike any previous Arabic readability formulae, as it includes diacritics in the calculation process to count the number of syllables per word. El-Haj and Rayson criticised previous Arabic formulae, particularly that of Al-Heeti. They stated that, 'although appealingly simple, the formula does not work as a good indicator of Arabic text readability, especially given that Arabic is a highly inflectional and derivational language and word length by itself does not reflect difficulty' (El-Haj and Rayson, 2016, p. 251). The OSMAN formula includes seven variables that measure Arabic text readability, as follows:

- number of words
- number of sentences
- number of 'difficult' words (words of more than five letters; i.e., long words)
- number of syllables in a word
- number of characters (excluding digits)
- number of 'complex' words (words of more than four syllables)
- number of *faseeh* words (complex words with any of the following Arabic letters – ء, ئ, ؤ, ذ, ظ – or which end with وا, ون)

(El-Haj and Rayson, 2016, p. 253).

A brief overview is provided in the following sections on the developmental history of readability measures in English, divided into four sections: the definition of a 'formula', early studies of readability, CRFs and criticisms of CRFs.

3.5 What is a Readability Formula?

Before discussing the development of English readability formulae, it is essential to identify what a formula is and how an evaluator can measure and predict text readability. Klare (1974), a leading scholar in readability research, highlighted two ways of addressing readability: 'measuring' and 'predicting'. *Measuring readability* is a non-predictive approach that involves

reader judgements and comprehension tests, while readability can be predicted using a formula (Klare, 1974). A readability formula 'uses counts of language variables in a piece of writing in order to provide an index of probable difficulty for readers. It is a predictive device in the sense that no actual participation by readers is needed' (Klare, 1974, pp. 63). The creation of a readability formula reduces the subjectivity required to assess text difficulty or readability, providing quantitative objective measures of whether the variables in a text 'yield a measurable difference in reader performance or behaviour' (Klare, 1974, p. 129).

3.6 Early Studies of Readability

At the beginning of the 20th century, leading educationists attempted to develop systematic and objective ways of measuring text difficulty based on assessing the words most commonly used in children's textbooks. Others have validated these measures and considered other readability measures that might affect text comprehension and the difficulty of passages. These early studies started by estimating the level of text difficulty in elementary-school texts, gradually moving onto adult reading materials. This growth reflected an increased understanding of issues that arise when measuring readability, highlighting the need for easy and accurate measures of content difficulty. A brief overview of the most influential studies in early readability research from the 1920s to 1935 is provided in the following paragraphs.

Edward Thorndike (1921)

In 1921, Edward Thorndike published *The Teacher's Word Book* (Chall, 1974; Klare, 1963; 1968), which presented a list of 10,000 words used in children's literature. The words were used in Biblical and English classics: elementary textbooks, cookbooks, texts on sewing, farming and trades and daily newspapers. Thorndike highlighted the frequency of usage of the respective words and proposed that texts with fewer common words would be less familiar and more challenging to comprehend. Klare (1968, p. 20) explains the role of word frequency in readability:

> Not only do humans tend to used [sic] some words much more often than others, they recognise more frequent words more rapidly than less frequent, prefer them, and understand and learn them more readily. It is not surprising, therefore, that this variable has such a central role in the measurement of readability.

Thorndike's word-frequency list inspired new methods of measuring word difficulty in children's textbooks. It helped teachers choose the most appropriate schoolbooks for their students' ages. Thorndike's work 'laid the foundation for almost all the research on readability that would follow' (DuBay, 2004, p. 12). One problem, however, is that reliance on such a list can be time-consuming, with researchers and teachers mostly preferring less time-intensive ways of detecting word-usage frequency (see Section 3.9).

Lively and Pressey (1923)

Thorndike's work paved the way for the study of Lively and Pressey in 1923, titled *A Method for Measuring the Vocabulary Burden of Textbooks*. This was the first publication of a readability formula in English (Klare, 1963, 1968), and it focused primarily on children's textbooks. To assess the 'word difficulty' of a children's textbook, the following variables were measured:

- number of different words or vocabulary range
- number of technical words not on Thorndike's list of 10,000 words
- median index number of the words on the Thorndike list

(cited in Klare, 1963; Chall, 1947, 1974).

DuBay (2004, p. 14) states that the third variable in the list above is 'the best indicator of the vocabulary burden of these reading materials: the higher the index number, the easier the vocabulary; the lower the index, the harder the vocabulary'. Lively and Pressey – and Thorndike – use 'vocabulary' as the primary criterion for measuring text difficulty. The Lively–Pressey formula, however, has not been incorporated into any software for practical use and does not use grade levels to reflect the difficulty level of a text.

Vogel and Washburne (1928)

Shifting from the work of Lively and Pressey, Vogel and Washburne (1928) attempted to consider word difficulty and include other features related to sentence structure. The authors cite the following six features as determining the grade level of a text: vocabulary difficulty, sentence structure, parts of speech, paragraph construction, general structure and physical makeup. In their paper, the authors examine each of these features and measure their use in various books to identify 'whether there was a definite rise or fall from grade to grade' (Vogel and Washburne, 1928, p. 375). Based on their investigation of these features, they propose the Winnetka formula, which includes the following variables:

- number of different words per 1000 words
- number of prepositions
- number of words not on the Thorndike list
- number of simple sentences in a sample of 75 sentences

(Vogel and Washburne, 1928, p. 377).

Their work was the first study of 'the structural characteristics of the texts' to predict the grade level of a children's book (Chall, 1974, p. 21). Klare (1963) stated that their formula was the first study to provide a basis for constructing a formula, concluding that their formula provides a prototype for contemporary readability formulae.

Dale and Tyler (1934)

Early studies sought to objectively assess the reading difficulty of textbooks for children. However, these studies did not explain the factors that could hinder adult readers with different reading skills, nor did they elucidate the appropriate text features or criteria for adult readers of various levels. The first attempt to estimate text difficulty for adult readers with limited reading ability was published in 1934 by Dale and Tyler, titled *A Study of the Factors Influencing the Difficulty of Reading Materials for Adults of Limited Reading Ability*. Dale and Tyler (1934) used a variety of comprehension tests to assess the comprehension of adults with poor reading skills. The texts concerned health-related topics and were taken from magazines, newspapers and textbooks. The authors sought to identify the written properties that influence reading difficulty based on eight factors from the surface features of the language. They found that for adults with limited reading ability, the difficulty of written materials was influenced by three factors, as shown in their following formula:

- number of different technical words
- number of different 'difficult' non-technical words
- number of indeterminate clauses

(Dale and Tyler, 1934, p. 401).

Gray and Leary (1935)

In 1935, Gray and Leary further explored the factors that make books appropriate for adult readers. They published a significant book on readability under the title *What Makes a Book*

Readable (1935). Their work focused on adults with limited reading capacity, examining 800 adults in the United States. The authors conducted a large-scale survey of publishers, librarians and teachers to identify the factors that hinder text readability and comprehension for adult readers with low reading skills. Gray and Leary (1935) identified 228 elements affecting text readability, which they classified under four headings in the following order of importance:

- content
- style of expression
- format
- features of organisation

The authors conclude that content and style of expression (word length, sentence length, word familiarity, sentence simplicity and idea flow) are the most important factors, with organisation and design being less important. Based on the four categories, Gray and Leary propose a formula that focuses on style factors to measure a text's difficulty in terms of five components: the number of words per sentence, the number of different 'difficult' words, the number of personal pronouns, the percentage of different words and the number of prepositional phrases.

Lorge (1939) adds another variable to Gray and Leary's formula – namely, 'a weighted index of word difficulty', which is based on the Thorndike list. He ran multiple correlations between Gray and Leary's variables and comprehension scores based on McCall-Crabbs standard test lessons in reading,[4] concluding that his combination outperformed Gray and Leary's formula. He concludes that vocabulary is 'the most significant indicator or predictor of passage difficulty' (Lorge, 1939, p. 229). Lorge was first to use the McCall-Crabbs test to construct a regression equation (Klare, 1963; Chall, 1974; Bailin and Grafstein, 2016).

The work of Gray and Leary 'stimulated an enormous effort to find the perfect formula, using different combinations of the style variables' (DuBay, 2004, p. 19). The two works described above focus on various factors – in addition to word difficulty – in their measurements of text readability for adults with limited reading abilities. Other early readability studies conducted during the 1920s can be found in Chall (1947; 1974) and Klare (1963).

[4] 'These are a series of 376 passages of children's readings already graded in difficulty on the basis of comprehension of questions at the end of each passage' (Chall, 1947, p. 6).

3.7 Classic Readability Formulae

Early studies of readability laid the testing ground for CRFs. These studies explored various linguistic variables to predict text readability or difficulty, but none of the variables were used to produce a grade level or a reading ease score. Klare (1963, p. 51) demonstrates that the focus of classic formulae was 'efficiency and simplicity of use'. Most CRFs are easy to apply, and they provide quantitative scores and grade levels that can be used as approximations of reading ease (Chall, 1947; Klare, 1963; Gilliland 1972; Binkley, 1988). A brief overview of the best-known and most-used classic formulae published between the 1940s and the 1970s is provided in the following sections.

The Flesch Formula (1943)

In 1943, Rudolf Flesch published his first readability formula for predicting the difficulty of adult reading material in his PhD dissertation *Marks of a Readable Style*. Flesch found that the existing adult formulae had various limitations: (1) they were inappropriate for adult materials, as they heavily emphasised vocabulary and word lists over other factors, and (2) they limitedly considered word abstractness, which was a significant measure of content difficulty (Klare, 1963; Chall, 1974). Flesch's formula (1943) comprised three variables: the number of affixes, the average sentence length and the number of personal references.

In 1948, Flesch produced his second formulae in two parts. The first part of his formula predicts 'human interest' based on two variables: the number of personal words (such as nouns and pronouns) and the number of personal sentences (such as quotation marks, exclamations). The second part of his formula predicts the reading ease of passages based on two variables: word length and sentence length. This is the Flesch formula (Harrison, 1980, p. 77):

$$\text{Reading ease score} = 206.835 - (84.6 \times \text{SYLLS/100w}) - (1.015 \times \text{WDS/SEN})$$

Where SYLLS/100w = syllables per 100 words
and WDS/SEN = number of words per sentence

The Flesch Reading Ease formula produces a reading score from 0 to 100, with 0 to 30 being 'very difficult' and 90 to 100 being 'very easy' to understand.

The value of Flesch's contribution lies in the formula's use of word length and its relationship with word frequency. Research findings have shown that short words are usually

concrete and frequent, while long words are usually abstract, rare and general (see Klare, 1968; Chall, 1974; McNamara et al., 2014). Another contribution of Flesch's formula is that it can produce a US grade level, as shown in the below Flesch's description of reading ease. Flesch's formula provides 'a more realistic picture of the difficulty of materials at the upper levels than the raw scores derived from an analysis based solely on the McCall-Crabbs passages' (Chall, 1974, p. 33).

Description of Style	*Average Sentence Length*	*Average No. of Syll. per 100 Wds.*	*Reading Ease Score*	*Estimated School Grades Completed*	*Estimated Percent of U.S. Adults*
Very Easy	8 or less	123 or less	90 to 100	4th grade	93
Easy	11	131	80 to 90	5th grade	91
Fairly Easy	14	139	70 to 80	6th grade	88
Standard	17	147	60 to 70	7th or 8th grade	83
Fairly Difficult	21	155	50 to 60	some high school	54
Difficult	25	167	30 to 50	high school or some college	33
Very Difficult	29 or more	192 or more	0 to 30	college	4½

Flesch's Description of Reading Ease Score (1949, p. 149)

In 1975, Kincaid et al. conducted a study using Flesch's formula and other older formulae, testing them for use on US Navy materials. They simplified the Flesch formula and converted it to a grade-level scale to produce the 'Flesch-Kincaid grade level'. In 1978, the US Department of Defence approved this as a new process for testing the readability of technical materials (DuBay, 2004). The Flesch-Kincaid formula is one of the most commonly used, tested and reliable measures in the history of readability research (Klare, 1963; 1968; Chall, 1974). It has been incorporated into Microsoft Word to check the reading ease and grade levels of their texts.

The Dale and Chall Formula (1948)

In 1948, Dale and Chall developed a formula to assess adults' written materials. Their formula is one of the most commonly used and reliable measures for predicting readability (Klare, 1963; 1968). The formula uses a word list of 3,000 words as a criterion for predicting vocabulary difficulty, although Dale was among those who criticised Thorndike's word frequency list (DuBay, 2004). Eighty percent of their list can be understood by fourth-grade readers.

Like previous researchers, Dale and Chall used the McCall-Crabbs test as a criterion for measuring comprehension. They acknowledged that while this criterion had deficiencies, it was

the best available at the time (Dale and Chall, 1948). The Dale-Chall formula has two variables: (1) percentage of hard words and (2) average sentence length, as shown in their formula (Dale and Chall, 1948):

Score = (0.1579 × PHW) + (.0496 × ASL) + 3.636
Where score = reading score of a pupil who could answer one-half of the test question.
PHW = percentage of hard words from outside their list of 3,000
ASL = average of sentence length

The Dale-Chall formula produces a reading score from 4.9 and below to 10.0 and above, with 4.9 and below equivalent to the 4th grade and 10 equivalent to the adult level. Dale and Chall (1948) found that their formula score correlated closely with the McCall-Crabbs scores but had a greater prediction value (for one factor – words outside of their list of 3,000 words) than the Lorge formula and the original Flesch formula (1943). They concluded that vocabulary load best contributes to text difficulty, with sentence length second.

Gunning's FOG Formula (1952)

In 1952, Robert Gunning, who was a counsellor in clear writing, published his book, *The Technique of Clear Writing*, with his FOG formula. Gunning published his formula after he noticed that business and newspaper writings were 'above the danger line of reading difficulty', because they contained a lot of 'fog' and unnecessary complexity in writing (Gunning, 1952, p. 30). Gunning was not convinced by the word list used in the Dale-Chall formula, believing that it implied that writers should limit themselves to a word list. He also rejected the idea of counting syllables, as required in Flesch's formula, as this is tedious for testing and distracts the 'writers' attention from the true cause of difficulty' (Gunning, 1952, p. 35).

The FOG index is similar to the Flesch formula, as both include average sentence length, but Gunning considers the percentage of polysyllabic words (words of more than three syllables) that require less time to count. This is the FOG index:

Grade level = 0.4 × (ASL + PPW)
Where ASL = average sentence length
PSW = percentage of polysyllabic words

The purposes of Gunning's index were to produce an easy formula to use due to the complexity of previous formulae and their difficulty for practical use and to provide a 'reliable measurement' for assessing readability (Gunning, 1952). The FOG index produces a grade score from 6 to 17, with 6 being understood by a sixth-grade pupil and 17 being understood by a college graduate. Gunning's polysyllabic word approach seems clever since it indirectly assesses specialised jargon. Longer words usually convey specialised meanings (Biber, 1988).

McLaughlin's SMOG Formula (1969)

On the assumption that the previous CRFs' primary goal was the simple and efficient prediction of reading difficulty, McLaughlin (1969) produced the 'SMOG' formula, which he claims is simpler, quicker and more valid than the previous CRFs. His formula focuses on word and sentence length, but these two measures are multiplied. Unlike the previous formulae, this involves a single variable (polysyllable count). The number of words of more than three syllables (polysyllable count) in 30 sentences was counted to obtain a grade level from 5 to 18.

Meri Coleman and T. L. Liau (1975)

Coleman and Liau (1975) argued that their formula enables the easy and accurate measurement of text readability. While previous CRFs involved counting the number of syllables per word, the accuracy of such counts could not be readily determined in computerised assessments (Coleman and Liau, 1975). The first variable in the Coleman-Liau formula is the number of letters per 100 words, and the second variable is the number of sentences per 100 words. The authors explain that 'there is no need to estimate syllables since word length in letters is a better predictor of readability than word length in syllables' (Coleman and Liau, 1975, p. 283).

Extensive reviews of other CRFs produced between the 1920s and 1970s can be found in Gray and Leary (1935), Flesch (1943, 1948), Chall (1974, 1988; 1996), Klare (1963; 1974; 1988), Gilliland (1972) and Harrison (1980). Most subsequent readability formulae are similar to those found in previous CRFs for the estimates of semantic/lexical and syntactic difficulty. Other advanced readability formulae have been produced for English since the 1980s. Stenner et al. (1988), for example, proposed a new way of measuring text readability with their 'lexile framework', which is not entirely different from CRFS and is based on two variables: the average sentence length and the frequency of word usage in the American Heritage Intermediate Corpus. The real breakthrough project in the readability research field since the 2000s has been Coh-Metrix. Coh-Metrix is discussed later in a separate section because it marked a huge departure from the previous CRFS towards a computational approach.

In conclusion, the most important development in readability research was the emergence of CRFs for assessing the writing-style difficulty or reading ease of a text. Readability researchers have attempted to objectively assess writing-style difficulty using formulae that consider word length and sentence length. However, since CRFs are based on two features of language and text comprehension, they are open to criticism.

3.8 Criticism of Readability Formulae

CRFs has been criticised as a method to assess content difficulty. Bruce et al. (1981) highlighted that CRFs does not incorporate current knowledge of reading theories or the reading process. Additionally, they do not consider 'other factors that make a particular text difficult, such as degree of discourse cohesion, number of inferences required, number of items to remember, complexity of ideas, rhetorical structure, dialect and background knowledge required' (Bruce et al., 1981, p. 50). Chall (1996, p. 28) rejects the claim of the atheoreticality of CRFs while conceding that:

> If classic readability is not grounded in modern cognitive psychology, it is a valid point. Classic readability came into being about 50 years before modern cognitive psychology. But classic readability has perhaps a longer theoretical and research base — that of the development of language and of reading comprehension.

Similarly, Kintsch and Vipond (1979) criticise the use of CRFs because it is not based on a model of reading comprehension, and the interaction between the reader and the text is ignored. However, Kintsch (1981) changed his opinion in an article written with another colleague, Miller, which concludes that CRFs 'indeed are correlated with the conceptual properties of text' and the two variables: long sentences are indicators of complex syntactic structures and infrequent words are also indicators of complex concepts; thus, both are best indicators of reading difficulty (Kintsch and Miller, 1981, p. 222). Furthermore, DuBay (2004) observes that, notwithstanding CRFs's limitations, many studies demonstrate that the surface features of a text – such as word length and sentence length – are indeed the best predictors of content difficulty.

Another weakness of CRFs is its lack of statistical basis (Bruce et al., 1981). Most CRFs have used the McCall-Crabbs standard test lessons in reading as a criterion for measuring comprehension. However, Bruce et al. argued that the McCall-Crabbs lessons were not designed as measures of text comprehension; they were intended only as practice exercises in

reading (1981). In response, Klare (1974, p. 75) states that 'these lessons have been convenient statistically because there are a large number of reading passages, resting upon extensive testing and providing detailed grade scores'. According to DuBay (2004), McCall-Crabbs lessons were an important criterion for developing and validating most readability formulas.

Other researchers have also criticised the use of CRFs. For example, Redish (2000) suggests that, as CRFs were developed with children's texts, they are not suitable for adult materials. Redish also states that while it is easy to measure word and sentence length, CRFs cannot measure other writing features, such as content and organisation. Some of Redish's criticisms are also identified as limitations of CRFs in Klare's list below, though some contradict the views of other researchers. For example, 'the record shows, however, that popular formulas such as the Flesch Reading Ease and the Kincaid formulas were developed mainly for adults and have been tested extensively on adult materials' (DuBay, 2004, p. 26). Furthermore, Chall (1996, p. 24) asserts that 'the classic approach to readability is still the most widely used for predicting comprehension difficulty of materials at the elementary, high school, college and adult level'.

DuBay (2004) observed that criticisms of CRFs tend to focus primarily on their limitations. However, as noted by Klare (1963), some researchers have criticised CRFs due to the inaccurate assumption that it was produced to measure all writing features. Klare (1963, pp. 24–5) highlights four particular limitations of CRFs:

- They measure only one aspect of writing-style and not other aspects such as organisation, word order, format or imagery.
- They do not measure the tone or 'dramatic effectiveness' of the writing.
- They do not measure 'difficulty perfectly'. This implies that formulae' scores might differ from one part to another part within the same document.
- They are not 'measures of good style'. This implies that they cannot identify whether a style is good for a particular document, but they measure style difficulty.

Many readability researchers (Zakaluk and Samuels, 1988; Binkley, 1988, 1988; Davison, 1988; Chall, 1996) have highlighted that CRFs are not designed for rewriting or as a writers' guide. Rather, they are designed to measure the difficulty of a written passage (Fry, 1988; Davison, 1984); thus, lowering a passage's difficulty level based on a CRFs measure does not always increase its comprehensibility (Klare, 1976). In her extensive review of readability research and the use and validity of readability measures, Chall (1974), however, concluded

that experimental studies have shown that simplifications based on formulae measures of vocabulary and sentence length significantly impact text comprehension.

In sum, CRFs are no more than a 'crude approximation' of content difficulty (Davison, 1984). CRFs only measures two types of writing-style difficulty: word difficulty and sentence complexity. The two measures of CRFs 'have validity as indexes of text difficulty. However, word length and sentence length alone explain only a part of text comprehension' (McNamara et al., 2011). Thus, CRFs are shallow measures, if not utilised together with other computational tools (such as Coh-Metrix) to provide a fuller picture of text readability. CRFs are widely used and generally accepted as accurate indicators of writing-style difficulty (see DuBay, 2004; 2007; Klare, 1963; 1976; 1974; Chall, 1996; 1974). CRFs was deemed convenient for the current research for the following reasons:

- To the researcher's knowledge, CRFs have not previously been used in any study to predict differences in the reading ease of English Qur'ān translations.
- CRFs are essential for exploring differences in reading ease in Qur'ān translations from different periods.
- The reading ease scores given by CRFs are beneficial for determining their prediction with comprehension scores made by human ratings.

CRFs is suitable when software such as Coh-Metrix and other research methods (such as human ratings for text comprehension) are combined to provide accurate measures of text readability and comprehensibility. CRFs are essential for preliminary investigations of content difficulty, but they are not sufficiently explanatory regarding which variables increase or reduce text difficulty. The best explanatory account of content difficulty is provided by individual indices or variables, such as the indices of Coh-Metrix, which attempt to overcome the criticisms outlined above and detect factors that CRFs cannot detect.

3.9 Coh-Metrix

> Coh-Metrix is a major departure from both the classic formulas and cloze tests. It is neither a formula nor a procedure, but rather a computational tool that facilitates the formulation and testing of hypotheses about readability and other reading comprehension issues (Bailin and Grafstein, 2016, p. 48).

Until the end of the 20th century – and before the publication of Coh-Metrix – there was no computer programme to measure other content features, such as cohesion, to provide accurate measures of text difficulty. Coh-Metrix can 'analyses texts on over 200 measures of cohesion, language, and readability. Its modules use lexicons, part of speech classifiers, syntactic parsers, templates, corpora, latent semantic analysis, and other components that are widely used in computational linguistics' (Graesser et al., 2004, p. 193). Coh-Metrix was developed and tested between 2002 and 2011 at the University of Memphis, and the first version was released for use in 2003 (McNamara et al., 2014). There are two versions of the computational software: (1) the Web Tool of Coh-Metrix (Coh-Metrix) version 3.0 and (2) the Coh-Metrix-TEA (Text Easability Assessor).

The Coh-Metrix TEA[5] has five key measures of 'text easability': narrativity, deep cohesion, referential cohesion, syntactic simplicity and word concreteness. This version is useful for new users and classroom purposes. It is simple to understand and free for public use. It allows educators to submit a short text (less than 1,000 words) and obtain a readability profile of the text quickly (Dowell et al., 2016). The process is easy to use (copy paste and analyse) and 'provides immediately interpretable results through an informative visual illustration and short explanation' (Dowell et al., 2016, p. 75–6).

Furthermore, the Web Tool of Coh-Metrix[6] is an extended and full version of Co-Metrix, with 108 measures. Coh-Metrix was designed to analyse large volumes of texts of up to 15,000 words each. These 108 measures cover all text difficulty levels, which might be overwhelming for novice users (Dowell et al., 2016). However, they are essential for scaling content difficulty on several levels of language and discourse (McNamara et al., 2014). The measures of Coh-Metrix are aligned with theories of text and discourse comprehension, which assume that the comprehension process of a text operates at multiple language and discourse levels (McNamara et al., 2014).

All previous CRFs focused on the surface characteristics of the text and are not intended to assess the deeper features of language and discourse. Coh-Metrix provides a global picture of the potential challenges readers might encounter in comprehending texts. It provides multiple indices (or variables) of language or text characteristics on the levels of words, sentences and discourse. Coh-Metrix uses variables from other fields and combines variables from computational tools to compute various linguistic, semantic and psycholinguistic features that accurately estimate the degree of content difficulty. For example, it uses the Medical Research

[5] The Coh-Metrix TEA is available online (http://tea.cohmetrix.com).

[6] The web tool of Coh-Metrix is available online (http://141.225.61.35/cohmetrix2017).

Council Psycholinguistic Database for computing psychological features of words (such as meaningfulness, concreteness and imageability) and WordNet for computing semantic features of words (such as hypernymy and polysemy) (McNamara et al., 2014). Coh-Metrix uses the Apple Pie parser to compute measures other than sentence length, such as the number of modifiers in noun phrases, the number of words before the main verb and the density of syntactic patterns (such as verb phrases, noun phrases and adverbal phrases) (McNamara et al., 2014).

No classic readability formulae provide indices of text cohesion, although sentence length serves as a standard proxy for cohesion (Fry, 1988; McNamara et al., 2014). Coh-Metrix computes three sources of text cohesion: referential cohesion, semantic coherence and connectives (see Chapter 4). Referential cohesion refers to the degree of word overlap or repetition across sentences and paragraphs (McNamara et al., 2014). Semantic coherence is based on Latent Semantic Analysis, which detects the semantic similarities between words, sentences and paragraphs. Coh-Metrix also provides several indices of connectives: causal connectives (*because, so*), logical connectives (*and, or*), adversative/contrastive connectives (*although, whereas*), temporal connectives (*first, until*) and additive connectives (*and, moreover*). These connectives draw explicit connections between sentences and events (McNamara et al., 2014). In sum, these three sources of text cohesion can provide more accurate estimates of text readability than traditional measures.

Some studies have used Coh-Metrix to explore linguistic features in various genres and to examine the content difficulty of different texts. For example, Louwerse et al. (2004) used some measures from Coh-Metrix to distinguish between spoken and written registers in English. McCarthy et al. (2006) used Coh-Metrix to detect shifts in writing styles across several authors. Lightman et al. (2007) scrutinised the level of cohesion ties and text difficulty in three genres. However, some recent studies have shown an interest in using Coh-Metrix to explore the features of translated texts. A study by Li et al. (2015) used measures of formality and cohesion from Co-Metrix to compare Google translations and human translations of Chinese into English. The study found that, on measures of formality and cohesion, Google English translations were close in quality to both the English texts translated by a human being and the original Chinese texts.

As far as the current researcher knows, this software has never been used in a study to explore the effect of retranslation in changing text readability and comprehension, particularly in English translations of the Qur'ān. The level of linguistic complexity and readability of these

retranslated versions of the Qur'ān has not been previously explored using Co-Metrix. This software was chosen for this research for the following reasons:

- It can provide an objective assessment of the readability of these retranslated versions of the Qur'ān.
- It includes a diverse range of measures and has the best, most widely tested and most commonly used predictors of text difficulty.
- It enables this research to correlate and compare variables that might increase or reduce text comprehension.
- It makes it possible to test whether these retranslated versions of the Qur'ān are similar regarding linguistic complexity.
- It allows for cross-validation between the results of the automatic assessment (see Chapter 5) and human assessments or ratings (see Chapter 6), allowing for the identification of convergence or discrepancies between the two sets of results.

3.10 The Text-Reader Model of Text Readability and Comprehensibility

Some of the early works in readability research were only interested in certain text characteristics, such as style variables, to predict text readability. Nevertheless, these works inspired subsequent studies to address other factors that could not be detected by CRFs but could be influencing text difficulty. For Gray and Leary (1935), the four primary contributors to text readability are content, style of expression and presentation, format and general organisation factors. This proposition is unique in the field of readability research, as it considers several text factors.

Other researchers have linked readability to both text and reader factors, marking the beginning of a 'text-reader model of readability' (Singer, 1988). This kind of approach has expanded since the 1980s, as researchers have brought in views from cognitive psychology and linguistics studies to propose a more complex picture for determining text readability (see Miller and Kintsch, 1980; Zakaluk and Samuels, 1988). One of the important studies combining text and reader factors was that of Harrison in 1980. Harrison (1980) divided the factors that would make schoolbooks easy or difficult to understand for children into two groups. The first group is related to reader factors (such as reading ability, interest and motivation), and the second group is related to text factors such as 'legibility of print', 'illustration and colour', 'vocabulary', 'conceptual difficulty', 'syntax' and 'organisation'.

Zakaluk and Samuels (1988, p. 140), proposing a new approach to predicting text comprehensibility, opined that 'text factors alone cannot determine readability; readers' prior knowledge and understanding influence comprehensibility and recall'. They argue that text comprehensibility need not be viewed from the perspective of text factors alone, but can also be viewed from text-reader interaction. The approach of Zakaluk and Samuels to analysing readability and comprehensibility has a psychological orientation, as the process of reading comprehension requires an interaction between the reader and the text to form a coherent presentation of the text (Miller and Kintsch, 1980; Kintsch and van Dijk, 1978).

Singer (1988) argues, however, that the approach of Zakaluk and Samuels is time-consuming, as it measures both text and reader factors. Figure 3.1 provides a visual description of the text-reader model of text readability and comprehensibility. It seems that text comprehension difficulties might arise from either the reader or the text itself. The factors, whether on the text side or the reader side, influencing text readability and comprehensibility are discussed in the following sections.

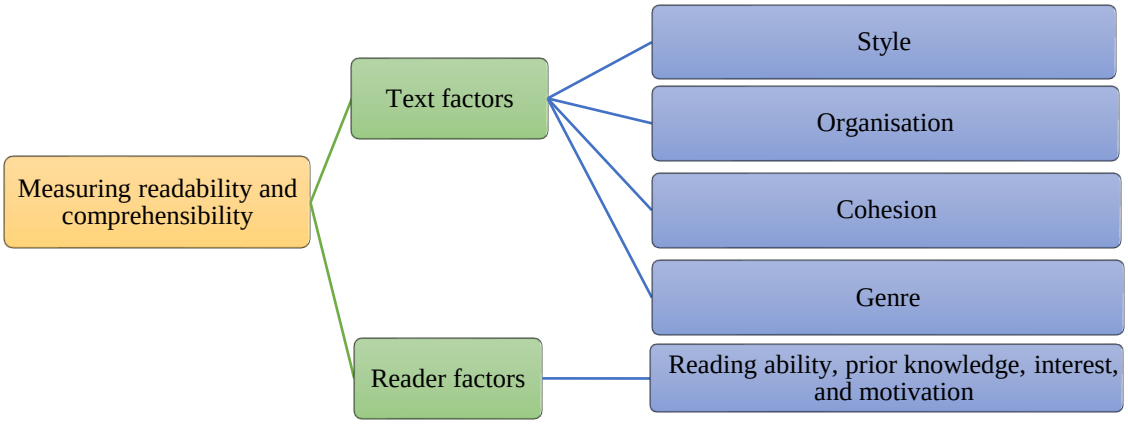


Figure 3.1: The Text-Reader Model of Text Readability and Comprehensibility

3.10.1 Reader Factors

As previously mentioned, the text and reader are the main factors influencing the comprehensibility of a text, as the reading process requires a text-reader interaction. Klare (1980; 1988) produced a model illustrating the performance of a reader during the reading process. According to Klare's model, a reader's performance depends on the level of his/her competence, motivation, prior knowledge and readability of writing.

Klare (1976) published an article titled *A Second Look at the Validity of Readability Formulas*, which analytically examines 36 studies on the influence of 'readability variables upon reader comprehension' (1976). One of these studies was by Denbow (1973, cited in Klare, 1976, p. 141), who examined two versions of a text with different levels of reading difficulty: 'one higher in measured interest-value and one lower'. Denbow found that the version with high readability was higher in information gain than the version with low readability, while the information gain from the version with higher readability was more significant for the topic with a lower interest value. This indicates that higher readability might benefit those with lower interest values.

Klare also discusses a study conducted by McLaughlin (1966) that compared two versions of a text (one easier and one harder) and tested the effect of motivation on the two groups. McLaughlin found that the difference in comprehension scores for the easier and more difficult versions was insignificant when the subject was highly motivated, but they were significant when the subject was not highly motivated, inducing his conclusion that the easier version was preferred. Klare (1976, p. 141) notes that McLaughlin's result might appear contradictory, but the readability of a text is 'less important when motivation is high than when it is low, i.e., with less interesting material'. If a text is linguistically complex, readers with low interest or limited motivation might struggle to understand it properly. According to Chebat et al. (2003, p. 601), 'some empirical studies on linguistic complexity support arguments favouring high readability'.

Another critical factor when measuring a text's readability is the degree of prior knowledge of the subject. Several researchers have explained the use of prior knowledge on text comprehension within 'schema theory', which posits that a reader's 'schema' – their background knowledge and shared experience of the world – influences how they comprehend, learn and recall information from the text (Armbruster, 1984; Zakaluk and Samuels, 1988; McNamara et al., 2014). Research findings have shown that what readers bring to a text can influence their understanding, and readers with more prior knowledge might recall faster and understand better than readers with less or no prior knowledge (Zakaluk and Samuels, 1988; Klare, 1988; Binkley, 1988; Moravcsik and Kintsch 1993). According to Klare (1976; 1988), many studies show that when prior knowledge of the subject is high, the effect of readability is low, whereas when prior knowledge is low, the effect of readability is high. Entin and Klare (1985) examined the impact of different readable versions on reader's interest and prior knowledge of the subject. The authors found that texts with high readability were more

beneficial for those with limited knowledge and low interest in the subject than for those with high knowledge and high interest.

The current research will consider the importance of reader's prior knowledge, alongside text factors, in estimating the readability of Qur'ān translations. Thus, this research employs two groups to rate the comprehension ease of the chosen Qur'ān translations. These groups are (1) readers with high or moderate levels of knowledge of the Qur'ān and (2) readers with little or no knowledge of the Qur'ān (for more details of these participants, see Chapter 4). This comparison between the two groups will allow us to determine whether the factors that influence text comprehension are due to some readers' backgrounds or some text factors found in some Qur'ān translations. The text factors that affect the comprehensibility and readability of a text are explored in the next section.

3.10.2 Text Factors

> What the reader brings to the text, however, is not the only factor that affects learning from text. Characteristics of the text itself, [influence] the reader's ability to construct a coherent model of the text's meaning.
>
> (Armbruster, 1984, p. 203)

Since the 1920s, readability research has generated a considerable body of literature related to examining text factors that influence the comprehensibility of a text. Some of the analysis in these early readability studies was based on shallow traditional measures (e.g., word length, sentence length), ignoring other factors that also influence readability and comprehension, such as cohesion and organisation (Klare, 1988). The study by Gray and Leary in 1935 was the first to investigate text factors other than word length and sentence length regarding their contributions to text readability and comprehension.

This research attempts to summarise the most important text factors found in readability and comprehension studies (see Gray and Leary, 1935; Harrison, 1980; 1986; Klare, 1988; 1976; 1963; Danielson, 1987; Fry, 1988; DuBay, 2004; McNamara et al., 2014), which can affect and assess text comprehension and readability. These text factors fall under four broad categories:

a) Style
    1. Word Difficulty
    2. Syntactic Complexity
b) Organisation
c) Cohesion and Coherence

d) Genre

We will consider each of these in turn.

3.10.2.1 Style

Style refers to the stylistic elements of a writing, such as word and sentence choices, that make a text easy or difficult to read and understand (Gray and Leary, 1935; Klare, 1963). Style has been the most important factor determining text readability since the early studies of readability and CRFs. Chall (1956, p. 95) notes that 'elements of style, so far, have been measured reliably and have been included in readability formulas'. The study by Gray and Leary in 1935 surveying readers' opinions supports the view that the 'style of expression presentation of material' plays a significant role in whether a reader finds a text readable or not.

Moravcsik and Kintsch (1993) argued that stylistic elements of writing can contribute to text comprehension. They examined the effect of different styles of writing on readers' comprehension and concluded that 'good writing cannot entirely compensate for a comprehender's lack of domain knowledge. It has strong positive effects, just as domain knowledge does … Good writing helps comprehenders form a coherent textbase' (1993, p. 371). Four stylistic elements are reliable and significantly related to content difficulty: 'vocabulary load', 'sentence structure', 'idea density' and 'human interest' (Chall, 1956, 1974). Chall cites vocabulary load as the most significant of the four factors for making a text readable.

The difficulty or ease of a writing style in Qur'ān translations can be reliably measured by Coh-Metrix or CRFs. Several style variables indicative of writing style difficulty are classified in this research in terms of 'word difficulty' and 'syntactic complexity'.

- ***Word Difficulty***

Word difficulty has been a basic measure in readability research since the 1920s. It has been used as a robust predictor of text difficulty and is the most frequently studied variable (Klare, 1963; 1976; Gilliland, 1972; Chall, 1974; Harrison, 1980). Many methods have been used to measure word difficulty in texts. One of the most common and foundational methods in early studies was to measure 'vocabulary frequency' using a word list (see Thorndike, 1921). Such word lists have become less common with technological advances that allow researchers to measure vocabulary frequency in the British National Corpus or CELEX, for example. CELEX, a lexical database of nearly 17.9 million words, was designed by Baayen et al. (1995) to provide 'estimates of how frequently English words are used in a very large corpus of documents' (cited in McNamara et al., 2014, p. 40). CELEX has been incorporated into Coh-Metrix. Word length is another central measure of readability for gauging word difficulty. It is

measured in CRFs by counting the average number of syllables, polysyllables, monosyllables, hard words and so on (see Chapter 4). The underlying assumption of the previous two methods is that 'the more rare, uncommon, or long words, the harder it is for the readers to grasp the meaning' of a text (Chall, 1956, p. 95).

Unlike these traditional style variables, such as word length and word frequency, there are several methods to detect various elements of word difficulty reliably. Research on comprehension and readability has demonstrated that psycholinguistic lexical features (such as meaningfulness, concreteness, imagability) and semantic lexical features (such as hypernymy) can contribute to making a text easy or difficult to read and understand (Paivio et al., 1968; Paivio, 1969; Richardson, 1975; Klare 1975, Harrsion, 1980; McNamara et al., 2014). These psycholinguistic variables in a text usually make words more vivid, more interesting and more engaging because imageable words can evoke a mental image, whereas concrete words can make the reader see, hear, feel and smell (McNamara et al., 2014).

Several psychological studies have shown that highly imageable, concrete and meaningful words are faster to understand and remember than abstract words (Paivio, 1969; Paivio et al., 1968; and Richardson, 1975; Schwanenflugel and Stowe, 1989). Hypernymy is another measure of word difficulty, which explains words with more or less specific meanings; in fact, hypernym is an indirect measure of abstractness (McNamara et al., 2014). Thompson (2014, p. 102) notes that these psycholinguistic variables can capture cognitive parts of reading difficulty 'not directly addressed by the surface vocabulary and syntax features of traditional formulas'. These semantic and psycholinguistic variables have been adopted in Coh-Metrix, making it possible to measure multiple aspects of word characteristics that might cause reading difficulty.

- ***Syntactic Complexity***

There are several aspects of syntactic complexity that appear to be associated with reading difficulty. One of the most common is sentence length. As seen previously, in early studies of readability, sentence length is frequently used as a robust indicator of syntactic complexity in most readability formulae (see Davison, 1984; Rabin, 1988; Chall, 1974; 1996; McNamara et al., 2014). Long sentences make a text less comprehensible (Fry, 1988) because they put higher demands on working memory (Millis et al., 1993; McNamara et al., 2014). This is not to say that no sentence should be long, but 'on the average [sic] sentences should be shorter for better communication' (Fry, 1988, p. 81).

Anderson and Davison (1988) have argued that texts with short sentences are understood alongside long sentences because long sentences usually contain markers of connection, which

provide more explicit clues to the meaning of a sentence. Other researchers have clarified that average sentence length is positively correlated with complex syntactic structure – such as the use of phrase density (i.e., noun, verb and prepositional phrases), subordination and connectives (Harrison, 1980; Fry, 1988; Anderson and Davison 1988; Millis et al., 1993; McNamara et al., 2014). Research findings have found that sentence length could explain a small part of text comprehension (Anderson and Davison, 1988; McNamara et al., 2011). Other syntactic features have also been used to indicate complexity and ambiguity, which cannot be measured by classic formulae, such as 'the degree to which sentences have embedded constituents, and [with] the load on working memory. Working memory is taxed when there are noun-phrases with many modifiers and when many words must be held in working memory before the reader receives the main verb of the main clause' (McNamara et al., 2014, p. 79). All these variables of syntactic features have been incorporated into Coh-Metrix.

3.10.2.2 Organisation

These aspects of writing style are not the only factor making a text readable, but a text's difficulty also depends on its degree of organisation. 'Organisation' refers to format variables (such as paragraphs, subheadings and typographies) that may add to text comprehensibility (Gray and Leary, 1935; Harrison, 1980; 1986). Studies on text organisation have shown the effects of summaries in a text on recall and the effect of headings on recall, search and retrieval (Hartley and Trueman, 1982; 1983).

The most prominent study in readability literature is that of Gary and Leary (1935), who examined the opinions of readers regarding what factors make text readable for readers with limited reading ability. The researchers found that format variables (book size, colour of print, illustrations, graphs, font size and style and chapter headings) were rated third in importance, while content and style were rated of higher importance. A study found that readers understood a paragraphed passage better than an unparagraphed one that was supposed to be the original organisation of the text (Harrison, 1980). Features of format such as 'good paragraphing, introductory phrases in bold type, and short chapters' support 'the reader's effort to grasp the author's thoughts' (Chall, 1974, p. 7).

Another way of explaining some aspects of text organisation is for the use of 'signalling' (Meyer, 1979, cited in Armbruster, 1984), 'advance organisers' (Ausubel, 1960, cited in Harrison, 1980) and 'adjunct comprehension aids' (Zakaluk and Samuels, 1988). These coined terms, which take a learning perspective, are often used to describe aspects of text organisation

that help readers understand the content and highlight information in the text. Armbruster notes (1984, p. 204) the following types of signalling used by writers:

(1) Explicit statements of the structure or organisation;
(2) Previews or introductory statements, including titles;
(3) Summary statements;
(4) Pointer words and phrases such as "an important point is…"
(5) Textual cues such as underlining, italics and boldface.

A study found that students remember more from texts that use signalling devices (Meyer, 1979, cited in Armbruster, 1984). In a recent study, it has been shown that global cohesion (i.e., adding headlines, connections between paragraphs and summarising sentences) can contribute to reading comprehension in expository texts, but not in literary ones (Schmitz et al., 2017). Zakaluk and Samuels (1988, p. 148) emphasise the influence of 'adjunct comprehension aids' on text comprehensibility:

> In addition to text readability level, another outside-the-head factor that influences comprehension is adjunct aids, such as statements of objectives or study questions located within the text or at the beginning or the end. These adjunct comprehension aids highlight important information and increase the depth of text processing. Their presence in the material adds to text comprehensibility.

The text organisation that a translator includes in a target text might impact the understanding of the translation. Regarding the content organisation of English Qur'ān translations, some translators have variously employed different text layouts and with 'adjunct comprehension aids', such as introductory statements on the meaning of the Qur'ānic chapter's titles, introductions to the background and general structure of Qur'ānic chapters and footnotes and commentaries on the meanings of certain passages. In contrast, some translators have not used any aspect of comprehension aids. This research aims to empirically test the influence of different Qur'ānic text layouts and comprehension aids that have been employed in different Qur'ān versions on text understanding. One of the limitations of all classic formulae is that they do not measure the effects of organisation, comprehension aids or even cohesion on text understanding.

3.10.2.3 Cohesion and Coherence

In 1935, Gray and Leary introduced the term ‘content’ into readability research, referring to the ‘nature of the subject matter’ and the ‘unity of content’. They found that content was the first-most crucial factor of text difficulty. Klare (1988) noted that studies started to focus on other factors influencing text readability, such as cohesion, coherence and organisation, in the 1980s.

The major inputs that hang a text together to make it comprehensible are cohesion and coherence. ‘Cohesion’ refers to the surface relations of linguistic features in a text (Graesser et al., 2004; McNamara et al., 2014). It generally links sentences, paragraphs and ideas via lexical and grammatical devices. ‘Coherence’ refers to the interplay between the text and the reader’s attributes (e.g., world knowledge, domain knowledge, and reading ability), thus creating a mental textual world that readers might build from the text (Graesser et al., 2004; McNamara et al., 2014). A non-cohesive text might make readability difficult. Concurrently, if a text is cohesive, it is not necessarily coherent for all readers because of their knowledge.

However, cohesion might help readers organise the comprehension process of a text. Many studies on cohesion in reading comprehension have shown that cohesive ties facilitate reading comprehension for many readers (Britton and Gulgoz, 1991, McNamara 2001; Zwaan and Radvanksy, 1998), as they ‘help readers construct more coherent mental representations of text content’ (McNamara et al., 2012, p. 6). However, cohesion can affect reading comprehension for readers with different background knowledge. Research findings have revealed the effects of cohesive ties on those with low prior knowledge of the topic, who are thus likely to benefit from high-cohesion texts, and on those with high prior knowledge, who can benefit from low-cohesion texts (McNamara and Kintsch, 1996; McNamara et al., 1996; Kamalski et al., 2008). This is because a text with high cohesion has fewer conceptual gaps, which reduces the reader’s cognitive load to generate interference to link the parts of the text (Lightman et al., 2007; McNamara et al., 2014).

Altogether, these studies suggest that cohesion ties can make a large contribution to readability. A study by Binkley (1988) demonstrated the importance of cohesive ties on readers’ understanding, learning and recall. She discussed the role of cohesive analysis in assessing readability and explaining the patterns of ties that often occur in specific texts. Linguistic pattern or register in specific texts is considered a distinct factor, which is discussed in the genre section below.

3.10.2.4 Genre

‘Genre’ refers to the type or category of a text (Biber, 1988). According to McNamara et al. (2014), the genre of a text can be narrative, expository, persuasive or descriptive. Some texts are not ‘pure’ regarding the text genre (McNamara et al., 2014) – for example, the Qur’ān is a mixture of narrative, legal and instructional passages (Abdul-Raof, 2019).

These categories of text have different linguistic features and discourse registers. For example, science writings usually repeat the same nouns throughout, whereas social science writings usually depend on synonyms (Binkley, 1988). According to McNamara et al. (2014, p. 100), narrative writings usually include people, places, events and oral language that are ‘substantially easier to read, comprehend and recall than are other genres of text such as science, history and other expository domains’. Binkley (1988) highlights that differences in argumentation style can influence the comprehensibility of a text and might have different implications for readers.

Lightman et al. (2007) scrutinised the distribution of cohesion ties and text difficulty across three genres (narrative, science and history). They found that these three genres showed different levels of text readability. In their findings, narrative texts were easier than history and science texts because they contained more familiar words and simple sentences. The authors also observed that science texts were more cohesive than narrative and history texts. The cohesion in science texts is required to support the reader as they encounter more terminology (McNamara et al., 2014). Lightman et al. (2007) suggest that differences in genres present comprehension obstacles to their readership.

Research findings have shown that expectations regarding genre can affect reading comprehension (Zwaan, 1994; Schmitz et al., 2017). In Zwaan’s experiment, when readers read the text with literary expectations, they had longer reading times than when reading the text with newspaper expectations. Readers showed better memory of surface features of language in literary texts but poor memory of situational information in newspapers (Zwaan, 1994). Similarly, Schmitz et al. (2017) suggest that readers take different approaches to reading processing when they interact with the challenge and comprehension of different genres.

Overall, the source of reading difficulty might arise from the linguistic characteristics of specific texts and registers. As Biber (1992 p. 133) states, although many studies take a comprehension perspective and address the issue of ‘what linguistic characteristics of a text make it more or less understandable or readable’, there are also two major aspects on this issue: ‘complexity differences in the language produced by different social groups, and complexity differences in the language produced in different situation’. Studies analysing the linguistic

characteristics of English Qur'an translations are lacking. Thus, a multidimensional approach[7] to textual variation introduced by Biber (1992; 1988) is needed to determine the sources of text difficulty in respect of register variation in English Qur'an translations.

3.11 Chapter Summary

This chapter's focus was on the concepts of readability and comprehensibility. The relationship between them was explained, and the two concepts were defined. The chapter also reviewed the development of CRFs, and identified the text variables that might influence text comprehension, and discussed software that could provide more information than the CRFs. The chapter has shown that evaluating text comprehension and readability is challenging because it is intertwined with two critical factors: the influence of reader characteristics on text comprehension and the influence of text properties on reader comprehension. Readability is primarily examined from the text side, and comprehension is influenced by both the reader and the text side. Despite CRFs' limitations, they nonetheless provide a good approximation of content difficulty. Along with the individual indices from Coh-Metrix, this CRFs – and human ratings of the readability and comprehensibility of the chosen Qur'ānic samples – will provide an accurate global picture of content difficulty in Qur'ān translations. All measures from Coh-Metrix and the CRFs used in the current research – and in conjunction with other methods such as human ratings of readability and comprehensibility – are described further in the next chapter.

[7] This is a statistical technique applied through a factor analysis that allows the researcher to describe the range of linguistic variation in specific texts or registers (see Chapter 4).

Chapter 4: Research Methodology

4.1 Chapter Overview

This chapter reviews the philosophical stance adopted in this research. The chapter is divided into three main parts. The first begins by describing the difference between multimethod and mixed-method research and the reasons for choosing a multimethod approach in this research. In this first section, other issues of validity and reliability and the three primary sources for collecting data are also discussed. The second part briefly covers Schriver's (1989) three categories of text-quality evaluation: (1) text-focused, (2) expert judgement-focused, and (3) reader-focused approaches. These three categories are helpful and relevant to the three quantitative approaches used in this research for measuring comprehension and text difficulties, namely: the classic approach, the computational approach, and the human judgement approach, which are intended to allow an analysis of multiple perspectives. The classic approach involves five readability formulae for predicting comprehension difficulty. The computational approach, employing Coh-Metrix, which is the primary approach in this research, presents multiple variables that can measure comprehension difficulties in texts from various angles. The three corpora used for Coh-Metrix analysis and the manner in which the Qur'ānic corpora have been cleaned to obtain consistent scores are detailed in the Coh-Metrix section. The human judgement approach involves the survey design, rating scale, pilot study, sampling, survey administration, survey type, and ethical considerations. The last part of this chapter introduces a different set of statistical procedures and tests for data analysis. An overall chapter summary is given at the end before moving on to the data analysis in chapter 5.

4.2 Multimethod Research

The term ‘multimethod’ itself is used synonymously with the term ‘mixed methods’ in some studies, which creates confusion, although there are researchers who make a clear distinction between the two terms. For example, Anguera et al. (2018) discuss this issue, stating that some authors see no difference between the two terms, whereas others treat the term ‘multimethod’, in a general sense, as meaning different methods or styles of research involving either multiple methods of quantitative data, multiple methods of qualitative data, or multiple methods of both qualitative and quantitative data (i.e., mixed methods).

The term ‘multimethod’ denotes a distinctive style of research that is clearly different from that of mixed methods, according to some authors. Creswell (2015, p. 3), who distinguishes between multimethod and mixed methods, argues that ‘when multiple forms of qualitative data (or multiple forms of quantitative data) are collected, the term is multimethod’. Other authors adopt a similar view in distinguishing between the two terms. Thus, the research design for a multimethod study can be defined as in Hesse-Biber et al.’s (2015, p. 6) words:

> In a multimethod research design a qualitatively driven project may call on a second qualitative method as its auxiliary component: the second qualitative method would take on a secondary role (qual) in the service of a primary QUAL method. The addition of a second qualitative method would serve a supplementary function in that it answers a different question, but its primary aim is to support the core qualitatively driven approach and question. This qualitatively driven design would be called a multimethod design by its use of two different qualitative methods.

Anguera et al. (2018) point out that the definition above is ‘ingenious’ in interpretating the process of multimethod research. Hesse-Biber et al.’s (2015) explanation, quoted above, is relevant to the methodological design of this research, which is explained in the following section.

Hunter and Brewer (2015) view multimethod research as a combination of two or more different methods — which may be purely qualitative, quantitative, or a combination of both qualitative and quantitative. They argue that the epistemology of multimethod research is rooted in pragmatism. The pragmatic approach allows the researcher to employ the most relevant and practical tools to address the questions the research aims to answer.

Mark and Shotland (1987, cited in Mark, 2015) explained three purposes of multimethod rather than mixed-methods research, including triangulation, bracketing, and complementary goals. Triangulation acts to enhance the validity of the research; the results of study through triangulation might not converge, but might be consistent or even contradictory from one method to another (Preissle et al., 2015). Mark and Shotland's (1987) view is that the results of different methods, associated with shortcomings inherent in any type of method, can inform alternative estimates or answers on the same topic within the research. The last benefit of multimethod research identified by Mark and Shotland (1987) is that it is complementary. This means that a number of different methods are used to handle other functions of the research to clarify and enhance findings derived from one type of method to another one in the same study.

4.2.1 Multimethod Design

The primary purpose of this section is to encapsulate how and why multimethod research has been chosen as the design for this research. This study adopts multimethod research to examine the readability and comprehension of the three English Qur'ān translations. It employs three quantitative approaches to the measurement of text readability, comprehension, and cohesion, which are: the classic approach (i.e., classic readability formulae [CRFs]), the computational approach (i.e., automated evaluation of text and discourse with Coh-Metrix) and the human judgement approach (i.e., human ratings for text readability and comprehension). The rationale behind using such a design is to overcome the limitations of each method, and so that they can serve both supplementary and complementary purposes, allowing us to interpret the questions of this research from differing perspectives. According to Brewer and Hunter (2006), some research questions and problems might be better handled when research combines either two forms of quantitative method or two forms of qualitative method in the same study.

Based on our extensive literature review in Chapter 3, reader characteristics and text properties in readability research are the two leading indicators for estimating text and comprehension difficulty. The classic approach predicts comprehension difficulty based on the surface characteristics (i.e., word length and sentence length) of the text. CRFs are not developed to measure most of the characteristics of texts which cause comprehension difficulties. The classic approach provides preliminary analysis in this research to predict the ease with which certain text can be read. Thus, the computational approach employing Coh-Metrix plays a primary role in this research, allowing us to measure 'text difficulty at various levels of language, discourse, and conceptual analysis. Text difficulties include problems with vocabulary, syntactic composition, meaning, and cohesion' (McNamara et al., 2004, p. 5). This

software is relevant for the purposes of this research because the indices or measures which it uses that are compatible with theories of text and discourse comprehension, have been validated in thousands of comprehension and discourse studies (see McNamara et al., 2014). The classic and computational approaches are used to measure and describe comprehension difficulties more from the perspective of the text itself.

The human judgement approach plays a secondary role in this research; it allows us to ascertain the human reaction to those text difficulties identified by the previous approaches, in addition to providing human ratings for readability and ease of comprehension. This human feedback is obtained through an online self-completion survey, aiming to recruit participants with different levels of prior knowledge of the Qur'ān, to rate the level of readability and ease of comprehension of translations of three sections of the Qur'ān. The ratings of those with different levels of prior knowledge are likely to offer a more complete understanding of the nature of any comprehension difficulties from the perspective of the reader and in relation to the text, allowing us to identify which translators of the Qur'ān produce an easy or difficult version for English readers who either have prior knowledge or no knowledge of the Qur'ān. This aspect of the research, assessing human judgement may be complementary or confirmatory in terms of the findings of each approach used in this research. Each of these three approaches is discussed in the following sections.

4.2.2 Reliability and Validity

One of the critical issues in designing research is to ensure that the tools used are valid and reliable for addressing the research questions. According to Frey et al. (2000), reliability is generally related to the consistency and accuracy of a tool, while validity refers to the appropriateness of a tool – whether it measures what it is intended to assess. Classic readability formulae are widely recognised among scholars as reliable and valid measures for estimating certain aspects of comprehension difficulty (see Kintsch and Miller, 1981; Chall, 1996; Klare, 1988; 1976; DuBay, 2007; McNamara et al., 2011; 2014). The indices provided by the computational approach of Coh-Metrix are also reliable and valid in measuring multiple aspects of comprehension difficulties in texts, and have been validated in a considerable number of studies (see McNamara et al., 2014).

The scores for these two approaches are highly accurate when they are repeated following on from cleaning the data in an identical manner for each approach. The reliability of a tool (such as the survey used in this research) can be tested through the internal consistency of a scale and a pilot study (Saldanha and O'Brien, 2014; Frey et al., 2000). According to Frey et

al. (2000), when the number of errors is reduced after a pilot study has been conducted, the reliability of a tool is likely to increase. A pilot study was carried out on the survey used in this research, and the internal consistency of each scale used to measure each construct was examined, as reported in Section 4.6.4.

Frey et al. (2000) point out several threats to the validity in research findings, associated with both internal and external validity. Threats to internal validity can include factors such as how the research has been conducted, how the research participants have been selected, and the researcher's influence on the research (Frey et al., 2000). The threats to external validity, which generally relate to the generalisability of research findings, can be due to sampling, ecological validity, and replication (Frey et al., 2000). The following sections explain how this research was conducted, along with all the variables used in each approach and how the text samples were selected and computationally analysed, allowing for an increase in the replication of our findings, particularly in the text automated analysis. However, replication is not easy to achieve in a participant study (Saldanha and O'Brien, 2014). The selected participants, detailed in Section 4.6.8, were not known to the researcher, nor did they have any relationship to the researcher (Frey et al., 2000). Since the surveys were online self-completion surveys, the influence and interaction with people in the responses was low. Online surveys are intended to test people in a natural setting, allowing them to read and easily assess translations in their free time. Lastly, the Qur'ān translations were selected as the primary sources of the data collection based on their significance and popularity from among translations of the Qur'ān and among English readers.

4.2.3 Qur'ānic-Samples Sources

All the Qur'ānic chapter samples that are described in Section 4.5.5.1 were taken from three reputable translations of the Qur'ān, as listed below.

- *The Holy Qur-ān: English Translation of the Meanings and Commentary* by Abdullah Yusuf Ali (1934, revised edition 1989).
- *The Koran Interpreted* by Arthur John Arberry (1955, revised edition 1983).
- *The Qur'ān: A New Translation* by Muhammad A. S. Abdel Haleem (2004).

These translations are faithful to the original Arabic text, and were done by two Muslims and one non-Muslim. As discussed in Chapter 2, these translations are considered retranslations following the first publication of a Qur'ān translation. This research treats these retranslations

according to this classification: (1) Abdel-Haleem's work, published in the 21st century, is a recent retranslation; (2) the other two translations, published in the 20th century are initial translations. This research selected two widely used versions from the 20th century and another recognisable one from the 21st century to reflect textual changes in text readability and comprehensibility.

4.3 Approaches to Evaluating Text and Comprehension Difficulties

The approaches for assessing text and comprehension difficulties fall under Schriver's (1989) three general classes of tests to evaluate text quality. Schriver classifies methods for assessing text quality into the following categories: (1) text-focused, (2) expert judgement-focused, and (3) reader-focused approaches, which are very relevant to the methods used in this research.

The first category, text-focused methods, scrutinises a text's properties and does not directly examine readers' responses (Schriver, 1989). The methods of this category yield either manual-based guidance (which can be produced using checklists and guidelines for evaluating the quality of a text) or automatic-based scores (which can be produced using readability formulae or other computer-based text analysis software such as Coh-Metrix), both of which help researchers to explore and identify the readability and comprehensibility of a text. Most earlier computer-based text analysis programmes or formulae, such as classic readability formulae, determine text difficulty according to word and sentence length, and ignore other readability and comprehension factors (see Danielson, 1987; Schriver, 1989; DuBay, 2004; Graesser et al., 2011; Crossley et al., 2011b; McNamara et al., 2014). The two measures of CRFs absolutely 'have validity as indexes of text difficulty. However, word length and sentence length alone explain only a part of text comprehension' (McNamara et al., 2011, p. 294).

Although classical readability formulae were subject to criticism in the 1980s and 1990s, they have subsequently regained some popularity, according to François (2015). In the first decade of the current millennium, a new approach to readability assessment emerged via advances in computational linguistics. 'Advances in psycholinguistics, discourse processes, and cognitive science provide a theoretical foundation for characterising texts on multiple levels' (McNamara, 2011, p.2). The computational linguistic software Coh-Metrix was developed to combine different features of texts from these fields on a number of levels, both shallow and deeper, to capture text readability and comprehension difficulties. The indices of the computational approach that allow for comprehensive automated text analysis for text and comprehension difficulties are explained further in subsequent sections.

Schriver's (1989) second category, expert judgement-focused methods, can be undertaken by experts who have a high knowledge in assessing the quality of a text. Lastly, Schriver's (1989) third category for assessing text quality is reader-focused methods, which are intended to obtain a direct response from a reader or test the text according to the reader's perspective. Schriver (1989, p. 247) classifies reader-focused methods into two types: (1) concurrent tests, 'which evaluate the real-time problem-solving behaviours of readers as they are actively engaged in comprehending and using the text for its intended purpose', and (2) retrospective tests, 'which elicit feedback after the reader has finished with reading and using the text'.

Examples of concurrent tests are cloze test, behaviour protocols (sometimes called motor protocols), performance testing, and thinking-aloud verbal protocols (also known as think-aloud protocols; Schriver, 1989). The cloze test, for example, is the most common analytical tool in readability research to measure text comprehension. In a cloze test, words are taken out of sentences and readers are asked to fill in the blanks. A major weakness of the cloze test is that it is not always appropriate for testing all types of texts, such as Qur'ānic passages. For example, if this kind of test is given to readers who have no or little knowledge of the subject matter, they will not fill in the blanks but simply leave them without answers. Another essential tool that is not in Schriver's (1989) concurrent tests is eye-tracking. The eye-tracking device was originally used in cognitive science to record eye fixations or movements on a computer screen. Eye-tracking is a common tool in current comprehension and readability studies, and measures eye movement and cognitive processes during reading extracts (for more on eye tracking studies, see Duchowski, 2007; O'Brien, 2009; Hvelplund, 2017; Saldanha and O'Brien, 2014).

Examples of retrospective tests are surveys, interviews and focus groups (Schriver, 1989). According to Göpferich (2009, p. 32), reader-focused tests 'undoubtedly provide the least speculative and most reliable results on text comprehensibility because this is a relative text quality which depends on the audience, whose comprehension and comprehension problems are central for its evaluation'. Retrospective tests, such as the use of surveys, provide comprehension information when readers have been asked to read and evaluate the degree of text comprehension and readability. The use of rating scales in surveys is also an extremely popular technique in testing the comprehensibility and readability of translations (see Section 4.6.1). A survey was chosen for this study as a convenient method for fulfilling the research objectives, exploring whether the deficiencies that influence text comprehension in the chosen translations are associated with the backgrounds of different readers or with the complexity of textual features.

The text-focused and reader-focused methods appear to be the most appropriate approaches for investigating readability and comprehension difficulties in translations of the Qur'ān for two reasons. First, the classic approach and the computational approach both provide an objective assessment of multiple aspects of textual features contributing to readability and comprehension difficulties or making a Qur'ān translation more or less difficult to understand. Second, they provide a fuller understanding of comprehension difficulties in translations of the Qur'ān when comprehension scores are obtained through the human judgement approach. As Klare (1976, p. 146) comments: 'formula scores are, at best, first approximations to difficulty for readers, and human judgements are needed along with the scores'. This applies even to Coh-Metrix scores.

4.4 The Classic Approach

Since the 1920s, many readability formulae have been developed in an attempt to estimate objectively the comprehension difficulty of written materials from the elementary level right up to the adult level. There are many classic readability formulae, but this research uses five widely known major formulae: the Flesch-Kincaid Grade Level (1948), Dale-Chall formula (1948), Gunning Fog Index (1952), SMOG Formula (1969) and Coleman-Liau Index (1975). The classic approach predicts comprehension difficulty or reading ease for texts based on two factors: aspects of word difficulty and sentence complexity (Chall, 1996; Zakaluk and Samuels, 1988).

Table 4.1: Readability Formulae Variables for Estimating Reading Ease

Variables	Flesch-Kincaid	Dale-Chall Formula	Gunning Fog	Smog Index	Coleman-Liau Index
Sentence Length	√	√	√		
Syllables per Word	√				
Hard Words		√			
Polysyllabic Words per 100 words			√		
Polysyllabic Words per 30 Sentences				√	
Letters per 100 Words					√
Sentences per 100 Words					√
Reading Ease Scores	5-16	4-16	6-12	5-18	1-16

Table 3.3 provides a summary of the readability formulae variables for estimating reading ease. Most of these formulae measure two variables to estimate reading ease. Some of them share the same variable of 'average sentence length' to measure sentence complexity, but they differ in terms of other lexical variables for measuring word difficulty. The SMOG Index and Gunning Fog formula use a similar variable (the percentage of polysyllables). But they use different criteria for counting polysyllables. The SMOG Index calculates the percentage of polysyllables per 30 sentences. The Gunning Fog formula calculates the percentage of polysyllables per 100 words. The Dale-Chall formula uses a unique variable, the percentage of difficult words compared to their word lists, meaning that the variables it uses are different from those used in the other formulae. The Coleman-Liau Index also uses two distinctive variables (average number of sentences per 100 words and letters per 100 words).

Gunning's formula can indicate the percentage of polysyllabic words (particularly jargon words). The Dale-Chall formula indicates simplicity in terms of the use of lexical items. Their formula measures the percentage of hard words appearing in their list. Eighty percent of their list is 'known to fourth-grade readers' (DuBay, 2004, p. 23). Flesh's formula is fundamental in reflecting 'the aspects of the difficulty of reading single words and sentences, the word factor may be seen as a measure of semantic difficulty since syllable counts are a reflection of word length […] The higher syllable counts will tend to measure rarity of words and therefore difficulty of meaning' (Gilliland, 1972, p. 91). All these five formulae were developed to measure adult materials (see Flesh, 1948; Gilliland, 1972; Klare 1976; Harrison, 1980, DuBay, 2004).

There are two main reasons for taking the chosen five formulae together. First, this research explores how each formula agrees with the other formulae in terms of the degree of reading ease for translations of the Qur'ān. This comparison might increase the validity of the findings in calculating the scores based on five formulae. Second, this research correlates the formulae scores with the human ratings of comprehension difficulty. However, these formulae are limited to specific variables which cannot predict all the characteristics that make text comprehension difficult (McNamara et al., 2010; 2014). These formulae are also best utilised with other new computational techniques such as Coh-Metrix, which provides a comprehensive automatic readability analysis, as explained in the following section.

4.5 The Computational Approach

The computational approach used in Coh-Metrix moves beyond partial indicators of comprehension difficulty in the classic approach, by analysing difficulties in texts based on

'over 200 measures of cohesion, language, and readability' (McNamara et al., 2004, p. 193). It should be noted, however, that it would not be feasible for this research to analyse the chosen data according to 200 measures. Each of the Coh-Metrix variables or indices for detecting most of the textual features causing comprehension difficulties is described in the following sections. This research makes the most use of multiple levels, operating at the word level, sentence level, cohesion level, and genre level, in order to scale comprehension difficulties in translations of the Qur'ān. The study examines a large set of variables to determine the following factors: the ease of lexical choices, the ease of syntactic choices, the explicit ideas of cohesive elements, and the discourse complexity of the translations of the Qur'ān. Using many multiple variables at each level allows us to view our results from different angles, covering most lexical, syntactic, semantic, and discoursal aspects that cause comprehension difficulties in translations of the Qur'ān. The variables, taken from Coh-Metrix, allow us objectively to examine the chosen versions across a wide range of cognitive, semantic, stylistic, and linguistic features operating at multiple levels of language and discourse, as described in the sections below.

4.5.1 Lexical Measures

This research uses six important variables from Coh-Metrix to measure the most important aspects of word difficulty. Some of them are psycholinguistic-based lexical features, and others are lexical-semantic features.

Word length: This variable is a widely used indicator of word difficulty and reading ease for most of the classic formulae. A lower score in the number of syllables per word indicates words that are shorter and easier to process (McNamara et al., 2014).

Word Frequency: The main obstacle to and predictor of readability is the use of infrequent words (Klare, 1976). Coh-Metrix checks each word's frequency against the CELEX corpus, 'the database from the Dutch Centre for Lexical Information (Baayen, Piepenbrock, and Gulikers, 1995) that analyzed 17.9 million words' (McNamara et al., 2014, p.42). A word frequency score indicates how frequently particular words appear in the English language.

Word Hypernym: Coh-Metrix uses the WordNet corpus (Fellbaum, 1998) to measure this crucial variable. A lower score of hypernym indicates 'an overall use of less specific words', while a higher score indicates 'an overall use of more specific words' (Crossley et al., 2011a, p. 176). According to Duran et al. (2007, p. 234), a lower average score for hypernyms in a text is 'a proxy for word abstractness because the word has few distinctive features'.

MRC Variables: Psycholinguistic-lexical variables were developed by the Medical Research Council (MRC) Psycholinguistic Database, which has been incorporated into Coh-

Metrix (McNamara et al., 2014). The MRC variables are meaningfulness, imageability, concreteness, age of acquisition and familiarity. The last two variables are not used in this research as another variable, 'word frequency', as described above, can detect the unfamiliarity of words. Collins-Thompson (2014) pointed out that these MRC variables can capture cognitive components of reading, which are not immediately addressed by the surface vocabulary and syntactic elements of these classic formulae. These MRC variables are based on human ratings for evaluating content word difficulty (McNamara et al., 2014). Higher scores in these variables reflect a degree of lexical difficulty used in text. A higher average score for concrete words in a text indicates a greater reference to things that can be watched, heard, felt, smelled and tasted, and thus contributes to greater text comprehension (McNamara et al., 2014). A higher average score for word imageability indicates there are words that can easily arouse mental images of things or events in people's minds (McNamara et al., 2014). A higher average score for word meaningfulness indicates words that are strongly associated with other words (e.g., 'people'), whereas words with low scores signify very weak association with other words (e.g., 'amorphous') (McNamara et al., 2014).

4.5.2 Syntactic Measures

Coh-Metrix is significantly different from the classic approach, as it measures many aspects that make the syntactic structure of sentences less or more complex. This research used seven important variables from Coh-Metrix to measure multiple characteristics of syntactic complexity.

Sentence Length: A higher average score for sentence length reflects the fact that 'longer sentences tend to place a greater load on working memory and thereby increase comprehension difficulty' (McNamara et al., 2014, p. 13).

The Number of Words before the Main Verb: A higher average score for the number of words before the main verb in a sentence indicates greater grammatical complexity, which 'places a burden on the working memory of the comprehender' (Graesser et al., 2011, p. 76).

Number of Modifiers per Noun Phrase: This variable has been tested in many studies due to its impact on text processing (McNamara et al., 2014). A higher average score in the number of modifiers per noun phrase indicates greater grammatical complexity as it is more taxing on working memory (McNamara et al., 2014; Graesser et al., 2011).

Syntactic Pattern Density: Syntactic complexity is associated with the density of specific grammatical patterns, word types, and phrase types (McNamara et al., 2014). Coh-Metrix gives a percentage measure for noun density, verb density and adverbial density in a text. The higher

the score for these phrase types, the greater the text difficulty and the higher the proportion of information conveyed using informational density with complex syntax (McNamara et al., 2014).

4.5.3 Cohesive Measures

For the sake of understanding the cohesive variables taken from Coh-Metrix, three main measures for cohesive elements are used in the analysis of the data: (1) six variables for referential cohesion, (2) four variables for semantic coherence, and (3) six variables for connectives. Each of these three measures has its own variables, as described below.

Referential Cohesion: This refers to overlap between content words (i.e., nouns, pronouns, or noun-phrase arguments) in the text (McNamara et al., 2014). Coh-Metrix identifies referential cohesion in texts according to three types: noun overlap (e.g., forgive, forgive), argument overlap (e.g., forgive, forgives) and stem overlap (e.g., forgive, forgiveness, forgiver). These three types measure lexical overlap cohesion at the local and global levels. The local level involves a measure of lexical overlap cohesion between adjacent sentences. By contrast, the global level involves a measure of lexical overlap cohesion (i.e., noun overlap, argument overlap and stem overlap) between all the sentences being analysed (McNamara et al., 2014). All these six variables operating at the local and global levels are useful in estimating the number of lexical overlaps used by the three translators. A high score in these variables indicates greater lexical cohesion and lexical repetition across sentences. In addition, the similarity between Qur'ān translations in scores for these lexical repetition measures might indicate a higher degree of literalness. This is because lexical repetition is the most common stylistic pattern in Qur'ānic genre or discourse (Abdul-Raof, 2001; 2018). By contrast, when there are differences between the translations in these measures, the translator might avoid closeness to the ST cohesive patterns and might select more words which are not from the same root, unlike what Arabic usually does.

Semantic Coherence: Coh-Metrix uses Latent Semantic Analysis (LSA) to measure semantic similarities or relatedness between sentences and paragraphs (McNamara et al., 2014). It gives measures for semantic similarities in passages of text according to the four variables: LSA similarity between adjacent sentences, LSA overlap between all sentences, LSA overlap between adjacent paragraphs, and LSA between given/new sentences. The latter variable is 'a proxy for how much given versus new information exists in each sentence in a text, compared with the content of prior text information' (McNamara et al., 2014, p. 66). LSA measures have been used to estimate the readability and comprehensibility of written materials (Foltz, 2007). Coherence of a passage of text is a crucial factor in comprehensibility (Foltz,

2007). A higher score in these four variables indicates greater semantic similarity between words and sentences.

Connectives: Coh-Metrix provides different measures for all types of connectives based on Halliday and Hasan's taxonomy (1976). These measure all the connectives in a passage of text: causal connectives ('because,' 'so'), adversative and contrastive connectives ('although,' 'whereas'), expanded temporal connectives and temporal connectives ('first,' 'until'), additive connectives ('and', 'moreover'), and logical connectives ('and', 'or', 'if' and 'then') (McNamara et al., 2014). A higher score in these variables indicates more connectives between ideas within the text. Logical connectives are essential for two reasons. First, they indicate syntactic complexity since the presence of logical operators in texts correlates with higher demands on working memory (see Grossley et al., 2007; McNamara et al., 2014). Second, English makes frequent use of asyndetic linkage (i.e., no connector between sentences, in particular), but syndetic linkage (with و *wa-*, or ف *fa-* in particular) is the norm in Arabic (Dickins, 2017). It seems logical to assume that the dense use of the 'and' connector across verses is not typical of English. Thus, dense use of 'and' reflects a high level of literalism on the part of some translators. The fact that some translators remain very close to the original Arabic grammatical structures of the Qur'ān results in unfamiliar grammatical structures in English, and can sound unnatural, thus affecting text processing for readers. For these two reasons, it is worth measuring the use of 'and' in these translations statistically. This analysis might prove the existence of this peculiarity in some translations and identify it as a grammatical issue that influences text readability and comprehensibility.

4.5.4 Genre Measures

As demonstrated previously, Coh-Metrix provides a complete picture of the source of comprehension difficulty at the word level, sentence level, and cohesion level. The characteristics of specific texts (i.e., the genre of a text) are also another source of difficulty that should be considered in assessing the complexity of a passage of text (Binkley, 1988; McNamara et al., 2014). These measures, taken from Coh-Metrix, operating at the word level, sentence level, and particularly cohesion level, are used as predictors, classifying 'particular discourse genre' (McNamara et al., 2014).

Genre is generally about the type of text, which is primarily classified as either narrative, expository, persuasive, or descriptive (McNamara et al., 2014). Coh-Metrix includes a narrativity component to measure the degree of narrativity in texts. This component captures story-telling aspects such as 'characters, events, places, and things that are familiar to the

reader. Narrativity scores indicate the extent to which a text is likely to contain more familiar, oral language that is easier to understand' (McNamara et al., 2014, pp. 85-87).

Narrativity is an important variable for measuring one aspect of text difficulty; however, it is not enough to reflect the primary source of difficulty in translations of the Qur'ān. To determine thoroughly what aspects or dimensions of translations of the Qur'ān translations pose difficulties, this current study employs a factor analysis that was used in previous studies (Biber, 1988, 1995; Louwerse et al., 2004; Bu et al., 2020) to cluster some of the chosen Coh-Metrix measures into dimensions of shared linguistic features that frequently co-occur within the discourse of translations of the Qur'ān. The factor analysis is described in Section 4.7.7.

4.5.5 Coh-Metrix as a Text Analysis Corpus Study Tool

A study using Coh-Metrix is not a participant study. It is rather 'a text analysis corpus study' and the type of corpus used in Coh-Metrix studies is not meant to be similar to corpora such as the British National Corpus or Brown Corpus (McNamara et al., 2014). For the purposes of Coh-Metrix analysis, McNamara et al. (2014, p. 146) define a corpus as: 'a set of written, representative and balanced, computationally readable texts that form a reasonable point of departure as a thematically related language variety, register, genre, or text-type'. This definition focuses on four distinct aspects of building a corpus. The first part, 'computationally readable texts' is relevant in that Coh-Metrix only analyses texts in readable documents, rather than CDs, DVDs, or recorded speech.

The second aspect, 'thematically related', means that every text in the corpus should fall under a single text theme, such as a corpus of newspapers or Qur'ānic translations. The third aspect, 'representativeness and balance', is meant to include major members of texts in terms of including 'a suitable diversity of types of text and a suitable frequency of examples of these types' (McNamara et al., 2014, p. 147). The two concepts, representativeness and balance, make gathering the 'perfect corpus' highly time consuming and costly (McNamara et al., 2014). However, it is not usually necessary to build a huge and perfect corpus for a Coh-Metrix study (McNamara et al., 2014). In the following subsections, we describe the three corpora used in this research for the Coh-Metrix analysis, and then explain how this research cleaned each corpus in a systematic way to increase reliability and validity.

4.5.5.1 Description of Corpora

Three English Qur'ānic corpora, taken from three different translators, were built specifically to allow automated text analyses to be conducted. Each Qur'ānic corpus contains 40 complete Qur'ānic chapters, which differ in terms of their subject-matter and lengths. The three corpora together consist of 120 complete Qur'ānic chapters totalling 61,983 words, as shown in Table 4.2 below. The total number of words in the first corpus is 19,605 and the passages are taken from Abdel-Haleem's version; the second corpus includes 19,823 words from Arberry's version, and the third corpus is made up of 22,555 words from Ali's version. No single chapter in any of the Qur'ānic corpora is less than 100 words in length. Each chapter is a full independent passage.

Table 4.2: Description of the Three Qur'ānic Corpora

	First Corpus	Second Corpus	Third Corpus
Total words in short chapters	2,505	2,380	2,871
Total words in medium chapters	5,156	5,106	6,072
Total words in long chapters	12,162	12,119	13,612
Total words in all chapters	19,823	19,605	22,555
Number of chapters	40	40	40
Number of chapters in the three corpora	120		
Size of the three corpora	61,983 words		

This research follows the common practice of selecting samples (see, Lightman et al., 2007; McNamara et al., 2014; Biber, 1988; 1993) and uses stratified sampling in order to select Qur'ānic samples to avoid the issue of subjective selection. The Qur'ān generally has three types of chapters: short, medium, and long. In choosing 40 Qur'ānic samples, we took 13 chapters of approximately 100-300 words each to represent short chapters (ch. 90, ch. 93, ch. 86, ch. 91, ch. 96, ch. 87, ch. 82, ch. 92, ch. 98, ch. 88, ch. 84, ch. 85, ch. 81), 13 chapters of approximately 300-500 words to represent medium-length chapters (ch. 89, ch. 80, ch. 62, ch. 63, ch. 83, ch. 75, ch. 78, ch. 61, ch. 73, ch. 79, ch. 71, ch. 70, ch. 66) and 14 chapters of approximately 600-1,400 words to represent long chapters (ch. 64, ch. 72, ch. 65, ch. 67, ch. 60, ch. 49, ch. 59, ch. 58, ch. 45, ch. 31, ch. 48, ch. 57, ch. 47, ch. 46).

The 40 Qur'ānic samples collected for each corpus are sufficiently balanced and representative in terms of the diversity of chapters; there is a wide variation among chapter types. They are also adequate in terms of the size of each corpus. McNamara et al. (2014, p. 153), the originators of Coh-Metrix, explain: as a rule of thumb for choosing the corpus size,

researchers are recommended to include ‘at least 20–30 texts for each variable in the analyses they conduct’. Biber (1990, pp. 261-262) argues that ‘10-text samples are large enough to reliably represent a genre. It seems safe to conclude that the 2,000-word and 5,000-word texts in the standard corpora are reliable representatives of their respective text categories for analyses of this type’.

Lastly, the extent to which a corpus is representative largely depends on the variation among text types (Saldanha and O’Brien, 2014). These authors argue that the language of text types such as fiction or journalistic texts is varied, meaning that more texts are required for sources of this type. They add that the language of other text types such as weather forecasts is restricted, meaning that fewer texts are needed. The language of English translations of the Qur’ān is not very varied. As mentioned above, this research has selected two major translations from the 20th century (Ali’s work published in 1934 and Arberry’s in 1959) and another recent and important translation from the 21st century (Abdel-Haleem’s in 2004). The three translators in each corpus have different mother tongues and different levels of academic knowledge about Qur’ānic and Arabic studies. Such variety in selecting different translators might increase the extent to which the corpora are representative. A wide range of textual differences between the three translators is likely to exist since their work covers a long period of time.

4.5.5.2 Cleaning the Corpora

The way in which translators structure paragraphs is the biggest challenge in using Coh-Metrix or readability formulae and it requires significant consideration. The translators do not use similar conventions for paragraphing as one another. For example, Arberry and Ali use the continuous paragraph (i.e., stanza), while Abdel-Haleem uses the prose paragraph. The paragraph differences between the translators were removed to test the variability of scores in Coh-Metrix and the readability formulae. The scores were not wholly different but not highly consistent. If Abdel-Haleem’s paragraphs were retained as they were originally written, a reliable score would not be obtained as the other two translators’ paragraphs are not similar enough to Abdel-Haleem’s. According to McNamara et al. (2014), Coh-Metrix’s results appear to be flawed when texts in the corpus are not cleaned or when the same treatment is not given to all texts. To address this issue, Dowell et al. (2016, p. 85) suggest two ‘best practices’ for cleaning data in the corpus for users of Coh-Metrix, as follows:

- If there is not a good reason to take it out, the researcher should leave it in.

- What the researcher does to one text, should be done to all.

Abdel-Haleem's paragraph structure was altered so that it was parallel to Ali's and Arberry's paragraphs to avoid any inconsistency or misleading scores attributable to paragraph structure, although Coh-Metrix and even the readability formulae do not detect the influence of the structure of a paragraph on text comprehension. Thus, all the translators' paragraphs were made consonant in a systematic manner to obtain reliable and valid results. The footnote numbers and letters given between verses and the numbers of verses in each version were removed to avoid inconsistent scores. All the Qur'ānic chapters in the three corpora were manually cleaned and transferred into plain texts. They were then entered individually one by one into Coh-Metrix (Version 3.0) using the Coh-Metrix Web Tool and the Automatic Readability Checker website for the five readability formulae.

4.6 The Human Judgement Approach

The oldest approach to estimating comprehension difficulty is the use of human judgement (Chall, 1996; Van Slype, 1979). Human judgement has been used to develop most classic readability formulae (Chall, 1996) and even other more recent automated linguistic and psychological measures (see McNamara et al., 2014). The main purpose in using the human judgement approach in this thesis is to obtain readability and comprehensibility ratings, and to correlate the ratings, as judged by humans, with automated measures derived from the previous two approaches. In the following subsection, a brief overview of the techniques used in previous studies to assess readability and comprehensibility is provided. The design stage for re/building the survey of this study is then discussed.

4.6.1 Rating Scale

Van Slype (1979) suggests several techniques that can be employed to assess translation comprehension specifically: (1) a rating scale, (2) a cloze test (discussed previously in Section 4.3), (3) a multiple-choice questionnaire, and (4) a knowledge test. The multiple-choice questionnaire and knowledge test are more concerned with the fidelity or accuracy of information transfer between a TT and an ST. These last two techniques for assessing translations of the Qur'ān typically require general agreement about the meaning of a verse among Qur'ānic exegetes when respondents' answers are assessed. These two techniques do not serve the purposes of this research since it focuses only on the readability of TTs rather

than the accuracy of information transfer between a TT and an ST. This research only explores the factors that make translations more readable and comprehensible.

The use of a rating scale for assessing text quality (see Klare, 1963; Gilliland 1972; Davison et al., 1988) or translation quality (see Halliday and Briss 1977; Van Slype, 1979; Tiselius, 2009; Angelelli, 2009; Saldanha and O'Brien, 2014) has a long history. A rating scale is commonly used for obtaining responses from respondents on measures such as adequacy, fidelity, acceptability, readability, and comprehensibility. The former two measures are generally used to test a TT and its relationship to an ST. In comparison, the latter three measures (i.e., acceptability, readability, and comprehensibility) are usually used to focus on a TT and its relationship to the norms of the TT language and culture. In many studies, readability and comprehensibility are measured using various rating-scale formats or pre-established criteria in surveys in which the evaluator uses these criteria to assess text or translation quality (for a review see Carroll, 1966; Sinaiko and Klare, 1972; Halliday and Briss 1977; Van Slype, 1979; Charrow 1988; Zanón, 2005; Fiederer et al., 2009; Kandula and Zeng-Treitler, 2009; Tiselius, 2009; O'Brien, 2010; Hansen-Schirra et al., 2015; Mujiyanto, 2016.)

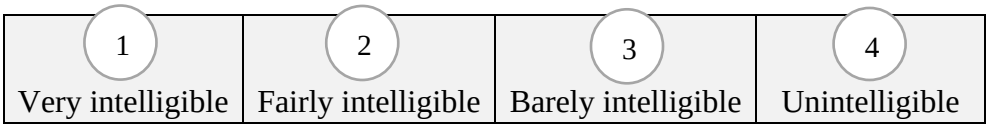


Figure 4.1: An example of a 4-point intelligibility scale

When a scale contains more than 7 points for its rating description, this can distract or confuse the respondent. About half of these studies employed an odd number of choices while the others were designed with an even number of response options. Most of the studies developed their own scale descriptors while some used previous scales. Some of these studies used a long description for each number, while others used a short one as shown in Figure 4.1, the example being taken from Van Slype (1979). The former option might give the evaluators a more complete description of how passages or sentences should be rated, but it is somewhat hard on evaluators because it takes considerable time to remember and it increases the cognitive burden on the respondent if each value involves a long description (Carroll's scale (1966) is a good example of a long description). On the other hand, the latter option gives evaluators a short description of how passages or sentences should be assessed. This option is much easier on evaluators than the former option because it is quicker to remember and respond to. Lastly,

some of the data samples for rating the quality of translations or texts focus on isolated sentences while others involve longer excerpts.

4.6.2 Qur'ānic Chapter-Samples

The Qur'ānic chapter-samples presented to participants for them to evaluate were taken from two complete chapters: chapter 91, which has 15 verses, and chapter 86, which has 17 verses, giving a combined total of 32 verses for evaluation. These two chapter-samples from the three translators were included in the web-based survey for individuals without any manipulation of their original production.

These two chapters were chosen because they each express a complete thought. They include different Qur'ānic themes that might or might not require prior knowledge. The main theme of the Sun, chapter 91, is about 'purifying or corrupting the soul, with the tribe of Thamud given as an example of corruption' (Abdul-Haleem, 2004, p. 423). This chapter as a topic is not hard to understand, but contains some cultural and historical references which require prior knowledge. The Night-comer, chapter 86, 'focuses on a series of examples of things coming out: the piercing night-star, spurting semen, the baby that bursts out of the womb, and plants that sprout out of the ground. All of these are used to illustrate resurrection from the grave' (Abdul-Haleem, 2004, p. 417). This chapter does not include any cultural or historical words; thus, it might not require prior advanced background knowledge.

These two complete samples were chosen as they offered a convenient sample for the survey. Participants could read and wholly evaluate these chapters more easily than longer chapters. It is preferable not to give readers isolated verses to evaluate without context, as this might prevent readers from understanding the meaning of the verse in relation to its context. Some verses in the Qur'ān can be better understood from the details in the previous or subsequent verses. Some Qur'ānic translators provide introductory information about Qur'ānic chapters to enhance text comprehension. This study will test the influence of this additional information on ease of readability of the texts compared to translations which do not include introductory information. The complete text is the best way of assessing text quality in any parameter, reflecting how readable and comprehensible the overall content of a passage of text is. In Halliday and Briss's study (1977, p. 11), 'test subjects were asked to evaluate complete texts, rather than single sentences or textual units. Therefore, in this test, context played an important part in determining intelligibility'. Giving readers a whole chapter to read and evaluate the degree of text comprehension and readability is important.

4.6.3 Survey Design

Prior to designing the survey, it was necessary to consider the broader context for this research design. Including three comparable surveys for three translators (and even the three comparable corpora in the previous section) are clearly point to a cross-sectional research design. Cross-sectional research, in Dörnyei's words (2007, p. 78), is 'a snapshot-like analysis of the target phenomenon at one particular point in time; focusing on a single time interval. It allows us to establish relationships between variables and to find out about the participants' thoughts, attitudes, and emotions as well as various cognitive and personality traits'.

An initial survey design was undertaken which consisted of main three sections for completion. The first section described the purpose of the survey and asked participants to rate the readability and comprehensibility of three English versions of the Qur'ān. The evaluation was framed so as to rate: (i) the ease of comprehension of translated verses and titles, (ii) the influence on comprehension of an introduction and footnotes for text, and (iii) the simplicity and clarity of non/linguistic elements such as word usage, grammar usage, cohesion usage, and text formatting (i.e., the presentation of the page layout) in each English version of the Qur'ān.

The second section of the initial survey was made up of nine different closed-ended questions about the general background of the participant, including: age, gender, education, specialism, religion, knowledge of Arabic, knowledge of English, and knowledge of the Qur'ān (see Participant Background Information, in Appendix B). The last section of the survey included two instructions for completing the survey. The first explained that the chosen translations were labelled Version A, Version B and Version C. There was a separate table for rating each version. The second instruction asked participants to rate the lexical usage, grammatical usage, cohesion usage (links between elements), page layout and ease of comprehension for each version. Each of these five categories required participants to answer on a 5-point Likert scale and they were asked to agree or disagree with various statements. The total number of statements for each version was 31. A pilot study was carried out based on this design, as reported in the next section.

4.6.4 The Piloting of the Initial Survey

The main point of piloting the initial survey was to make sure that the survey was feasible and valid for measuring what this research intended to assess; it was also an opportunity to pick up any errors made in the initial survey. The pilot study, which was conducted on the campus of Leeds University in November 2019, involved responses from two participants: an expert and a student.

The expert respondent was a male professor in intercultural studies, a native speaker of English, and a non-Muslim with a moderate knowledge of the Qur'ān. The student respondent was a female doing a bachelor's degree in linguistics, a native speaker of English, and a non-Muslim without prior knowledge of the Qur'ān. The results of the pilot study indicated that two out of the three translations were clearly different in terms of their level of text readability and comprehension. This pilot study indicated that the design of the initial survey was valid for measuring what this research intended to measure since it included different levels of textual variables to rate translation readability and comprehensibility.

The pilot study obtained very constructive feedback from the two respondents, who were asked several questions focusing on issues related to the practicalities of the survey and improving the survey design, such as: its comprehensibility and the logical sequence of the survey questions, the simplicity and clarity of the survey instructions, the clarity of the attached images in the survey, the appropriateness of the verse division for evaluation, and the length of the survey. Both respondents agreed that the images attached for reading were clear, the survey instructions were simple and clear, and the questionnaire questions were comprehensible and logically presented. Such agreement between the two respondents on these factors supports the validity of the survey design. The appropriateness of the verse division for evaluation was another important issue that was identified during the pilot study. In the pilot survey, three verses were combined into one statement (e.g., 'all meanings of verses 1, 2, and 3 are easy to understand') and the overall average comprehensibility of all three was asked about on a 5-point scale. One of the respondents said that the first verse, for example, was easy to understand while the second verse was not easy. In the final survey design, statements regarding the comprehensibility of verses were modified such that each verse was assessed separately.

Lastly, the length of the survey in terms of completion time was a very important issue. The researcher timed how long it took each respondent to complete the survey. The expert respondent completed the survey in 35 minutes, while the student finished it in 42 minutes. To tackle the issue of survey length, the number of constructed statements were reduced, and another scale procedure was adopted to reduce the length of the survey (see Section 4.6.5). To reduce the time it would take to complete the survey and minimise the response bias of individuals for a particular translation, an individual survey for each translation was adopted in the final version of the survey. The new individual survey for each translator takes 15-20 minutes (approximately) to complete. The pilot study was important in offering these research lessons and insights and led to an improved survey design. The following sections will highlight several changes that were considered before the final distribution of the survey.

4.6.5 Multi-Item Scale for Textual Variables

A Likert scale was adopted in the initial version of the survey, which had 31 statements for each translation. The total number of statements for the three translators was 93. These statements were based on five main levels: word usage, grammar usage, cohesion usage, page layout, and the ease of comprehension of a Qur'ānic chapter. Figure 4.2 provides a sample from the initial survey for rating word usages in a Qur'ānic chapter.

Q.1 After reading this translation, to what extent do you agree or disagree with the following statement? 1= strongly agree, 2= agree, 3= neither agree nor disagree, 4= disagree, 5= strongly disagree.	Please Tick One
The words are largely familiar.	1 □ 2 □ 3 □ 4 □ 5 □
Some words are archaic.	1 □ 2 □ 3 □ 4 □ 5 □
Some words are not understandable, as they are not normal English words.	1 □ 2 □ 3 □ 4 □ 5 □
The meaning of some words is ambiguous.	1 □ 2 □ 3 □ 4 □ 5 □
The length of words in most verses is short.	1 □ 2 □ 3 □ 4 □ 5 □

Figure 4.2: Sample from the Initial Version of the Survey

In the pilot study, the use of a Likert scale was shown to be problematic. It was not suitable for the current study, which attempts to rate multiple aspects of readability and comprehension, but it would increase the time needed to complete the survey if many statements were included, as demonstrated in the pilot study. The use of a Likert scale with many statements might cause participants to give up, thereby reducing response rates. Another rating scale technique, semantic differential (SD), was adopted in the final version of the survey to overcome the difficulties observed in the pilot study.

Q1. Overall, how would you rate the word usages of Chapter 1 on the following criteria?						
	1	2	3	4	5	
Familiar words						Strange words
Modern words						Archaic words
Frequent English words						Infrequent English words
Clear words						Confusing words
Short words						Long words

Figure 4.3: Sample from the Final Version of the Survey

SD is a type of rating scale, which was developed by Osgood et al. (1957). The SD technique is usually used when asking people to rate an issue or subject using a group of adjectives with opposite meanings. This semantic differential rating scale used in the final version had many advantages over the initial scale such as: its ease of use, to the fact that it can be completed quickly compared to reading many statements, and that one question covers many aspects. Originally, two opposite adjectives only are used in an SD. In this research, nouns were combined with adjectives to facilitate comprehension on the part of evaluators, as shown in the example below. Figure 4.3 is a sample from the revised version for rating the word usages above. The full and final version of the survey for each translator is provided in Appendix E. When the evaluators had read a whole chapter, they were asked in the first stage of the survey to rate the comprehensibility of the passage they had just read based on four variables:

- The comprehensibility of translated titles.
- The comprehensibility of translated verses. The translated verses in each chapter were rated on a verse-by-verse basis on the SD scale from 1 to 5 (with 1 being a comprehensible verse and 5 an incomprehensible verse; see the full survey in Appendix E).
- The influence of any introduction provided for text comprehension.
- The influence of any footnote given for text comprehension.

These variables were included in the survey in each translation which was rated, except those translations that did not include any footnotes or introductions to the chapters in the original. The framework for all these variables of comprehensibility was extracted from the chosen translations to assess objectively the factors that contributed to text comprehension. In the second stage of the survey, the evaluators were asked to rate the readability of each Qur'ānic chapter based on four primary levels of ratings:

- Six pairs of adjectives for rating the page layout.
- Five pairs of adjectives for rating the word usages.
- Four pairs of adjectives for rating the sentence usages.
- Three pairs of adjectives for rating the cohesion usages.

Each of these main levels included multidimensional variables, which provided complete coverage of the various aspects of comprehension difficulties in translations. The selection of

two adjectives with opposite meanings for these main levels was taken into consideration. All the negative adjectives were positioned on the right-hand side and all the positive adjectives were positioned on the left to avoid any confusion, as shown above. The adjectives were chosen with care to avoid self-selection, a systematic selection of two adjectives with opposite meanings being chosen from *Roget's Thesaurus* (1990) as an external source and some suggestions being obtained from my supervisor, James Dickins. Osgood et al.'s study (1957, p. 47) used *Roget's Thesaurus* to 'obtain a logically exhaustive sampling of semantic dimensions which would also be independent of our own theoretical biases'.

As discussed in Section 4.2.2, the reliability of a survey can be determined by its internal consistency. The internal consistency of the multi-item scales for the three surveys after distribution was measured using Cronbach alpha reliability coefficients and the reliability analysis in SPSS. The multi-item scales for readability and comprehensibility measures showed a high level of internal reliability, with a Cronbach's alpha of above .70 in each question in Abdel-Haleem's, Arberry's, and Ali's surveys. If the Cronbach's alpha is above .60, it is satisfactory (see Dörnyei, 2007).

4.6.6 Prior Knowledge

Reader variables, particularly relating to different prior knowledge of the Qur'ān, are another important factor in this research. As stated in Chapter 3, the reader variables interact with the textual variables, the latter showing the effects of text readability and comprehension on readers with differing amounts of prior knowledge. This research distinguishes between two groups with regard to prior knowledge, as follows: (1) those with high or moderate knowledge of the Qur'ān, and (2) those with low or no knowledge of the Qur'ān. The emphasis of this research is on the latter group, those with low or no knowledge. The first group might be not critical enough due to their language skills or high prior knowledge, which both usually compensate for complexity in texts. The differences between the results for those with low and high prior knowledge of the Qur'ān will help us to recognise if the deficiencies that influence comprehension difficulty are due to different reader backgrounds or to some textual features in certain translations of the Qur'ān.

4.6.7 Survey Type and Distribution

A web-based survey was implemented using https://www.survey.leeds.ac.uk (BOS), which is designed in accordance with the Data Protection Act. The use of a web-based survey was

intended to maximise the external validity of this research compared to the printed surveys that were used in the pilot study. The web-based survey employed two modes of distribution.

Mode 1 was adopted in the first stage of the survey collection and involved emailing random potential participants, who worked and studied at the University of Leeds. The School of Modern Languages and Cultures also forwarded three individual links from the researcher (the first link was designed for Abdel-Haleem's survey, the second for Ali's and the third for Arberry's) to some potential students and/or staff at the University of Leeds. Using this mode, the first survey for Abdel-Haleem received 12 responses, that for Arberry 10 responses, and that for Ali 13 responses. Five participants were excluded because they were not English native speakers, although it was stated clearly in the information sheet for each survey that participants should be native English speakers.

Mode 2 was employed in the second stage of data collection; the remaining participants were recruited by the researcher in person. The links for the three surveys were downloaded onto an iPad device with the goal of obtaining 10 male and 10 female participants for each survey. The remaining participants were found in Laidlaw Library, Brotherton Library and Edward Boyle Library at the University of Leeds. This mode was used as a strategy to increase the response rate and achieve stratified sampling. Lastly, an incentive was provided in the web-based survey to attract participants to complete the survey. Some researchers view the use of incentives as 'ethically unsound' (Saldanha and O'Brien, 2014), although many are in favour of this as it increases response rates (see Frey et al., 2000; Dörnyei and Taguchi 2009).

4.6.8 Sampling

Through these multi-mode survey collections, this research recruited 20 participants for each survey,[8] meaning that the total number of participants across the three surveys was 60. A descriptive breakdown of participants' characteristics in each survey (gender, age, religion, mother tongue, education level, and prior knowledge of the Qur'ān) is provided in the table below.

Each survey had a gender balance of 10 males and 10 females within an age range of 18 to 64 years old, except the survey for Arberry which had no participant of 64 years old. There was a fairly even spread across the 60 participants in terms of the number of Muslims and non-Muslims in each survey. All 60 participants were native English speakers, except for one (in Abdel-Haleem's survey), who was a native speaker of both English and Arabic. The other 59

[8] According Dörnyei (2007), there is not an absolute rule for the optimal size of a population sample, but it is advisable to have at least15 participants for each group in comparative procedures.

participants were not native speakers of Arabic, but had either low or high background knowledge of Arabic.

Table 4.3: Descriptive Breakdown of Participants

		Abdel. Survey	Arberry Survey	Ali Survey
	Total participants	20	20	20
Variable	Category			
Gender	Male	10	10	10
	Female	10	10	10
Age	18-24	9	15	13
	25-34	3	1	5
	35-44	5	3	0
	45-54	2	1	1
	55-64	1	0	1
Level of Education	High school	9	15	13
	Bachelor's degree	3	1	5
	Master's degree	5	3	0
	Doctoral degree	2	1	1
Language	Native speakers of English	20	20	20
	Native speakers of Arabic	1	0	0
	Low prior knowledge of Arabic	14	17	17
	High prior knowledge of Arabic	6	3	3
Qur'ān Knowledge	Low prior knowledge	11	13	12
	High prior knowledge	9	7	8

Overall, the participants in each survey had different levels of education and diverse majors. The participants were asked to rate their knowledge of the Qur'ān according to three factors: whether they had read the Arabic Qur'ān, whether they had read English translations of the Qur'ān, and whether they had read interpretations of the Qur'ān. Abdel-Haleem's survey had 11 participants with low prior knowledge of the Qur'ān and 9 participants with high prior knowledge of the Qur'ān. Ali's survey had 12 participants with low prior knowledge of the Qur'ān and 8 participants with high prior knowledge of the Qur'ān. Arberry's survey had 13 participants with low prior knowledge of the Qur'ān and 7 participants with high prior knowledge of the Qur'ān.

4.6.9 Ethical Considerations

During the earlier stages of this research, the ethical aspects of research involving human participants were clarified in the Research with Human Participants course, which was conducted by Research Ethics, the University of Leeds. In attending this course, many ethical

aspects such as sensitive topics, data storage, vulnerable participants, potential risks to the participants and the researcher, avoiding coercion and anonymity and confidentiality were clarified before ethical approval was obtained. Before obtaining approval, the risks and purposes of this project, the methods of collecting the data, the type and number of participants, the type of data to be collected, the protection of data after collection and the participants' right to privacy were clarified for the Faculty of Arts, Humanities and Cultures Research Ethics Committee. The Research Ethics Committee at the University of Leeds granted ethical approval for this research in 2019.

The potential participants were provided with a Participant Information Sheet (see Appendix A) for each survey, which gave details of what participants were required to do for the survey, why the research was being carried out and what it would involve. Before answering the online survey, a Participant Consent Form was also provided and the potential participants were asked to tick a box confirming that they had read the statement and understood the information provided. The Participant Consent Form can be found in Appendix B. Each survey covered the following ethical aspects: the purpose of this research; instructions for completing the survey (see Appendix D); the need for targeted participants to be native English speakers; the length of the survey, which is about 15-20 minutes; the fact that the answers would only be used for the purpose of this research; the fact that all of the data obtained would be treated as confidential and stored securely as required by the Data Protection Act; that taking part in this study was voluntary; that participants could withdraw at any time without giving a reason, and that any responses already provided would be retained due to the anonymity of the responses.

4.7 Statistical Analysis

This research used several dependent variables to measure readability and comprehensibility between the three different translations. The variables, taken from the automatic assessment, were organised according to their conceptual nature: five formulae for reading ease, six variables for lexical difficulty, seven variables for syntactic complexity, six variables for referential cohesion, four variables for semantic coherence, and six variables for connectives. For the human ratings, five levels of rating were included in the survey design: four variables for rating text comprehension, six variables for page layout, five variables for word usages, four variables for syntactic usages, and three variables for cohesion usages. To determine the differences and relationships between these variables in each translation, this research used different types of statistical tests for the automatic assessment and human ratings. In analysing

the scores from the automatic assessment, this research employed the following statistical procedures:

- ❖ A one-way MANOVA was used to examine the differences between the three translations based on the combination of the dependent variables.
- ❖ Another fellow-up test, a one-Way ANOVA with post-hoc analysis, was employed to determine how these translations differed from each other in terms of individual variables.
- ❖ Factor analysis was conducted through principal component analysis to provide a shared group of linguistic features that frequently co-occur in English translations of the Qur'ān.

In analysing the results from human ratings, this research applied the following statistical procedures:

- ❖ Independent-samples t-tests were used to compare human ratings for readability and comprehensibility in each survey.
- ❖ A correlation analysis using Pearson's correlation coefficient was used to examine the relationship between readability scores and comprehension scores.
- ❖ A multiple regression analysis was conducted to determine which of the automated measures were most predictive for comprehension difficulty in each translation of the Qur'ān.

All the scores from the human ratings and automatic assessment were manually entered into Statistical Package for the Social Sciences (SPSS) in order to conduct all these statistical analyses, based on step-by-step SPSS tutorials from the Laerd Statistics website (Laerd Statistics, 2015). The following subsections will explain these statistical tests.

4.7.1 Testing for Normality

In using the parametric tests below, the data are assumed to be normally distributed, although linguistic data are often skewed (Saldanha and O'Brien, 2014). The Shapiro-Wilk test is used to detect data normality. If the p-value is less than .05 in the Shapiro-Wilk test, this means that the data are not normally distributed and vice versa. When data are not normally distributed, it is not a serious matter in parametric tests; the non-normality of data does not influence a Type

I error[9] because the parametric tests are considered to be 'robust' to non-normality (see Elliott and Woodward, 2007; Maxwell and Delaney, 2004 cited in Laerd Statistics, 2015; Mellinger and Hanson, 2016). Any departure from normality for any variable will be reported in the results of the parametric tests. Lastly, two-tailed significance is used for all the tests below as our main hypothesis is non-directional (see Chapter 1); the non-directional hypothesis does not specify which translation of the Qur'ān is expected to have higher text readability and comprehensibility.

4.7.2 Independent-Samples T-Test

The independent-samples t-test is also referred to as a student's t-test (Laerd Statistics, 2015). It investigates differences in the mean values of the same dependent variable between two different groups.

4.7.3 One-Way MANOVA

A one-way multivariate analysis of variance (MANOVA) is an extension of the ANOVA (Laerd Statistics, 2015). For this study, the MANOVA test allows us to include two or more independent variables (i.e., the three translators are treated as independent variables) and two or more dependent variables (i.e., all six chosen lexical variables are treated as dependent variables), allowing us to examine all of them at once. It mainly examines if the three independent variables/translators differ significantly based on the combination of the dependent variables. When a MANOVA finds there is a significant difference, a follow-up test such as ANOVA, along with a post-hoc test, is typically used to determine where the differences in the dependent variables have occurred across the translators.

4.7.4 One-Way ANOVA

A one-way analysis of variance (ANOVA) is an extension of the t-test that is used to investigate only one dependent variable in the context of two or more independent variables (Laerd Statistics, 2015). This test enables us to examine, for example, which one of the chosen lexical dependent variables results in a significant difference across the three translators. The ANOVA

[9] There are two types of errors which may occur in a hypothesis test: a Type I error and Type II error (Elliott and Woodward, 2007). A Type I error occurs when the null hypothesis is rejected although it is true, while a Type II error occurs when the null hypothesis is false and is not rejected. To reduce the chance of committing a Type I error, the p-value of .05 should be reduced (see Section 4.7.3).

does not indicate how these translators differ from each other in terms of these individual lexical variables, however. A Bonferroni post-hoc test is used after the ANOVA to determine how these translators differ from each other in terms of each individual variable. The reason for using the Bonferroni post-hoc test is that the cut off for statistical significance is .05, which is the default p-value used in SPSS and all statistical tests. This research adjusted the p-value of .05 to reduce the danger of a Type I error when the Bonferroni post-hoc analysis made multiple comparisons between translators for each variable (Elliott and Woodward, 2007). The Bonferroni-adjusted significance criterion is used for these comparisons. In this adjustment, the p-value of 0.05 was multiplied by the number of pairwise comparisons (i.e., .05/3= .016). The p-value of 0.016 is considered statistically significant in reporting the results of the pairwise comparisons.

4.7.5 Correlational Analysis

Most of the above statistical tests determine the significant difference between variables and not the linear relationship between two variables. In correlational analysis, a correlation coefficient can be estimated either from Pearson's correlation or Spearman's correlation to 'look at two variables and evaluate the strength and direction of their relationship or association with each other' (Dörnyei, 2007, p. 223). The value of the correlation coefficient between two variables ranges from 1.00 to +1.00. A higher value for the correlation coefficient indicates a positive correlation between two variables, while a lower value for the correlation coefficient indicates a negative correlation between two variables.

4.7.6 Multiple Regression

Multiple regression is also referred to as 'multiple linear regression'. It is commonly used in readability research to estimate the highest predictors of comprehension difficulty or text difficulty (see Klare, 1988; Zakaluk and Samuels, 1988; Chall, 1996). Multiple regression is a form of linear regression that uses only one explanatory independent variable to predict a dependent variable. Unlike simple linear regression, multiple regression makes it possible to incorporate two or more explanatory independent variables to predict a dependent variable (Laerd Statistics, 2015).

In the first phase of analysing automated measures, this research identified several automated variables from the classic approach and the computational approach, on the assumption that these automated measures are the most predictive of comprehension difficulty

for translations of the Qur'ān. The best approach for identifying predictors of comprehension difficulties in Qur'ān translations is to use a multiple regression analysis, which makes it possible to represent the greatest number of predicators or the highest loadings on the regression model to estimate comprehension difficulty for translations of the Qur'ān. In conducting the regression model, the outliers and multicollinearity (i.e., those variables that are highly correlated with each other) for these identified automated measures, variables should be removed. Any high correlation between these automated variables (i.e., multicollinearity) can be identified by inspection of correlation coefficients (Laerd Statistics, 2015). Multicollinearity and non-significant predictors among these automated variables are excluded from the regression model analysis to understand which variables are most predictive of comprehension difficulty in translations of the Qur'ān.

Statistical multiple regression analysis is critical for further explaining the findings of automated measures in Chapter 5 in relation to the results for the human comprehension ratings in Chapter 6. It is also very useful for determining if the chosen twenty-nine Coh-Metrix indices are more predictive than the chosen five readability formulae for comprehension difficulty in translations of the Qur'ān.

4.7.7 Factor Analysis

Factor analysis is a multidimensional approach to 'textual variation' (Biber, 1988) or classifying a 'particular discourse genre' (McNamara et al., 2014). The most common investigation of linguistic variation is Biber's (1988) study. Biber (1988) used a principal factor analysis (PFA) to compare linguistic variation across spoken and writing registers based on 67 linguistic features at the word level. Another piece of research conducted by Louwerse et al. (2004) used a principal component analysis (PCA) to compare linguistic variation across spoken and written registers. Their study used 236 linguistic features at the text level (i.e., the word level, to sentence, paragraph, and discourse level). More recently, Bu et al. (2020) used 67 linguistic features identified by Biber (1988) and applied a PCA to explore the linguistic co-occurrence patterns in car discourse.

The similarity between a PCA and a PFA is that they use a similar procedure (i.e., an item or variable 1, 2, 3 measures one construct and a variable 4, 5, 6 measures another construct, etc.). In other words, they cluster and reduce a larger set of variables into groups that are referred to as 'factors' or 'dimensions'. The difference between the two is that 'a PCA attempts to account for all of the variance in the data, while a PFA attempts to account for only the

shared variance' (Biber, 1988, p. 82). This study opted for a PCA[10] using a Promax rotation because the co-occurring patterns of the selected linguistic features were not already known in English translations of the Qur'ān (Biber, 1988, 1995; Louwerse et al., 2004).

The PCA gathered the Coh-Metrix indices into groups that co-occur frequently within the discourse of English translations of the Qur'ān. Those 29 linguistic features taken from Coh-Metrix operated at the word level, the sentence level and the cohesion level. This study followed these studies in terms of the absolute value of loadings (where a loading is a value to signify an important or salient variable in the analysis) in each variable. Loadings that had an absolute value of less than .35 were removed from the analysis (Louwerse et al., 2004; Biber, 1988). This allowed only the salient and important linguistic features or indices to be retained in the analysis. Lastly, this study also followed the statistical procedure of those previous studies that standardised the scores of all variables to a mean of 0.0 and a standard deviation of 1.0 on a single scale (i.e., Z-scores) before the scores were computed (Biber, 1988, 1995; Louwerse et al., 2004).

This factor analysis is significant for this study for two reasons. First, it provides a shared group of linguistic features that co-occur frequently in English translations of the Qur'ān. The co-occurrence patterns of the linguistic features for each factor are interpreted based on their functional relationships to the studies of linguistic complexities (Biber, 1988;1992; McNamara et al., 2014) and to the ST features examined by Abdul-Raof (2019; 2001). Second, it allows this research to explain the source of difficulty within translations of the Qur'ān from a broader perspective.

[10] Varimax and Promax are commonly used (Biber, 1988). Varimax requires factors that are not correlated, while Promax requires 'minor correlations among the factors' (Biber, 1988, p. 85). A Promax rotation is preferred as there is no need to assume that the factors in textual variation are completely uncorrelated (Biber, 1988). A PCA is also used 'because the underlying structures were undefined. In such cases, a PCA is appropriate because it reduces the variables to meaningful units' (Crossley et al., 2014, p. 201).

4.8 Chapter Summary

This chapter has discussed the multiple methods and measures that were relevant to this research. The difference between multimethod and mixed methods and the reasons behind using multimethod in this research were clarified. This introductory section was followed by a synoptic outline of the three quantitative approaches used for measuring text readability and comprehension: the classic approach, the computational approach – Coh-Metrix, and the human judgement approach. It also elucidated certain issues relating to validity and reliability, the sources of Qur'ānic chapter-samples, the methods and measures used to assess comprehension difficulty, the description of the three corpora for the automated text analysis, the pilot study for the survey design, the sampling selection, the reader type required for this research, the type and administration of surveys, and a description of ethical considerations. Lastly, this chapter concluded with a discussion of the statistical analysis procedures, allowing us to quantitatively discover patterns of association between each translation and establish the differences between variables in each translation. The following two chapters present the data analysis of all the variables, taken from human ratings and automatic assessment tools, i.e., the classic approach and the computational approach.

Chapter 5: Data Analysis of Automated Measures

5.1 Chapter Overview

This chapter investigates three corpora of English Qur'ān translations – namely, Abdel-Haleem (2004), Arberry (1955), and Ali (1934) – based on several automated measures of text readability, cohesion, and comprehension, adopted from two automated approaches to textual analysis: the classic formulae and the computational Coh-Metrix. The approach taken to assess the readability of these translations includes a 'multilevel theoretical framework for language and discourse processing' (McNamara et al., 2014, p.15) and a multidimensional analysis of the discourse complexity of specific texts within the same register (Biber, 1992).

The automated measures under investigation in this chapter are five classic readability formulae [CRFs] of reading ease, six variables of lexical difficulty, six variables of syntactic complexity, six variables of referential cohesion, four variables of semantic coherence, seven variables of connectives, and narrativity (a measure of the genre of a text). As stated in Chapter 4, the analysis of these variables is divided into two parts. The first part examines whether there are differences between the mean values of the variables for these three translators. The second part uses a multi-dimensional analysis of textual variation and explores a set of linguistic features and patterns that frequently co-occur in these three Qur'ān translations. This is intended to identify the sources of discourse complexity in the linguistic characteristics of English Qur'ān translations. These two parts describe translation readability in relation to different levels of text factors: *style*, *cohesion*, *literalness*, *genre*, *register*, and *retranslation*. A summary and discussion of these factors are provided at the end of this chapter, followed by a chapter recap.

5.2 Part One: Differences Between Groups

To examine each of the automated measures as a single group, according to their conceptual frameworks, this research involved six one-way multivariate analyses of variances (MANOVA) in SPSS. The MANOVAs examined whether there were differences between the three translators in their combinations of the dependent variables in each of the following frameworks: readability formulae of reading difficulty, variables of lexical difficulty, variables of syntactic complexity, variables of referential cohesion, variables of semantic coherence, and variables of connectives. There was homogeneity of variance for all the automated measures, as assessed by the Levene's test of homogeneity of variance ($p > .05$).

On examination of these five readability formulae as one group, the MANOVA revealed a statistically significant differences in 'reading ease' for the chosen Qur'ān translations: $F (10, 106) = 11.802$, $p < .0005$; Wilks' $\Lambda = .224$; partial $\eta 2 = .527$. Second, the MANOVA found a statistically significant difference between the translators in their combinations of the six-predictor lexical variables: $F (14, 102) = 4.803$, $p < .0005$; Wilks' $\Lambda = .363$; partial $\eta 2 = .397$. Third, the MANOVA revealed a statistically significant difference between the translators in their combinations of dependent variables of syntactic complexity: $F (18, 98) = 2.828$, $p < .0005$; Wilks' $\Lambda = .433$; partial $\eta 2 = .342$. Fourth, the MANOVA found that there was a statistically significant difference between the translators for the combined dependent variables (i.e., the six variables for referential cohesion), $F (12, 104) = 2.712$, $p < .0005$; Wilks' $\Lambda = .580$; partial $\eta 2 = .238$. Fifth, the MANOVA found a statistically significant difference between the translators in their combined dependent variables (i.e., the four variables of semantic coherence): $F (8, 228) = 2.880$, $p < .0005$; Wilks' $\Lambda = .825$; partial $\eta 2 = .092$. Finally, the one-way MANOVA found a statistically significant difference between the translators in their combined dependent variables (i.e., the six connective variables): $F (12, 104) = 3.136$, $p < .0005$; Wilks' $\Lambda = .539$; partial $\eta 2 = .266$.

The results of the MANOVAs suggest significant differences between the three translations on multiple computational variables of text characteristics including lexical, syntactic, semantic, and discourse features. The lexical, syntactic, semantic, and cohesive differences can describe what the text factors are that make a translation of the Qur'ān more readable and comprehensible. Follow-up ANOVAs with the Bonferroni post hoc analysis were conducted to further interpret these differences and determine how these three translators differed for each of the variables in these six constructs. These results for each variable are discussed in the following sections in relation to their conceptual frameworks.

5.2.1 Analysis of Reading Ease

As revealed in the MANOVA results, the reading ease of the three Qur'ān translations differed when assessed using the five CRFs: fog index, Flesch-Kincaid grade level, Coleman-Liau index, SMOG index, and the Dale-Chall formula. When tested by these formulae, the lower the score or grade reading level, the easier the text is to read and understand (and vice versa). Table 5.1 summarises the results for these formulae for the three Qur'ān translations and across three types of Qur'ānic chapters (short, medium, and long). Abdel-Haleem's translation was found to have a consistently low grade level for each formula, across the short, medium, and long chapters, whilst Arberry's was consistently high for each formula across these three types chapters, and Ali's was of a medium level.

Table 5.1: Descriptive Statistics for Reading Ease of Five CRFs in Three Translations

		Qur'ānic Chapters			
Formula	Translator	Short (n.13)	Medium (n.13)	Long (n.14)	All Chapters (n.40)
Fog Index	Abdel-H.	4.61 (1.64)	5.95 (2.05)	7.32 (1.84)	5.96 (2.12)
	Arberry	10.4 (2.58)	11.5 (2.66)	11.5 (1.18)	11.0 (2.25)
	Ali	8.00 (2.20)	6.22 (1.43)	7.05 (1.00)	7.07 (1.74)
Flesch-Kincaid Grade Level	Abdel-H.	3.01 (1.45)	3.44 (1.91)	4.82 (1.79)	3.44 (2.03)
	Arberry	7.91 (2.32)	8.91 (2.65)	8.28 (1.25)	8.38 (2.16)
	Ali	6.10 (2.08)	3.74 (1.77)	4.96 (1.26)	4.90 (1.96
Coleman-Liau Index	Abdel-H.	3.62 (1.71)	4.71 (1.64)	5.15 (1.14)	4.50 (1.62)
	Arberry	6.08 (1.50)	5.36 (1.34)	5.92 (1.04)	5.78 (1.31)
	Ali	6.69 (1.00)	5.93 (1.07)	6.23 (0.73)	6.28 (1.96)
SMOG Index	Abdel-H.	3.74 (1.18)	4.61 (1.33)	5.34 (1.21)	4.56 (1.38)
	Arberry	6.41 (1.77)	7.04 (1.21)	7.16 (1.00)	6.88 (1.37)
	Ali	6.29 (1.24)	4.95 (1.14)	5.49 (0.63)	5.56 (1.16)
Dale-Chall Formula	Abdel-H.	5.51 (0.44)	5.81 (0.39)	5.83 (0.32)	5.72 (0.41)
	Arberry	6.62 (0.53)	6.74 (0.61)	6.48 (0.22)	6.62 (0.49)
	Ali	6.78 (0.53)	6.26 (0.51)	6.41 (0.27)	6.47 (0.49)

Note: Standard deviations are in parentheses.

The table also displays the overall mean for all chapters as one group in each formula. Abdel-Haleem's translation received the lowest grade-level score for each formula, Arberry's received the highest for each formula, and Ali's fell between the two. A one-way ANOVA with the Bonferroni post hoc test was conducted to determine how these translations differed from one another for the five CRFs. There was a homogeneity of variances for all five, as assessed by the Levene's test of homogeneity of variance ($p > .05$). All five formulae were normally distributed for each translator, as shown by the Shapiro-Wilk test ($p > .05$), except two of Abdel-Haleem's scores in the fog index and the Flesch-Kincaid grade level, which did not indicate extreme outliers, as assessed by boxplot. The ANOVA revealed statistically significant differences between the grade levels for the three translations as assessed by the fog index (F

(2, 78) = 81.105, p < .001); Flesch-Kincaid grade level (F (2, 78) = 75.644, p < .001); Coleman-Liau index (F (2, 78) = 37.170, p < .001); SMOG index (F (2, 78) = 57.869, p < .001); and Dale-Chall formula (F (2, 78) = 5.199, p < .001).

Table 5.2: Bonferroni Results for Reading Ease of Five CRFs in Three Translations

Formula	Pairwise Comparisons		Mean difference
Fog Index	Abdel-Haleem	Arberry	5.07 **
	Abdel-Haleem	Ali	1.11 *
	Arberry	Ali	3.96 **
Flesch-Kincaid Grade Level	Abdel-Haleem	Arberry	4.94 **
	Abdel-Haleem	Ali	1.46 **
	Arberry	Ali	3.47 **
Coleman-Liau Index	Abdel-Haleem	Arberry	1.27 **
	Abdel-Haleem	Ali	1.77 **
	Arberry	Ali	0.50 *
SMOG Index	Abdel-Haleem	Arberry	2.31 **
	Abdel-Haleem	Ali	0.99 **
	Arberry	Ali	1.31 **
Dale-Chall Formula	Abdel-Haleem	Arberry	0.90 **
	Abdel-Haleem	Ali	0.75 **
	Arberry	Ali	0.14

Note: ** p < .001; * p < .05

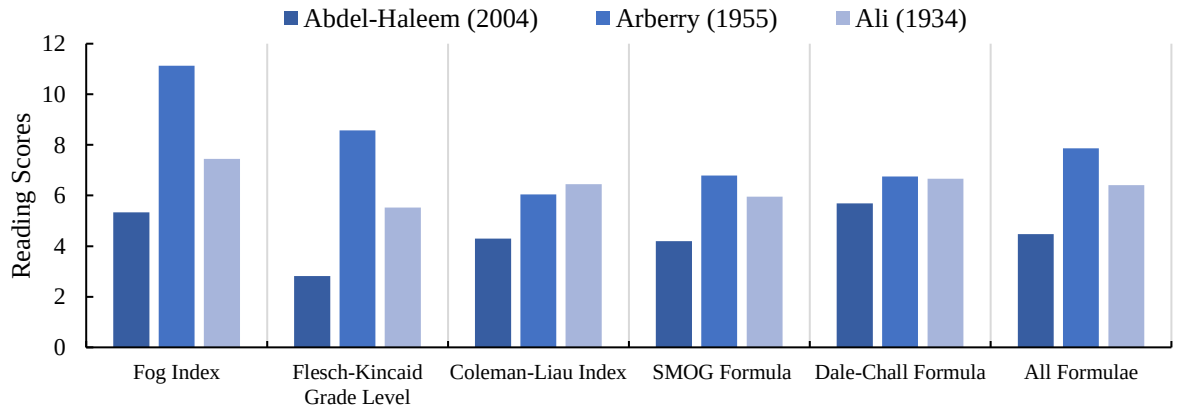


Figure 5.1: Differences in Reading Ease of Three Translations, as Measured by CRFs

The Bonferroni analysis revealed a significant change in grade level in all formula from Abdel-Haleem to Arberry and to Ali (see Table 5.2). There was a significant change in all grade levels from Ali to Arberry, except for the Coleman-Liau index and the Dale-Chall formula, which were not statistically significant. The bar chart above illustrates the highest and lowest grade levels for each translation. According to these formulae, the most readable English Qur'ān translation of the three was Abdel-Haleem's, indicating simpler words and sentences, as explored in the following sections.

5.2.2 Analysis of Lexical Difficulty

The MANOVA results suggest that there are differences between the three Qur'ān translations in their respective combinations of the six lexical variables. The first three variables in Table 5.3 are measures of lexical-semantic features, whilst the later three are measures of lexical-psycholinguistic features. The findings indicate that, in general, Abdel-Haleem's translation was consistently low in lexical difficulty for each variable across the three types of Qur'ānic chapters (short, medium, and long), whilst the lexical choices of Arberry's translation were of a consistently 'medium' lexical difficulty across the three types, whilst Ali's translation was high in lexical difficulty.

Table 5.3: Descriptive Statistics for Variables of Lexical Difficulty in Three Translations

		Qur'ānic Chapters			
Variable	Translator	Short (n.13)	Medium (n.13)	Long (n.14)	All Chapters (n.40)
Word Length	Abdel-H.	1.26 (0.06)	1.29 (0.04)	1.30 (0.04)	1.28 (0.05)
	Arberry	1.29 (0.07)	1.33 (0.06)	1.33 (0.05)	1.31 (0.06)
	Ali	1.33 (0.05)	1.32 (0.05)	1.33 (0.03)	1.33 (0.07)
Hypernym	Abdel-H.	1.44 (1.12)	1.39 (0.14)	1.40 (0.09)	1.41 (0.10)
	Arberry	1.22 (1.16)	1.32 (0.09)	1.30 (0.12)	1.28 (0.14)
	Ali	1.24 (1.18)	1.28 (0.09)	1.29 (0.06)	1.27 (0.12)
Word Frequency	Abdel-H.	1.08 (0.34)	1.13 (0.26)	1.16 (0.16)	1.12 (0.26)
	Arberry	0.81 (0.17)	1.00 (0.21)	1.08 (0.12)	0.90 (0.20)
	Ali	0.87 (0.26)	1.04 (0.33)	1.01 (0.39)	0.97 (0.33)
Word Concreteness	Abdel-H.	368 (27.4)	372 (18.5)	360 (12.3)	366 (20.6)
	Arberry	364 (18.6)	376 (21.7)	359 (15.8)	366 (19.6)
	Ali	357 (21.9)	366 (16.4)	354 (11.9)	359 (17.5)
Word Imageability	Abdel-H.	410 (21.2)	412 (16.5)	402 (12.6)	408 (17.2)
	Arberry	400 (23.8)	411 (17.2)	399 (13.7)	403 (19.0)
	Ali	393 (21.3)	401 (15.4)	391 (10.7)	395 (16.5)
Word Meaningfulness	Abdel-H.	441 (17.3)	435 (9.36)	434 (7.06)	437 (12.1)
	Arberry	431 (12.8)	432 (11.0)	429 (9.24)	431 (10.9)
	Ali	414 (12.5)	421 (11.4)	415 (8.26)	417 (11.2)

Note: Standard deviations are in parentheses.

Table 5.3 displays the overall means of all chapters as one group for each variable and for each translation. To determine more precisely where differences emerge in each of the variables, a one-way ANOVA was conducted, with post hoc Bonferroni analyses (see Table 5.4). There was homogeneity of variance for all six lexical variables, as assessed by the Levene's test of homogeneity of variance ($p > .05$). The variables were normally distributed for each group, as shown by the Shapiro-Wilk test ($p > .05$), except in Arberry's scores for concreteness and imageability, which showed a moderate departure from normality. These were retained in the analysis, as they did not seem to be extreme outliers, as assessed by boxplot. The ANOVA

revealed that the three translations had statistically significant differences in word length ($F(2, 78) = 19.12$, $p < .001$); hypernyms ($F(2, 78) = 28.31$, $p < .001$); word frequency ($F(2, 78) = 5.199$, $p < .001$); meaningfulness ($F(2, 78) = 58.76$, $p < .001$); concreteness ($F(2, 78) = 7.978$, $p < .001$); and imageability ($F(2, 78) = 21.08$, $p < .001$). The results of the Bonferroni analyses for each variable are given below, including the overall means for all chapters as one group.

Word length (syllables per word). Abdel-Haleem's translation ($M = 1.28$, $SD = 0.05$) uses significantly shorter words than Ali's ($M = 1.31$, $SD = 0.06$) and Arberry's ($M = 1.33$, $SD = 0.07$). Arberry's translation uses shorter words than Ali's (see Table 5.4), but the difference was not found to be significant.

Word hypernym. Abdel-Haleem's translation ($M = 1.41$, $SD = 0.10$) scored significantly higher than the other two translators in terms of word hypernym, whilst there was an insignificant difference between Arberry's ($M = 1.28$, $SD = 0.14$) and Ali's ($M = 1.27$, $SD = 0.12$).

CELEX word frequency. The word frequency of Abdel-Haleem's translation ($M = 1.12$, $SD = 0.26$) was found to be significantly higher than either Arberry's ($M = 0.90$, $SD = 0.20$) or Ali's ($M = 0.97$, $SD = 0.33$). There was no difference between the other companions in terms of this variable.

Table 5.4: Bonferroni Results for Variables of Lexical Difficulty in Three Translations

Variable	Pairwise Comparisons		Mean Difference
Word Length (Syllables per Word)	Abdel-Haleem	Arberry	.030**
	Abdel-Haleem	Ali	.044**
	Arberry	Ali	.014
Hypernym	Abdel-Haleem	Arberry	.129**
	Abdel-Haleem	Ali	.140**
	Arberry	Ali	.011
CELEX Word Frequency	Abdel-Haleem	Arberry	.159**
	Abdel-Haleem	Ali	.149*
	Arberry	Ali	.010
Word Concreteness	Abdel-Haleem	Arberry	.209
	Abdel-Haleem	Ali	7.46*
	Arberry	Ali	7.25**
Word Imagability	Abdel-Haleem	Arberry	4.35*
	Abdel-Haleem	Ali	12.6**
	Arberry	Ali	8.28**
Word Meaningfulness	Abdel-Haleem	Arberry	5.88*
	Abdel-Haleem	Ali	19.7**
	Arberry	Ali	13.8**

Note: ** $p < .001$; * $p < .05$

Word concreteness. The word concreteness in Abdel-Haleem's translation ($M = 366$, $SD = 20.6$) was equal to that of Arberry's ($M = 366$, $SD = 19.6$), but significantly higher than that

of Ali's (*M* = 359, *SD* = 17.5). Arberry's used significantly higher word concreteness than Ali's.

Word imageability. Abdel-Haleem's translation (*M* = 408, *SD* = 17.2) had significantly higher scores for word imageability than the other two translations. There was a statistically significant difference between Arberry's translation (*M* = 403, *SD* = 19.0) and Ali's (*M* = 395, *SD* = 16.5).

Word meaningfulness. Abdel-Haleem's translation (*M* = 437, *SD* = 12.1) scored significantly higher than the other two translations for word meaningfulness. Arberry's translation (*M* = 431, *SD* = 10.9) had significantly higher word meaningfulness than Ali's (*M* = 417, *SD* = 11.2).

5.2.3 Analysis of Syntactic Complexity

Table 5.5 presents descriptive statistics for the measures of syntactic complexity in each translation and across the three types of chapters (short, medium, and long). In general, Abdel-Haleem's translation achieved consistently low scores for all measures of syntactic complexity across the three chapters. Arberry's and Ali's translations, in contrast, scored consistently higher in all measures of syntactic complexity across the three chapters. These results suggest that Abdel-Haleem's translation is less likely to use complex sentences than the other texts.

Table 5.5: Descriptive Statistics for Variables of Syntactic Complexity in Three Translations

Measure	Translator	Qur'ānic Chapters			All Chapters
		Short (n.13)	Medium (n.13)	Long (14)	(n.40)
Sentence Length	Abdel-H.	4.36 (2.84)	6.87 (3.5)	10.2 (2.69)	7.03 (4.04)
	Arberry	4.46 (2.90)	6.52 (4.2)	10.4 (3.01)	7.25 (4.08)
	Ali	5.14 (3.12)	8.19 (4.7)	12.4 (4.00)	8.59 (4.94)
Words Before the Main Verb	Abdel-H.	1.86 (21.2)	2.10 (1.2)	3.13 (1.07)	2.36 (1.23)
	Arberry	2.02 (1.24)	2.88 (1.6)	4.21 (1.46)	3.03 (1.69)
	Ali	2.38 (1.78)	2.85 (1.5)	3.48 (1.00)	2.90 (1.53)
Modifiers per Noun Phrase	Abdel-H.	0.58 (0.08)	0.51 (.05)	0.51 (0.06)	0.53 (0.07)
	Arberry	0.64 (0.11)	0.56 (.14)	0.53 (0.07)	0.57 (0.12)
	Ali	0.65 (0.11)	0.56 (.09)	0.54 (0.06)	0.58 (0.10)
Noun-Phrase Density	Abdel-H.	218 (18.2)	206 (17)	201 (17.8)	208 (18.7)
	Arberry	247 (30.7)	221 (40)	202 (19.0)	223 (35.8)
	Ali	271 (29.4)	240 (17)	229 (18.8)	246 (28.3)
Verb-Phrase Density	Abdel-H.	135 (26.8)	129 (18)	130 (12.4)	131 (19.6)
	Arberry	146 (26.2)	132 (19)	125 (7.74)	134 (20.9)
	Ali	123 (22.9)	126 (15)	125 (12.2)	124 (16.8)
Adverbial-Phrase Density	Abdel-H.	23.8 (6.76)	32.0 (7.7)	36.3 (17.4)	30.6 (12.5)
	Arberry	32.8 (9.48)	36.5 (9.4)	37.8 (18.2)	35.7 (12.8)
	Ali	27.5 (4.46)	32.7 (11)	32.6 (13.7)	31.0 (10.6)

Note: Standard deviations are in parentheses.

The descriptive table also shows the overall mean of all chapters as one group for each variable. An ANOVA was conducted after the MANOVA to examine whether there were statistically significant differences in the measures of syntactic complexity between the three Qur'ānic translations. There was homogeneity of variance for all six lexical variables, as shown by the Levene's test of homogeneity of variance ($p > .05$). These variables were normally distributed for each translation, as shown by the Shapiro-Wilk test ($p > .05$), apart from those for sentence length and 'words before main verb'. The sentence lengths used by the translators are dissimilar, which led to departure from normality. In inspecting this variable (i.e., the numbers of words before the main verb) using a boxplot, an outliner was spotted in the scores for Arberry's translation, though it was not an extreme outliner. Due to the large number of samples (40 for each translator), the mean values were used to perform parametric tests.

The ANOVA revealed significant differences between the three translations in terms of sentence length ($F (2, 78) = 15.02$, $p < .001$); number of words before the main verb ($F (2, 78) = 7781$, $p < .001$); number of modifiers per noun phrase ($F (2, 78) = 5.336$, $p < .001$); noun-phrase density ($F (2, 78) = 34.47$, $p < .001$); verb-phrase density ($F (2, 78) = 5.347$, $p < .001$); and adverbial-phrase density ($F (2, 78) = 5.081$, $p < .001$). A follow-up using the Bonferroni post hoc test was conducted after the ANOVAs to determine how these translations differ from each other in terms of these variables, as shown in Table 5.6. The result of the Bonferroni analyses for each variable, with the overall mean for all chapters as one group, is presented below.

Sentence length. Abdel-Haleem's translation ($M = 7.03$, $SD = 4.04$) and Arberry's ($M = 7.25$, $SD = 4.08$) used significantly shorter sentences than Ali's ($M = 8.59$, $SD = 4.94$). Abdel-Haleem's used shorter sentences than Arberry's, but the difference was not significant (see Table 5.6).

Words before main verb. The 'number of words before the main verb' in Abdel-Haleem's translation ($M = 2.36$, $SD = 1.23$) was significantly lower than that of either Arberry's ($M = 3.03$, $SD = 1.69$) or Ali's ($M = 2.90$, $SD = 1.53$). There was no significant difference between the other companions for this variable.

Number of modifiers per noun phrase. Abdel-Haleem's translation ($M = 0.53$, $SD = 0.07$) scored lower for this variable than either Arberry's ($M = 0.57$, $SD = 0.12$) or Ali's ($M = 0.58$, $SD = 0.10$). There was no significant difference between Ali's and Arberry's in respect of this variable.

Syntactic pattern density: According to McNamara et al. (2014, p. 72), the complexity of sentences can be identified by 'the density of particular syntactic patterns'. Abdel-Haleem's translation scored significantly lower for noun-phrase density than the other two translations, whilst Arberry's translation scored significantly lower than Ali's. Abdel-Haleem's translation scored significantly lower for adverbial-phrase density than Arberry's, but this variable showed an insignificant change between Abdel-Haleem's and Ali's and between Ali's and Arberry's translations. Finally, Arberry's translation scored higher for verb phrases than Abdel-Haleem's and Ali's, but the differences between the three were not significant.

Table 5.6: Bonferroni Results for Variables of Syntactic Complexity in Three Translations

Variable	Pairwise Comparisons		Mean Difference
Sentence Length	Abdel-Haleem	Arberry	.221
	Abdel-Haleem	Ali	1.55**
	Arberry	Ali	1.33**
Words Before Main Verb	Abdel-Haleem	Arberry	.677*
	Abdel-Haleem	Ali	.546*
	Arberry	Ali	.131
Number of Modifiers per Noun Phrase	Abdel-Haleem	Arberry	.043*
	Abdel-Haleem	Ali	.052*
	Arberry	Ali	.009
Noun-Phrase Density	Abdel-Haleem	Arberry	14.9*
	Abdel-Haleem	Ali	38.0**
	Arberry	Ali	23.0**
Verb-Phrase Density	Abdel-Haleem	Arberry	2.71
	Abdel-Haleem	Ali	6.90
	Arberry	Ali	9.62*
Adverbial-Phrase Density	Abdel-Haleem	Arberry	5.13*
	Abdel-Haleem	Ali	.403
	Arberry	Ali	4.73

Note: ** $p < .001$; * $p < .05$

5.2.4 Cohesive Elements

To better understand the variables of cohesive ties, as assessed by the Coh-Metrix, the results are organised here into their conceptual frameworks: (1) referential cohesion, (2) semantic coherence, and (3) connectives. This study used six variables for referential cohesion, four for semantic cohesion, and seven for connectives. Tables 5.7 and 5.8 are used to interpret the mean differences between these variables for the three Qur'ānic translations, as explained in the following sections.

Table 5.7: Descriptive Statistics for Cohesive Variables in Three Translations

	Translator	Qur'ānic Chapters			All Chapters (n.40)
		Short (n.13)	Medium (n.13)	Long (n.14)	
Referential Cohesion					
Noun Overlap (Adjacent Sentences)	Abdel.	0.11 (0.11)	0.19 (0.15)	0.22 (0.14)	0.17 (0.14)
	Arberry	0.14 (0.13)	0.22 (0.21)	0.29(0.15)	0.22 (0.18)
	Ali	0.18 (0.14)	0.22 (0.16)	0.29 (0.13)	0.23 (0.15)
Argument Overlap (Adjacent Sentences)	Abdel.	0.33 (0.15)	0.42 (0.25)	0.50 (0.19)	0.42 (0.21)
	Arberry	0.30 (0.18)	0.47 (0.27)	0.57 (0.16)	0.45 (0.23)
	Ali	0.33 (0.11)	0.46 (0.22)	0.52 (0.14)	0.43 (0.18)
Stem Overlap (Adjacent Sentences)	Abdel.	0.29 (0.15)	0.28 (0.16)	0.24 (0.14)	0.27 (0.15)
	Arberry	0.46 (0.11)	0.35 (0.22)	0.33 (0.16)	0.38 (0.17)
	Ali	0.58 (0.12)	0.50 (0.09)	0.47 (0.12)	0.52 (0.17)
Noun Overlap (All Sentences)	Abdel.	0.09 (0.11)	0.15 (0.13)	0.21 (0.13)	0.15 (0.13)
	Arberry	0.10 (0.11)	0.20 (0.20)	0.25 (0.15)	0.18 (0.17)
	Ali	0.11 (0.11)	0.17 (0.16)	0.25 (0.13)	0.18 (0.14)
Argument Overlap (All Sentences)	Abdel.	0.23 (0.16)	0.33 (0.22)	0.45 (0.18)	0.34 (0.20)
	Arberry	0.21 (0.11)	0.36 (0.25)	0.48 (0.17)	0.35 (0.22)
	Ali	0.22 (0.11)	0.33 (0.21)	0.44 (0.16)	0.33 (0.19)
Stem Overlap (All Sentences)	Abdel.	0.27 (0.17)	0.24 (0.15)	0.23 (0.13)	0.25 (0.15)
	Arberry	0.39 (0.11)	0.32 (0.25)	0.28 (0.15)	0.33 (0.17)
	Ali	0.50 (0.08)	0.45 (0.11)	0.43 (0.10)	0.46 (0.10)
Semantic Coherence (Latent Semantic Analysis [LSA])					
LSA Conceptual Overlap (Adjacent Sentences)	Abdel.	0.16 (0.06)	0.17 (0.05)	0.19 (0.05)	0.17 (0.05)
	Arberry	0.19 (0.08)	0.22 (0.07)	0.25 (0.04)	0.22 (0.07)
	Ali	0.19 (0.08)	0.22 (0.10)	0.28 (0.08)	0.23 (0.09)
LSA Conceptual Overlap (All Sentences)	Abdel.	0.11 (0.14)	0.12 (0.11)	0.14 (0.07)	0.12 (0.11)
	Arberry	0.07 (0.13)	0.16 (0.13)	0.20 (0.09)	0.14 (0.13)
	Ali	0.14 (0.20)	0.18 (.17)	0.21 (0.10)	0.18 (0.16)
LSA Conceptual Overlap (Adjacent Paragraphs)	Abdel.	0.17 (0.08)	0.23 (0.10)	0.29 (0.10)	0.23 (0.10)
	Arberry	0.20 (0.08)	0.27 (0.11)	0.32 (0.09)	0.27 (0.11)
	Ali	0.20 (0.09)	0.26 (0.15)	0.39 (0.13)	0.28 (0.14)
LSA Givenness of Information	Abdel.	0.28 (0.04)	0.29 (0.04)	0.32 (0.03)	0.29 (0.04)
	Arberry	0.28 (0.04)	0.32 (0.04)	0.35 (0.04)	0.32 (0.05)
	Ali	0.29 (0.04)	0.32 (0.07)	0.36 (0.05)	0.33 (0.06)
Connectives					
Causal Connectives	Abdel.	26.0 (13.4)	28.2 (10.8)	24.8 (5.79)	26.4 (10.4)
	Arberry	23.5 (16.4)	22.4 (7.19)	21.5 (5.37)	22.5 (10.5)
	Ali	31.7 (17.7)	27.0 (9.00)	29.2 (8.15)	29.3 (12.1)
Adversative and Contrastive Connectives	Abdel.	14.4 (6.46)	17.1 (6.12)	15.1 (4.13)	15.6 (5.64)
	Arberry	15.2 (6.79)	18.4 (7.44)	14.8 (4.07)	16.1 (6.32)
	Ali	20.2 (9.00)	18.4 (6.45)	20.9 (5.99)	19.8 (7.14)
Temporal Connectives	Abdel.	15.4 (9.59)	20.6 (9.05)	14.1 (4.30)	16.8 (8.33)
	Arberry	21.0 (14.8)	24.2 (11.9)	20.9 (9.67)	22.1 (12.1)
	Ali	18.6 (10.9)	21.2 (13.1)	19.0 (6.96)	19.7 (10.5)
Temporal Expanded Connectives	Abdel.	16.3 (11.1)	20.8 (8.77)	18.7 (9.28)	18.7 (9.70)
	Arberry	19.8 (13.1)	18.8 (5.99)	17.1 (10.1)	19.6 (9.90)
	Ali	21.8 (13.0)	20.6 (9.04)	21.0 (7.67)	21.2 (9.88)
Logical Connectives	Abdel.	24.7 (12.1)	35.1 (12)	37.8 (5.63)	32.6 (11.8)
	Arberry	33.8 (11.7)	39.6 (8.5)	40.6 (7.95)	38.8 (9.73)
	Ali	42.5 (12.9)	42.6 (13)	49.4 (8.70)	44.8 (11.9)
Additive Connectives	Abdel.	42.8 (14.7)	39.4 (13.4)	44.0 (9.10)	42.0 (12.5)
	Arberry	63.4 (17.9)	57.7 (21.1)	63.1 (9.22)	61.3 (16.7)
	Ali	65.4 (15.3)	54.3 (16.8)	60.3 (7.46)	59.9 (14.3)
All Connectives	Abdel.	85.3 (26.1)	90.5 (16.4)	89.5 (11.7)	88.5 (18.7)
	Arberry	108 (32.6)	107 (21.3)	111 (13.8)	109 (23.2)
	Ali	119 (29.5)	106 (19.1)	115 (9.69)	113 (21.2)

Table 5.8: Bonferroni Results for Cohesive Variables in Three Translations

Variable	Pairwise Comparisons		Mean Difference
Noun Overlap – Adjacent Sentences	Abdel-Haleem	Arberry	.046*
	Abdel-Haleem	Ali	.055*
	Arberry	Ali	.009
Argument Overlap – Adjacent Sentences	Abdel-Haleem	Arberry	.034
	Abdel-Haleem	Ali	.020
	Arberry	Ali	.014
Stem Overlap – Adjacent Sentences	Abdel-Haleem	Arberry	.112**
	Abdel-Haleem	Ali	.248**
	Arberry	Ali	.136**
Noun Overlap – All Sentences	Abdel-Haleem	Arberry	.032*
	Abdel-Haleem	Ali	.028
	Arberry	Ali	.004
Argument Overlap – All Sentences	Abdel-Haleem	Arberry	.011
	Abdel-Haleem	Ali	.008
	Arberry	Ali	.020
Stem Overlap – All Sentences	Abdel-Haleem	Arberry	.083**
	Abdel-Haleem	Ali	.214**
	Arberry	Ali	.131**
All Connectives	Abdel-Haleem	Arberry	20.8**
	Abdel-Haleem	Ali	25.3**
	Arberry	Ali	4.46
Causal Connectives	Abdel-Haleem	Arberry	3.90*
	Abdel-Haleem	Ali	2.85
	Arberry	Ali	6.75**
Adversative and Contrastive Connectives	Abdel-Haleem	Arberry	.443
	Abdel-Haleem	Ali	4.18**
	Arberry	Ali	3.74*
Temporal Connectives	Abdel-Haleem	Arberry	5.28**
	Abdel-Haleem	Ali	2.85*
	Arberry	Ali	2.43
Expanded Temporal Connectives	Abdel-Haleem	Arberry	.030
	Abdel-Haleem	Ali	2.48
	Arberry	Ali	2.51
Additive Connectives	Abdel-Haleem	Arberry	19.2**
	Abdel-Haleem	Ali	17.8**
	Arberry	Ali	1.41
Logical Connectives	Abdel-Haleem	Arberry	5.43**
	Abdel-Haleem	Ali	12.1**
	Arberry	Ali	6.7**
LSA Conceptual Overlap – Adjacent Sentences	Abdel-Haleem	Arberry	.048**
	Abdel-Haleem	Ali	.053**
	Arberry	Ali	.006
LSA Conceptual Overlap – All Sentences	Abdel-Haleem	Arberry	.019
	Abdel-Haleem	Ali	.055*
	Arberry	Ali	.036
LSA Conceptual Overlap – Adjacent Paragraphs	Abdel-Haleem	Arberry	.037*
	Abdel-Haleem	Ali	.054**
	Arberry	Ali	.017
LSA Givenness of Information	Abdel-Haleem	Arberry	.021**
	Abdel-Haleem	Ali	.030**
	Arberry	Ali	.010

Note: ** $p < .001$; * $p < .05$

5.2.4.1 Analysis of Referential Cohesion

Referential cohesion in texts can be identified at the local level (i.e., between adjacent sentences) and at the global level (i.e., between all sentences). Descriptive statistics for the measures of referential cohesion in each translation are presented in Table 5.7 (presented in the previous section). The descriptive table shows that both Ali's and Arberry's translations were generally high in referential cohesion across the three types of chapters (short, medium, and long). In contrast, Abdel-Haleem's translation was low in cohesion across the three types. The comparisons between these Qur'ānic chapters also indicate that the translations of Ali and Arberry were equivalent in the quantities of their repetitive aspects (such as lexical repetition and morphological repetition, as shown in Figure 5.2 below).

An ANOVA was conducted to identify whether there were statistically significant differences between the variables of referential cohesion in the three translations. No extreme outliers were revealed in a visual inspection of a boxplot. There was homogeneity of variance for all variables, as shown by the Levene's test of homogeneity of variance ($p > .05$). The ANOVA revealed that the three translations were significantly different for all variables of referential cohesion, except two ('argument overlap – adjacent sentences' and 'argument overlap – all sentences'), for which the differences were not statistically significant. A follow-up Bonferroni post hoc test was conducted after the ANOVA to determine how these translations differ from each other in these variables. The results of the Bonferroni analysis (see Table 5.8) for each variable, in terms of the overall mean for all chapters as one group, are reported below.

Noun overlap between adjacent sentences. Table 5.7 shows that the total mean for 'noun overlap' of Arberry's translation ($M = 0.22$, $SD = 0.18$) was higher than that of Abdel-Haleem's ($M = 0.17$, $SD = 0.14$), but nearly equal to that of Ali's ($M = 0.23$, $SD = 0.15$). The Bonferroni analysis table (see Table 5.8) shows that Abdel-Haleem's and Arberry's mean scores for this variable were statistically significantly different, as were those of Abdel-Haleem and Ali, whilst those of Arberry and Ali were not.

Stem overlap between adjacent sentences. The total mean for 'stem overlap' in Arberry's translation ($M = 0.38$, $SD = 0.17$) was higher than that of Abdel-Haleem's ($M = 0.27$, $SD = 0.15$) and lower than that of Ali's ($M = 0.52$, $SD = 0.17$). The mean scores for this variable in the Abdel-Haleem's and Arberry's translations were statistically significantly different, as were those of Abdel-Haleem and Ali and those of Arberry and Ali. Figure 5.4 illustrates these differences.

Argument overlap between adjacent sentences. The total mean for 'argument overlap' in Arberry's translation (*M* = 0.45, *SD* = 0.23) was somewhat higher than that of Abdel-Haleem's (*M* = 0.42, *SD* = 0.21), but it was nearly equal to that of Ali's (*M* = 0.43, *SD* = 0.18). This variable was not significantly different for any of the translations.

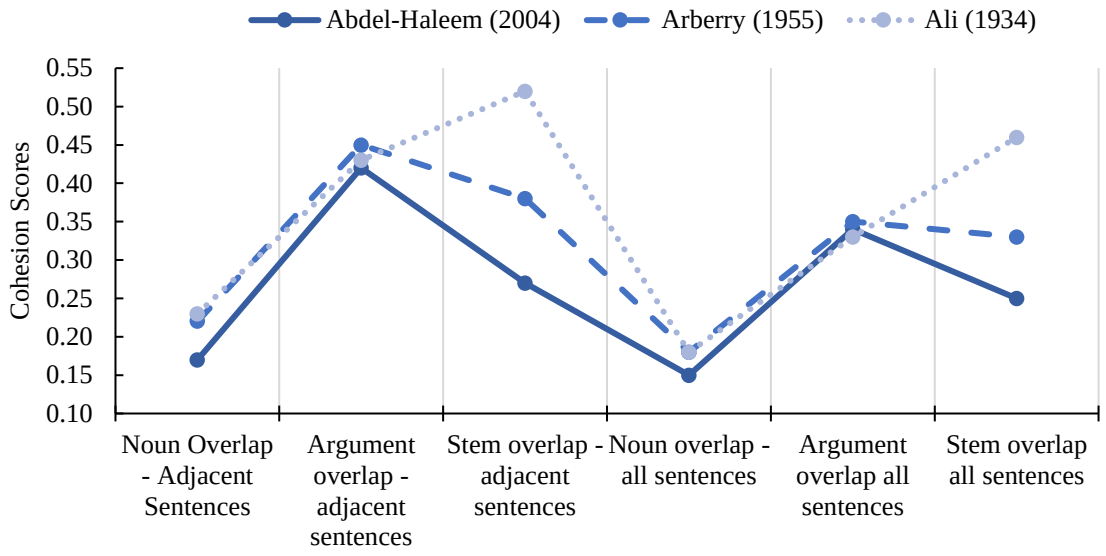


Figure 5.2: Differences for Referential Cohesion in Three Translations

Noun overlap between all sentences. The total mean for 'noun overlap' of Abdel-Haleem's translation (*M* = 0.15, *SD* = 0.13) was significantly lower than that of Ali's (*M* = 0.18, *SD* = 0.14) and of Arberry's (*M* = 0.18, *SD* = 0.17). The means of Arberry's and Ali's translations were not significantly different.

Stem overlap between all sentences. The mean scores for 'stem overlap' were statistically significantly different for Abdel-Haleem's translation (*M* = 0.25, *SD* = 0.15) and Arberry's (*M* = 0.33, *SD* = 0.17), for Abdel-Haleem's and Ali's (*M* = 0.46, *SD* = 0.10), and for Arberry's and Ali's.

Argument overlap between all sentences. The total mean for 'argument overlap' of Abdel-Haleem's translation (*M* = 0.34, *SD* = 0.20) was slightly lower than that of Arberry's (*M* = 0.35, *SD* = 0.13) and slightly higher than that of Ali's (*M* = 0.33, *SD* = 0.19). The scores were not significantly different for any of the translators.

In summary, Ali's translation had the highest scores for most of the local and global cohesion variables, Abdel-Haleem's translation had the lowest values for all variables, and Arberry's translation fell in between (in most variables, Ali and Arberry were equal in terms

of the averages). These results suggest that Ali's and Arberry's translations both use more elements of lexical and morphological repetition than Abdel-Haleem's translation, which might be intended to avoid the need for repetition by using near-synonyms. Another explanation for this finding may be that Ali's and Arberry's translations have similar patterns at most local and global levels of referential cohesion. Such similarity might mean that their translations are more likely to be affected by Qur'ān-specific repetitive patterns. Abdel-Haleem's translation, in contrast, has different patterns. Abdel-Haleem reduced the repetition in his translation by maintaining the norms of the target language.

5.2.4.2 Analysis of Semantic Coherence

Coh-Metrix includes measures from latent semantic analysis (LSA) to quantify the semantic coherence of a text by computing 'the semantic similarities between words, sentences, and paragraphs' (McNamara et al., 2014, p.2). For example, the word 'car' is related to words such as 'road' and 'driver'. This research used four variables to measure semantic similarities or relatedness: (1) LSA conceptual overlap – adjacent sentences, (2) LSA conceptual overlap – all sentences, (3) LSA conceptual overlap – adjacent paragraphs, and (4) LSA givenness of information (see Chapter 4). A high score for these variables indicates greater semantic relatedness across sentences and paragraphs, which is also an indicator of high cohesion (McNamara et al., 2014).

The descriptive statistics for the measures of semantic coherence in each translation are summarised in Table 5.7 (section 5.2.4). The descriptive table shows that Ali's and Arberry's translations scored consistently high for semantic coherence in the short, medium, and long chapters. Abdel-Haleem's translation, in contrast, scored consistently low for semantic coherence across the three types of chapters. An ANOVA was conducted to identify whether there were statistically significant differences between the variables of semantic coherence for the three translators. No extreme outliers amongst the scores were revealed by a visual inspection of a boxplot. There was homogeneity of variance for all variables, which was assessed using the Levene's test of homogeneity of variance ($p > .05$). The ANOVA revealed that the variables for semantic coherence were statistically significant different for all variables in the three translations. A follow-up using the Bonferroni post hoc test was conducted after the ANOVA to determine whether these translations differed from each other in these variables. The results of the Bonferroni analysis (see Table 5.8 in section 5.2.4) for each variable, in terms of the overall mean of all chapters as one group, are reported below.

LSA conceptual overlap – adjacent sentences. The total mean for this variable in Ali's translation (*M* = 0.23, *SD* = 0.09) was almost equal to that of Arberry's (*M* = 0.22, *SD* = 0.07), but significantly higher than that of Abdel-Haleem's (*M* = 0.17, *SD* = 0.05). The scores for this variable for Arberry and Ali were not significantly different. Figure 5.3 illustrates these differences.

LSA conceptual overlap – all sentences. The total mean for this variable in Ali's translation (*M* = 0.18, *SD* = 0.16) was higher than the mean values of Abdel-Haleem's (*M* = 0.12, *SD* = 0.11) and Arberry's (*M* = 0.14, *SD* = 0.13). The scores were only statistically different for Abdel-Haleem and Ali.

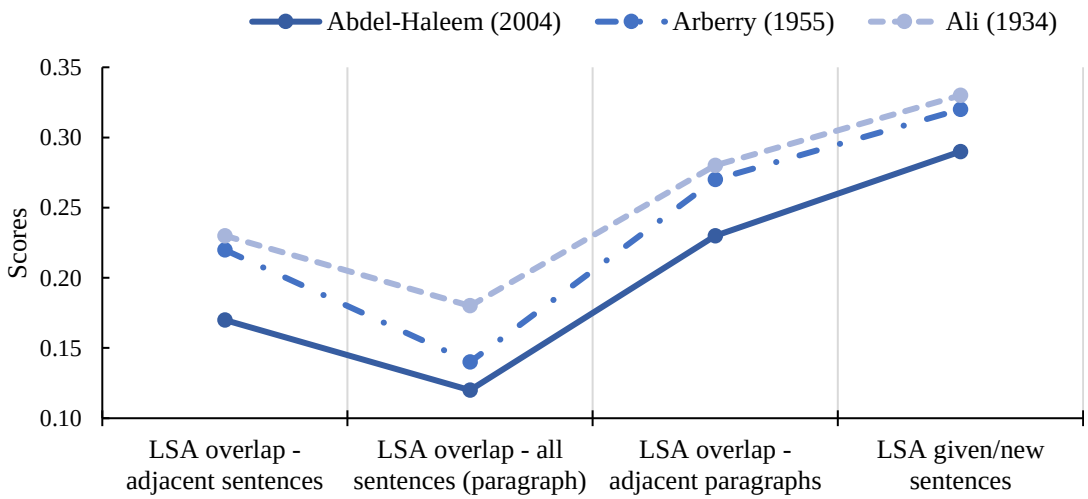


Figure 5.3: Differences for Semantic Coherence in Three Translations

LSA conceptual overlap – adjacent paragraphs. The total mean for this variable in Ali's translation (*M* = 0.28, *SD* = 0.14) was nearly equal to that of Arberry's (*M* = 0.27, *SD* = 0.11), but higher than that of Abdel-Haleem's (*M* = 0.23, *SD* = 0.10). The mean scores for this variable for Abdel-Haleem and Arberry were statistically significantly different, as were those of Abdel-Haleem and Ali, but not those of Arberry and Ali.

LSA givenness of information. The total mean for this variable in Ali's translation (*M* = 0.33, *SD* = 0.06) was nearly equal to that of Arberry's (*M* = 0.32, *SD* = 0.05), but higher than that of Abdel-Haleem's (*M* = 0.29, *SD* = 0.04). The mean scores for this variable in Abdel-Haleem's and Arberry's translations were statistically significantly different, as were those in Abdel-Haleem's and Ali's, but those in Arberry's and Ali's were not.

To summarise, Ali's and Arberry's translations were similar in three variables and scored higher than Abdel-Haleem's translation in all variables. Ali's and Arberry's translations showed similar patterns for the three variables. This similarity might indicate that their translations retain semantic relatedness or contextual similarities with the source text. In contrast, Abdel-Haleem's translation may have reproduced the semantic relatedness of the source text. This reproduction might have emerged due to the use of communicative translation.

5.2.4.3 Analysis of Connectives

This research used seven variables to measure types of connectives in translations: (1) adversative and contrastive, (2) causal, (3) additive, (4) temporal, (5) expanded temporal, (6) logical, and (7) all connectives. A high score for these variables indicates greater linking across sentences. The descriptive statistics for the measures of connectives in each translation are reported in Table 5.7 (section 5.2.4). The descriptive table shows that Ali's and Arberry's translations scored high for their use of connectives across the short, medium, and long chapters. Abdel-Haleem's translation scored low for its use of connectives across the three types of chapters. The results of this section do not contradict the results for syntactic complexity. Rather, they show that Abdel-Haleem's translation includes fewer complex sentences, resulting in fewer connectives.

An ANOVA was conducted to identify whether there were statistically significant differences between the variables of connectives for the three translators. No extreme outliers were revealed amongst the scores by a visual inspection of a boxplot. There was homogeneity of variance for all variables, which was assessed by the Levene's test of homogeneity of variance ($p > .05$). The ANOVA revealed that the three translations showed significant differences for all variables of connectives. A follow-up using the Bonferroni post hoc test was conducted after the ANOVA to identify whether the translations differed from each other in these variables. The results of the Bonferroni analysis (see Table 5.8 in section 5.2.4) for each variable, in terms of the overall mean for all chapters as one group, are reported below.

All connectives. The total mean for 'all connectives' in Ali's translation ($M = 113$, $SD = 21$) was nearly equal to that of Arberry's ($M = 109$, $SD = 23$), but higher than that of Abdel-Haleem's ($M = 88$, $SD = 88$). As shown in the Bonferroni table, the mean scores for this variable were statistically significantly different for Abdel-Haleem and Arberry, and for Abdel-Haleem and Ali, but not for Arberry and Ali.

Adversative and contrastive connectives. The total mean for 'adversative and contrastive connectives' in Ali's translation (*M* = 19, *SD* = 7) was higher than the means for Arberry's (*M* = 16, *SD* = 7) and Abdel-Haleem's (*M* = 15, *SD* = 5). As shown in the Bonferroni table, the mean scores for this variable were only statistically different for Abdel-Haleem and Ali.

Causal connectives. The total mean for 'causal connectives' in Ali's translation (*M* = 29, *SD* = 12) was higher than the means of Abdel-Haleem's (*M* = 26, *SD* = 10) and Arberry's translations (*M* = 22, *SD* = 10). The variables were statistically different for Abdel-Haleem and Arberry, and for Ali and Arberry, but not for Ali and Abdel-Haleem (see Table 5.8).

Additive connectives. The total mean for 'additive connectives' in Arberry's translation (*M* = 61, *SD* = 16) was almost equal to that of Ali's (*M* = 59, *SD* = 14) but higher than that of Abdel-Haleem's (*M* = 42, *SD* = 12). The mean scores for Abdel-Haleem's and Arberry's translations were statistically significantly different, as were those of Abdel-Haleem and Ali, but Arberry's and Ali's were not.

Temporal connectives: The total mean for 'temporal connectives' in Arberry's translation (*M* = 22, *SD* = 12) was nearly equal to that of Ali's (*M* = 19, *SD* = 10) but higher than that of Abdel-Haleem's (*M* = 16, *SD* = 8). There were statistically significantly differences between the mean scores of Abdel-Haleem and Arberry, and between those of Abdel-Haleem and Ali, but not between those of Arberry and Ali.

Expanded temporal connectives: The total mean for 'expanded temporal connectives' in Ali's translation (*M* = 21, *SD* = 9) was higher than the mean values of both Abdel-Haleem's (*M* = 18, *SD* = 10) and Arberry's (*M* = 19, *SD* = 9). There were no statistically significant differences between any of the translations.

Logical Connectives (*and, or, if, then*): High densities of logical operators (e.g., 'and', 'or', 'if', 'then') between sentences indicate greater complexity of a text and place a high load on working memory (Graesser et al., 2004; 2006). Abdel-Haleem's translation (*M* = 32, *SD* = 11.8) scored lower for density of logical operators than either Arberry's (*M* = 38, *SD* = 9.73) or Ali's (*M*= 44, *SD* = 11.09). Abdel-Haleem's translation scored significantly lower for this variable than the other two translations, whilst Ali's scored significantly higher than Arberry's.

The total frequency of all logical operators across the 40 translated chapters of Abdel-Haleem's text was found to be low (except for the logical operator 'if', when compared to Arberry's translation), as shown in Figure 5.4. In contrast, the total frequencies of all logical operators in both Arberry's and Ali's translations were significantly higher than in Abdel-Haleem's. The operator 'and' (و) is a major cohesive marker in the Qur'ān (Abdul-Raof, 2001;

2019). This operator is very rare in Abdel-Haleem's translation and very high in both Arberry's and Ali's translations. This result indicates that Abdel-Haleem may have chosen to avoid the marker to preserve the norms of the target text, whilst the other two translators chose to maintain it to preserve the norms of the source text.

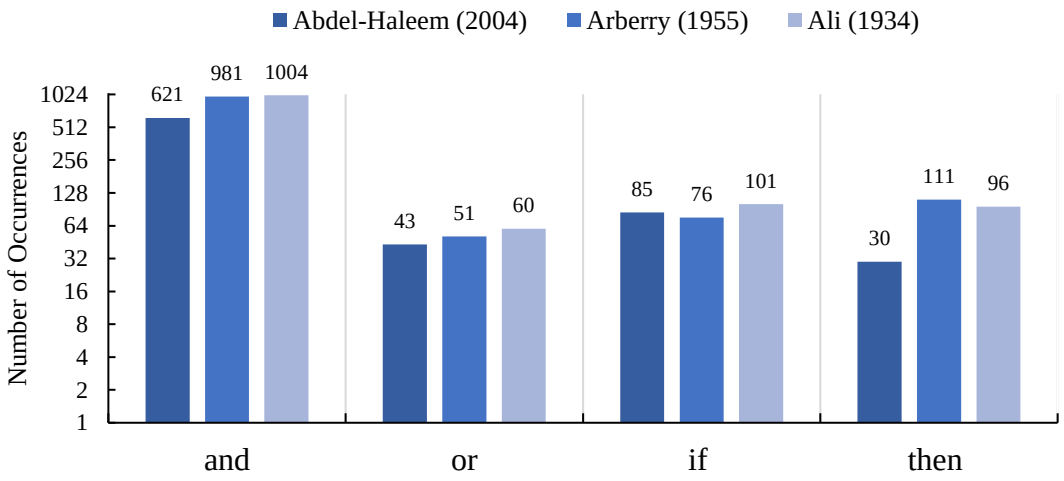


Figure 5.4: Frequency Counts for Logical Operators in Three Translations

5.2.5 Analysis of Narrativity

Descriptive statistics for the measure of narrativity in each translation are presented in Table 5.9. Abdel-Haleem's translation scored the highest for narrativity across the three types of Qur'ānic chapters, whilst Ali's translation scored the lowest and Arberry's fell between the two. The descriptive table also shows the overall mean for narrativity for all chapters as one group.

Table 5.9: Descriptive Statistics for Narrativity in Three Translations

		Qur'ānic Chapters			
		Short (n.13)	Medium (n.13)	Long (n.14)	All Chapters (n.40)
Narrativity	Abdel-Haleem	67 (21.0)	85 (6.59)	84 (9.2)	78 (15.6)
	Arberry	62 (24.2)	81 (12.8)	82 (10)	75 (18.6)
	Ali	57 (21.4)	76 (14.1)	80 (7.6)	71 (18.1)

Note: Standard deviations in parentheses.

An ANOVA and follow-up using the Bonferroni analysis were conducted. There was homogeneity of variance for all variables, as shown by the Levene's test of homogeneity of variance ($p > .05$). The scores for narrativity indicated a moderate departure from normality. This is because the themes of the short Qur'ānic chapters are more likely to be descriptive (e.g.,

the creation of the universe such as the moon, stars, sun, earth, people, heaven, etc.). The short chapters thus had lower values for narrativity, as the themes of the medium and long Qur'ānic chapters are different. The ANOVA revealed a statistically significant difference in narrativity for the three translators. The Bonferroni analysis gave Ali's translation a significantly lower score for narrativity than Abdel-Haleem's, with Abdel-Haleem's scoring higher than Arberry's. Arberry's score was also significantly higher than Ali's.

5.3 Part Two: Factor Analysis of Textual Dimensions

The 29 selected linguistic measures, taken from Coh-Metrix and used in the previous sections, operate at the word, sentence, and cohesion levels. This section explores the co-occurrence patterns of these selected linguistic measures in the linguistic characteristics of the selected English Qur'ān translations. As discussed in Chapter 4, this study follows previous investigations by using a multi-dimensional analysis to identify linguistic variation in spoken and written registers.

Before the textual variation analysis commenced, the three Qur'ānic corpora used in this chapter were combined. There are 19,823 words in Abdel-Haleem's corpus, 19,605 in Arberry's, and 22,555 words in Ali's (see Chapter 4). The combined corpus comprises 120 Qur'ānic chapters and 61,983 words. Biber (1990, pp. 261-262) states that '10-text samples are large enough to reliably represent a genre. It seems safe to conclude that the 2,000-word and 5,000-word texts in the standard corpora are reliable representatives of their respective text categories for analyses of this type'. These combined Qur'ānic samples are thus adequate to represent the distributions and patterns of linguistic features in the selected English Qur'ān translations.

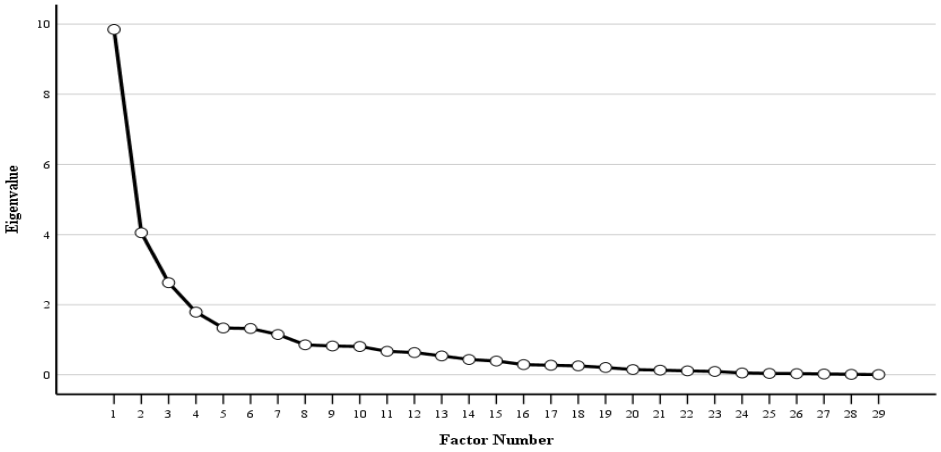
Table 5.10: First 10 Eigenvalues from the Principal Component Analysis

Factor Number	Eigenvalues	Percent of Variance	Cumulative Variance
1	9.847	33.9 %	33.957
2	4.052	13.9 %	47.929
3	2.628	9.6 %	56.99
4	1.789	6.1 %	63.159
5	1.335	4.6 %	67.762
6	1.321	4.5 %	72.319
7	1.151	3.9 %	76.288
8	0.857	2.9 %	79.243
9	0.823	2.8 %	82.08
10	0.809	2.7 %	84.87

A principal component analysis (PCA) was conducted in SPSS to identify the co-occurring patterns of these selected linguistic features in the translations. Before the PCA was performed, the suitability of a PCA was assessed. There are two assumptions: linearity or correlation between all variables, which can be detected from a correlation matrix, and sampling adequacy, which can be assessed using the Kaiser-Meyer-Olkin (KMO) and Bartlett's test of sphericity. First, an inspection of the correlation matrix showed that almost all the variables had at least one correlation coefficient greater than 0.3, except for 'verb-phrase density', which was less than 0.3 and was thus removed from the dataset. Second, the overall KMO measure was 0.74, which is a high and acceptable value, according to the Kaiser (1974, cited in Laerd Statistics, 2015) classification of values. Bartlett's test of sphericity was statistically significant ($p < .0005$), indicating that the variables in the dataset were correlated and were likely factorisable.

Since our data were deemed appropriate for the factor analysis, we needed to decide how many factors should be retained in the analysis. The first 10 eigenvalues for the factor analysis are reported in Table 5.10. As shown in the table below, Factor 1 accounted for 33.9% of the shared variance, whilst Factors 2 and 3 explained 13.9% and 9.6%, respectively. According to Biber (1988, p.82), a scree plot can also be used to specify 'the optimal number of factors' that must be retained in the analysis.

Figure 5.5: Scree Plot of the Principal Component Analysis



As shown in Figure 5.5, the clearest break in the scree plot occurs between the fifth and sixth factors, and another break occurs between the eighth and ninth factors. As Biber (1988, p.82) writes 'when faced with a choice between a larger or smaller number of factors, the more conservative procedure is to extract the larger number and then discard any unnecessary

factors'. Visual inspection of the scree plot reveals that five components should be retained, indicating that a five-factor solution is needed for a meaningful interpretation, as other factors have fewer than three variables, which is not sufficient for a full interpretation of the variables. In total, the five-factor solution explains 67.7% of the shared variance. The PCA revealed the five factors that had eigenvalues greater than one and explained 33.9%, 13.9%, 9.6%, 6.1%, and 4.6% of the total variance.

As discussed in Chapter 4, a Promax rotation was employed to clarify the description of the co-occurring patterns on each factor. The rotated solution exhibits a 'simplified structure' that 'facilitates the interpretation of the constructs underlying each factor' (Biber, 1988, p. 84), as shown in Table 5.11. Each of the five factors has a set of variables and the numbers 'in front of the linguistic features on each factor are referred to as factor "loadings" or "weights"' (Biber, 1988, p.80). Loadings that had an absolute value of less than .35 were removed from the analysis (Louwerse et al., 2004; Biber, 1988). This meant that only the salient and important linguistic features or indices were retained in the analysis.

Table 5.11: Summary of the Co-Occurring Linguistic Features for Each Factor

Factor 1	
Linguistic Features	**Loadings**
Noun overlap – all sentences	.945
Argument overlap – all sentences	.923
LSA givenness of information	.903
Noun overlap – adjacent sentences	.902
LSA conceptual overlap – paragraphs	.899
Sentence length	.897
Argument overlap – adjacent sentences	.861
LSA conceptual overlap – adjacent sentences	.829
Number of words before main verb	.805
LSA conceptual overlap – all sentences	.714
Word length	.550

Factor 2	
Linguistic Features	**Loadings**
Imageability for content words	.854
Meaningfulness for content words	.810
Concreteness for content words	.792
Word hypernymy	.571

Factor 3	
Linguistic Features	**Loadings**
Noun-phrase density	.892
Stem overlap – adjacent sentences	.812
Stem overlap – all sentences	.775
Modifiers per noun phrase	.651
Word length	.387

Factor 4	
Linguistic Features	**Loadings**
All connectives	.883
Additive connectives	.749
Causal connectives	.678
Adversative connectives	.643
Logical connectives	.413

Factor 5	
Linguistic Features	**Loadings**
Temporal connectives	.642
Expanded temporal connectives	.634
Adverbial-phrase density	.505
Word Frequency	.391

Key: latent semantic analysis (LSA).

Table 5.11 summarises the co-occurring linguistic features for each factor in the English Qur'ān translations. The co-occurring features underlying each of the five factors are interpreted based on the shared function of the variables and findings of previous studies, as discussed in the following sections. Each factor displayed at least five salient loadings, with

the exception of Factors 2 and 5, which displayed four. Each factor identified in the analysis has a shared group of linguistic features that frequently co-occur in the translations. As Biber (1988, p.13) writes 'when a group of features consistently co-occur in texts, those features define a linguistic dimension'. After interpreting the co-occurring of these features together on each factor, a textual dimension that contributes to the discourse complexity of Qur'ān translations is suggested at the end of each factor.

5.3.1 Interpretation of Factor 1: Density of Qur'ān-Specific Cohesive Ties

Factor 1 comprises 11 linguistic features that can be clustered into their conceptual frameworks, as listed below. Factor 1 explains 33.9% of the total variance. This is the largest of the five factors and very powerful, representing a significant feature of the textual variation that influences the reading difficulty of the translations. Ten of the features have positive weights of larger than .700 for this factor; word length is the exception, which has a positive weight of .550 and remains a large loading (see Table 5.11 in the previous section).

a) Four measures of referential cohesion (i.e., noun overlap – all sentences, argument overlap – all sentences, noun overlap – adjacent sentences, and argument overlap – adjacent sentences)
b) Four measures of semantic coherence (i.e., LSA conceptual overlap – adjacent sentences, LSA conceptual overlap – all sentences, LSA conceptual overlap – paragraphs, LSA givenness of information)
c) Syntactic complexity (i.e., number of words before main verb)
d) Grade-level difficulty (i.e., word length, sentence length)

All 11 features co-occur frequently in the translations. Most of these co-occurring features (i.e., the measures of both referential cohesion and semantic coherence) usually indicate complex linking structures in a text. Tannen (1982, p.3) points out that 'cohesion is established in writing through lexicalization and complex syntactic structures which make connectives explicit'. The measures of both referential cohesion and semantic coherence are the highest loading features for this factor, with weights of larger than .800 (see Table 5.11), representing a salient complex structure in some verses of the translations that are lexically overlapping and semantically coherent. The co-occurring features of referential cohesion are discussed here first.

The measures of referential cohesion, taken from Coh-Metrix, represent different types of lexical cohesion: noun overlap (e.g., table/table), argument overlap (e.g., he/he), and stem overlap (e.g., forgive/forgiving) in relation to local sentences (adjacent) and global sentences

(all sentences in a text; McNamara et al., 2014). Examples 1–3 below illustrate noun overlap that involves lexical repetition between adjacent verses.

1. (الْقَارِعَةُ. مَا الْقَارِعَةُ. وَمَا أَدْرَاكَ مَا الْقَارِعَةُ)

[The **Crashing Blow**! What is the **Crashing Blow**? What will explain to you what the **Crashing Blow** is? Q101: 1-3] (Abdel-Haleem, 2004).

2. (وَالسَّمَاءِ وَالطَّارِقِ. وَمَا أَدْرَاكَ مَا الطَّارِقُ)

[By heaven and the **night-star**. What shall teach thee what is the **night-star**? Q86: 1-2] (Arberry, 1955).

3. (الْحَاقَّةُ. مَا الْحَاقَّةُ. وَمَا أَدْرَاكَ مَا الْحَاقَّةُ)

[The **sure reality**! What is the **sure reality**? And what will make thee realise what the **sure reality** is? Q69:1-3] (Ali, 1934).

In Qur'ānic discourse, lexical cohesion is achieved through the repetition of lexical items (as shown in the above Arabic and English examples); and repetition is used to achieve the function of reiteration and 'continuity of meaning' between verses (Abdul-Raof, 2019). This stylistic feature – the lexical repetition of the word – was identified quantitatively by a factor analysis and found to co-occur frequently in the translations. As shown in the above examples, these three translators maintain this repetition in their translations, which is a style specific to Arabic and largely absent from English, which usually prefers lexical variation (see Johnstone, 1990; Dickins et al., 2017).

4. (أَفَلَا يَتُوبُونَ إِلَى اللَّهِ وَيَسْتَغْفِرُونَهُ ۚ وَاللَّهُ غَفُورٌ رَّحِيمٌ)

A. [Why turn they not to **Allah**, and seek His **forgiveness**? For **Allah** is Oft-**forgiving**, Most Merciful Q5:74] (Ali, 1934).

B. [Why do they not turn to **God** and ask His **forgiveness**, when **God** is most **forgiving**, most merciful? Q5:74] (Abdel-Haleem, 2004).

C. [Will they not turn to **God** and pray His **forgiveness**? **God** is All-**forgiving**, All-compassionate Q5:74] (Arberry, 1955).

Example 4 provides another illustration of both noun overlap and stem overlap. The co-occurring of a lexical item through its morphological root is another typical feature of Arabic Qur'ānic discourse, which is used to achieve 'continuity of thought' at the verse level (Abdul-Raof, 2019; 2001). This root repetition is identified quantitatively by a factor analysis and has been retained in these translations, as shown in the above examples.

On the other hand, four features of semantic coherence have a positive weight for Factor 1. The measures of semantic coherence describe semantic relatedness or contextual similarity between words, sentences, and paragraphs that have similarity in meaning (e.g., the word ‘hammer’ is associated with words such as ‘saw’, ‘nails’, and ‘constructing’; McNamara et al., 2014). Examples 5–7 below illustrate semantic similarity between words that are strongly associated with words from the same contextual function.

5. (أَفَلَمْ يَسِيرُوا فِي الْأَرْضِ فَتَكُونَ لَهُمْ قُلُوبٌ يَعْقِلُونَ بِهَا أَوْ آذَانٌ يَسْمَعُونَ بِهَا ۖ فَإِنَّهَا لَا تَعْمَى الْأَبْصَارُ وَلَٰكِن تَعْمَى الْقُلُوبُ الَّتِي فِي الصُّدُورِ)

[What, have they not journeyed in the land so that they have **hearts** to understand with or **ears** to hear with? It is not the **eyes** that are blind, but blind are the **hearts** within the **breasts**. Q22:46] (Arberry, 1955).

6. (وَمَا مِن دَابَّةٍ فِي الْأَرْضِ وَلَا طَائِرٍ يَطِيرُ بِجَنَاحَيْهِ إِلَّا أُمَمٌ أَمْثَالُكُم)

[All the **creatures** that **crawl** on the earth and those that **fly** with their **wings** are communities like yourselves. Q6:38] (Abdel-Haleem, 2004).

7. (وَإِلَى السَّمَاءِ كَيْفَ رُفِعَتْ وَإِلَى الْجِبَالِ كَيْفَ نُصِبَتْ وَإِلَى الأرْضِ كَيْفَ سُطِحَتْ)

[Do they not look at the Camels, how they are made? And at the **sky**, how it is raised high? And at the **mountains**, how they are fixed firm. And at the **earth**, how it is spread out? Q88:17-20] (Ali, 1934).

As shown in the above examples, the three translators have maintained the semantic similarity of the ST in their translations. The frequent co-occurrence of semantic similarity in the English translations might indicate that coherence is produced in verses through conceptual similarity, with a Qur’ānic argumentation technique used to represent God’s omnipotence, His signs, deductive reasoning, His creation, and proof of divinity (see Abdul-Raof, 2019). Example 7 describes God’s omnipotence with reference to His creation of the sky, mountains, and earth, which are semantically related words.

Some other positive features involving this factor (i.e., word length, sentence length, and number of words before the main verb) are often associated with discourse complexity and syntactic complexity (see McNamara et al., 2014; Biber, 1992). Sentences with more words before the main verb are more syntactically complex, more likely to be ambiguous, and therefore more difficult to comprehend (McNamara et al., 2014). The syntactic structure of the English translation often relies heavily on the complex structure of Qur’ān-specific cohesive ties between lexical cohesion and semantic relatedness to specify more explicitly the relationships between the meanings of verses. Such a structure usually expands the unit of the

verse to convey additional information, which results in more complex syntactic structures, longer words, longer sentences, and thus greater reading difficulty in Qur'ān translations.

Overall, these co-occurring features of referential cohesion and semantic coherence can show some complexities of Qur'ān-specific cohesive devices in English Qur'ān translations. In looking at the functions shared by the features underlying Factor 1, we suggest that the most insightful interpretive label for this dimension is 'density of Qur'ān-specific cohesive ties' in English Qur'ān translations, which employ a complex structure to achieve cohesion and coherence in the verses through lexical cohesion and semantic similarities.

5.3.2 Interpretation of Factor 2: Persuasive and Imaginal Style

Factor 2 concerns three MRC psycholinguistic lexical features and one semantic lexical feature, which explain 13.9% of the total variance. Factor 2, taken as a whole, is the second explanatory factor in our model, representing another salient textual variation between the English Qur'ān translations, which is another source of reading difficulty or ease. The three MRC lexical features (word imageability, word concreteness, and word meaningfulness) all have positive weights of larger than .700 for this factor, with word hypernymy having a positive weight of .571, which is still large loading (see Table 5.11 in section 5.3).

The English Qur'ān translations have high frequencies of all these lexical features, which are functionally related to tangible details. These co-occurring features are all used to represent statements on high-level sensory details that can be seen, touched, felt, and imagined and which are thus more meaningful and easier to process and comprehend than abstract details (McNamara et al., 2014; 2011). Word hypernymy is a hierarchical taxonomic feature of a word that usually tends to be more concrete and is highly correlated with MRC word concreteness (McNamara et al., 2014). Word imageability indicates the use of vivid language that evokes mental images in the mind of the reader (McCarthy et al., 2006; Paivio et al., 1968) and is highly correlated with word concreteness (Connell and Lynott, 2012). The presence of imageable words indicates the use of metaphorical language (Broadwell et al., 2013); and highly imageable words are used as a discourse technique for narrative persuasion in commercial industries (Green and Brock, 2002).

The frequent co-occurrence of these lexical features can have a persuasive and imaginative function. This can explain the argumentative structure of persuasion and imagination in either the English Qur'ān translations or the Arabic Qur'ān, which are marked by sensory and cognitive features evoking mental imagery in the reader's minds. Examples 1 and 2 illustrate the typical features of persuasion and imagination that are employed by highly concrete words,

imageable words, meaningful words, and hypernyms to evoke mental images and emotions about the Day of Resurrection. Example 3 is an illustration of the argumentative style of persuasion and imagination. It depicts God's omnipotence in sensory details and vivid images, intended to be convincing and appealing – for example, God's ability to send water from the sky and provide us with food.

1. (وَإِذَا الْجِبَالُ سُيِّرَتْ وَإِذَا النُّجُومُ انكَدَرَتْ إِذَا الشَّمْسُ كُوِّرَتْ)

[When the **sun** (with its spacious light) is **folded up**; When the **stars fall**, losing their lustre; When the **mountains vanish**. Q81:1-3] (Ali, 1934).

2. (يَوْمَ تَرَوْنَهَا تَذْهَلُ كُلُّ مُرْضِعَةٍ عَمَّا أَرْضَعَتْ وَتَضَعُ كُلُّ ذَاتِ حَمْلٍ حَمْلَهَا وَتَرَى النَّاسَ سُكَارَىٰ وَمَا هُم بِسُكَارَىٰ وَلَٰكِنَّ عَذَابَ اللَّهِ شَدِيدٌ)

[on the **Day** you see it, every **nursing mother** will think no more of her **baby**, every **pregnant female** will **miscarry**, you will think **people** are **drunk** when they are not, so severe will be **God's torment**. Q22:2] (Abdel-Haleem, 2004).

3. (ثُمَّ شَقَقْنَا الأَرْضَ شَقًّا فَلْيَنظُرِ الْإِنسَانُ إِلَىٰ طَعَامِهِ أَنَّا صَبَبْنَا الْمَاءَ صَبًّا)

[Let Man consider his nourishment [food]. We **poured** out the **rains** abundantly, then We **split** the **earth** in fissures. Q80:25-26] (Arberry, 1955).

Overall, the findings indicate that the most appropriate interpretive label for the dimension underlying this factor's features is 'persuasive and imaginal style' in English Qur'ān translations. This dimension indicates the degree to which persuasion and imagery are expressed in verses with large quantities of sensory detail, which might be used as a discourse strategy – not to give readers evidence for or arguments about Qur'ānic claims (see Abdul-Raof, 2019), but to move the reader or appeal to their emotions and senses (see example 2).

5.3.3 Interpretation of Factor 3: Integrative Structure

Factor 3 concerns five positive linguistic features: noun-phrase density, modifier per noun phrase, stem overlap – adjacent sentences, stem overlap – all sentences, and word length. This factor explains 9.6% of total variance in the factor analysis, representing a prominent pattern of textual variation between the English Qur'ān translations that could contribute to text difficulty or ease. Chafe (1982), in an article explaining four dimensions of textual variation in spoken and written language, proposes 'integration' to refer to how a large quantity of information is integrated and packed into a text with fewest words, using features such as nouns, prepositions, attributive adjectives, and nominalisations. Most of the five positive

linguistic features in Factor 3 share the function of informational density and integration in a text.

Word length, nouns and attributive adjectives are 'associated with a high informational focus and a careful integration of information in a text' (Biber 1988, p.104); they are also some features of discourse complexity (Biber 1992). Attributive adjectives are simply modifiers of phrases. A phrase modifier is used to elaborate on nouns by integrating more careful word choices to explain the noun's precise meaning, bringing a higher density of information to a text (Biber, 1988). As nominalisations and nouns are used in a text to achieve informational integration and to expand informational size (Chafe, 1982), noun phrases also perform the same function. A high frequency of nouns and noun phrases indicates a high density of information (Biber, 1988; McNamara et al., 2014). Stem overlap measures the nouns in one sentence (e.g., swimmer) that overlap with those in another (e.g., swimming), coming from the same semantic morpheme (McNamara et al., 2014). Stem overlap might have an integrative function and carefully integrating terms from the same root to provide cohesion and additional information.

The co-occurrence of these five features shares an integrative function, presenting information in a concise and precise manner (Biber, 1988), used 'to create maximum effect with the fewest words' (Tannen, 1982, p.3). The English Qur'ān translations have high frequencies of all these markers of informational integration. Accordingly, this study suggests that the most appropriate interpretive label for the dimension underlying this factor is 'integrative structure'.

5.3.4 Interpretation of Factor 4: Elaborative Structure

The five features with strong positive weights on Factor 4 (see Table 5.11) are additive connectives (e.g., 'and'); causal connectives (e.g., 'because', 'so'); adversative/contrastive connectives (e.g., 'although', 'though', 'whereas'); logical connectives (e.g., variants of 'or', 'and', 'not', 'if', 'then'); and 'all connectives'. This factor explains 6.1% of the total variance, revealing another important pattern in textual variation between the translations, which could contribute to the difficulty or ease of comprehension. The co-occurrence of these connectives or conjunctions contributes to semantic relationships between words, sentences, and clauses (Halliday and Hasan, 1976).

'And', as an additive connective, is used as a device for simple linking of sentences (Chafe, 1982; Beaman, 1984). It is associated with fragmented structure in spoken discourse (Chafe, 1982). This simple connector abounds in the Qur'ān, typically occurring between nouns, relative pronouns, verbs, adjectives, prepositional phrases, and nominalised nouns (Abdul-

Raof, 2019). Whilst other connectives associated with this factor (i.e., causal, conditional, and adversative) are often used as devices for complex linking of sentences (Beaman, 1984; Millis et al., 1993). Several authors have examined the functions of these connectives. Biber (1992) suggests that these connectives are features of 'structural elaboration', which is associated with linguistic complexity. These connectives are used for informational elaboration relating to the author's message, feeling, and attitude (Biber 1988; Millis et al., 1993), and they usually expand the unit of ideas or sentences in loose presentation (Biber 1988; Millis et al., 1993). Such an expansion results in higher syntactic complexity and increases reading difficulty (Binkley, 1988; Millis et al., 1993; McNamara et al., 2014). Finally, whilst the co-occurrence of logical connectives with the other features of Factor 4 requires further research, logical connectives are associated with complexity (Crossley et al., 2012). McCarthy et al. (2006) claim that frequent use of logical connectives (such as 'or', 'and', and 'not') is likely to put a larger load on working memory.

Overall, the co-occurrence of these connectives seems to share the function of informational elaboration that is syntactically constructed with simple or complex linking. English Qur'ān translations represent high frequencies of all these markers of informational elaboration. Accordingly, the label 'elaborative structure' is suggested for the dimension underlying this factor in English Qur'ān translations.

5.3.5 Interpretation of Factor 5: Narrative Style

Factor 5 comprises four linguistic features among English Qur'ān translations, which are temporal connectives, expanded temporal connectives, adverbial-phrase density, and word frequency. These features explain 4.6% of the total variance, reflecting another textual dimension in these English Qur'ān translations. Three of the linguistic features have positive weights of larger than .500 for this factor, whilst word frequency has a positive weight of .391, which remains a large loading on this factor (see Table 5.11).

Temporal connectives (e.g., 'before', 'then', 'later' 'when', 'until') and adverbial phrases (e.g., 'in a moment', 'sooner or later') are markers of temporality used to describe events, particularly in narrative texts (McNamara et al., 2012). The temporal features are used to depict 'internal event timeframes, such as an event that is complete or ongoing. The occurrence of events at a point in time can also be established by a large repertoire of adverbial cues such as before, after, then' (Duran et al., 2007, p.202). Duran et al. (2007) used temporal features to predict temporality between narrative texts and expository texts. Their results suggest that narrative texts incorporate more temporal cohesion markers than expository texts. In a similar

study, McCarthy et al. (2006a, no pagination) investigated textual differences between narrative (literary) texts and expository (science) texts. Their results found that expository texts are 'more inter-sententially and intertextually cohesive', whilst narrative texts integrate 'more inter-clausal and temporal cohesion markers'. Finally, McNamara et al. (2012) indicated that narrative texts intend to have more frequent words, high causal cohesion, and high temporal cohesion.

The features grouped in Factor 5 seems to be functionally associated with linguistic markers of narration. English Qur'ān translations show high frequencies of all these markers of narration. According, the label 'narrative style' is suggested for the dimension underlying this factor in English Qur'ān translations.

5.4 Discussion of Quantitative Findings

This section discusses the results of automated measures in relation to the first two research questions. The four research questions are as follows:

1. What are the primary textual factors that could affect the readability and comprehensibility of Qur'ān translations, as indicated by automated measures?
2. Does retranslation result in higher levels of readability and comprehensibility?
3. Which of the three translations under study is the most readable and comprehensible, and which is the least readable and comprehensible, as indicated by human judgement?
4. Do human judgements of readability and comprehensibility correlate with the automated measures and yield consistent findings?
5. Do readers with low (or high) levels of prior knowledge of the Qur'ān gain more understanding from lower (or higher) levels of readability?

The last three questions are concerned with human judgements (see Chapter 6). To identify the text factors affecting Qur'ān translation readability, this research used two types of statistical analyses and several measures of readability, cohesion, and language to determine differences between the three versions and to explore the language patterns of the three translations. The results of the MANOVAs and ANOVAs suggest significant differences between the three translations on a multiple range of text characteristics, including lexical, syntactic, semantic, and discourse features. The differences and similarities between the three versions of these text characteristics are discussed and systematised in light of six major text factors: *style, cohesion,*

literalness, *genre*, *register*, and *retranslation*. Each is discussed in turn in the following sections. The six content factors might describe what makes a Qur'ān translation more readable and comprehensible.

5.4.1 Reading Grade Level

The five classic readability formulae were used in this research to underpin a preliminary analysis, as they indirectly measure surface features of a text that cause difficulty for comprehension: word length and word frequency, which are features of word difficulty, and sentence length, which is a feature of syntactic complexity (McNamara et al., 2014; Klare, 1976; Chall, 1996).

The results of the classic formulae show that the Abdel-Haleem (2004) translation is estimated at a reading grade level of 3-5 (10–11 years of age) across all five formulae (see section 5.2.1), as derived from the total means of all 40 chapters. The five formulae give the Arberry (1955/1983) translation a reading grade level of 6-11 (16–17 years of age), as derived from the total means of all chapters. The five formulae estimated the Ali (1934/1989) translation as having a reading grade level of 6-7 (12–13 years of age), as derived from the total means of all chapters. The three translations can thus be roughly understood by readers of 10–17 years, according to Flesch's score description (1949) and other formulae. Abdel-Haleem's version was two grade levels lower than Ali's and four grade levels lower than Arberry's. In brief, the readability formulae suggest that the most readable version is Abdel-Haleem's translation, and the least readable is Arberry's.

The classic formulae are not sufficient to explain the other features of a text that contribute to text difficulty (see Chapter 3). In other words, they do not reveal which of the individual textual features contribute most to the readability of the Qur'ān translations. Thus, the theoretical framework of our analysis was not limited to one feature or subscale of text comprehension and discourse processing, as a text can be 'difficult according to some subscales but not for others' (McNamara et al., 2014, p.15). Thus, this research adopted a multilevel theoretical framework to assess text comprehension and readability to identify the difficulty of the three translations on the basis of various text characteristics, including lexical, syntactic, semantic, and discourse features.

5.4.2 Style

The content factor that most influences readability is the style of writing (Gray and Leary, 1935; Klare, 1968, 1988; Harrison, 1980). The most popular and commonly used style variables in readability studies and classic formulae are word length, word frequency, and sentence length. In this study, the writing styles of the three translations were measured according to two major aspects of style: word difficulty and syntactic complexity. The style aspects were assessed against several computational indices of lexical and syntactic features and now discussed in turn.

Based on the findings of CELEX word frequency and word length, Abdel-Haleem's writing style uses shorter and more common words than the others. Arberry's version uses shorter but less common words than Ali's, though the differences between the two are not significant. These two variables provide an excellent proxy for text readability and comprehension difficulty (McCarthy et al., 2006b; McNamara et al., 2014; Crossley and McNamara, 2011c; Klare, 1968, 1988; Harrison, 1980). Arberry and Abdel-Haleem are consistent in the frequency of their use of these features across the short, medium, and long chapters (see section 5.2.2). However, Ali's words in the medium chapters tend to be shorter than those in the short chapters. The same trend was not found in the other translations

The significant result for Abdel-Haleem's word lengths reflects the fact that shorter words are typically higher frequency; thus, longer words tend to be less common, in line with 'Zipf's law' (1949, cited in McNamara et al., 2014) on the relationship between word frequency and length. The non-significant results for these two variables for the other translators suggest that Ali's and Arberry's use of less common words might represents the stylistic flavour of an age that is now archaic. As Dickins et al. (2017, pp. 216-217) write, the use of archaisms in earlier English Qur'ān translations mirrors the language of earlier Bible translations. The use of archaisms in a text that is intended for a worldwide readership could limit the readers' understanding of the Qur'ān's meanings.

Word difficulty is also associated with the use of hypernyms, or 'word specificity' (McNamara et al., 2014; Crossley et al., 2017). Abdel-Haleem appears to use significantly more hypernyms across all 40 chapters than either Ali or Arberry does, indicating that the latter two use more general words. Whilst Abdel-Haleem uses more specific words (e.g., 'chair' and 'table' are specific words and hyponyms of 'furniture'). In addition, whilst Abdel-Haleem makes more use of hypernyms in short chapters than elsewhere, Ali and Arberry both use this feature less often in short chapters. The overall findings on hypernyms suggest that Ali's and Arberry's words tend to be more abstract than Abdel-Haleem's. According to McCarthy et al.

(2007) and McNamara et al. (2014), hypernyms are also associated with word concreteness. The findings of this study do not conflict with previous research, indicating that Abdel-Haleem's translation uses significantly more concrete words and hypernyms than the other translators, which contributes to more readable writing.

The complexity of the word items appears to be greater when some measures of lexical-psycholinguistic features – such as concreteness, imageability, and meaningfulness – are low: 'The impact of these psychological properties on text difficulty is intuitively straightforward' (McNamara et al., 2014, p. 42). Arberry's and Ali's translations scored significantly lower for two of these three variables than Abdel-Haleem's translation did, with the exception of word concreteness, for which Abdel-Haleem and Arberry scored equally. These results suggest that Abdel-Haleem's words are more easily visualised in the reader's mind than Arberry's and Ali's, whilst Ali's words are less easily visualised than Arberry's. Abdel-Haleem's words are more meaningful and concrete than those of the other, older translations, whilst Ali's words are less meaningful and concrete than Arberry's.

Arberry's and Ali's words are less likely to have a higher incidence of all five variables. A higher incidence of shorter words, word frequency, hypernyms, concreteness, imageability, and meaningfulness in the writing style of a translation might facilitate text comprehension because they allow the expression of meaning in a clear, simple, and vivid manner, which keeps the reader engaged. In addition, these four psycholinguistic variables are associated with the use of artistic imagery properties, which is a characteristic of Qur'ānic discourse.

Another aspect that makes a writing style more or less difficult to understand is syntactic structure and sentence types. The third-section analysis provides important insights into the translations with low and high syntactic structure complexity based on multiple syntactic indicators, including sentence length, phrase types (i.e., noun phrases and adverb phrases), modifiers per noun phrase, and number of words before the main verb. Abdel-Haleem's translation uses significantly shorter sentences than Ali's. It also uses shorter sentences than Arberry's translation, albeit the difference is not significant. Long sentences do not necessarily impede comprehension (Anderson and Davison, 1988), but different sentence structures can contribute to syntactic complexity and text difficulty (Klare, 1975; Harrison, 1980; McNamara et al., 2014). Compared to the translations of Ali and Arberry, Abdel-Haleem's translation has significantly lower noun-phrase density, lower adverbial-phrase density, lower verb-phrase density, fewer modifiers per noun phrase, and fewer words before the main verb. Ali's and Arberry's translations have equal or greater numbers of modifiers per noun phrase and words

before the main verb. Texts with high densities of these elements are usually structurally dense and syntactically complex; they impose a larger burden on working memory and are thus more difficult to understand (Graesser et al., 2006; McNamara et al., 2014). Abdel-Haleem's translations use significantly simpler and less embedded syntactic structures than the other two translations, which might have a substantial effect on text comprehension. The translations of Ali and Arberry, in contrast, use larger proportions of complex sentences with embedded constituents, which might cause difficulties when understanding sentences.

To conclude, the stylistic choices reflected in Ali's and Arberry's translations at the level of lexis and syntax are likely to be consistently higher in their complexity across short, medium, and long chapters than those in Abdel-Haleem's translation, which adopts a more contemporary and simple style of language across all chapters.

5.4.3 Cohesion

Cohesion is one of the most important content factors in terms of its effect on the level of translation difficulty. Several scholars in the fields of education, comprehension, readability, psychology, and linguistic studies have shown that text cohesion contributes to readability, learning, recall, and comprehension because it helps to tie the thoughts of the content together for many readers, particularly those with less experience and background (Carpenter and Just, 1977; Halliday and Hasan, 1976; Harrison, 1980, Freebody and Anderson, 1983; Keenan et al., 1984; Armbruster, 1984; Binkley, 1988; McNamara and Kintsch, 1996; McNamara et al., 2014). This research used 16 variables from the Coh-Metrix, falling into three main categories: referential cohesion, semantic coherence, and connectives. Each category is discussed here in turn.

In the present study, Ali's translation scored significantly higher than the others for both local and global referential cohesion and semantic coherence. Abdel-Haleem's translation had the lowest scores for all variables. The scores for Arberry's translation were not significantly different to those of Ali for most variables. Abdel-Haleem's translation uses fewer noun, argument, and stem overlaps between adjacent and all other sentences. In contrast, both Ali's and Arberry's translations use more noun, argument, and stem overlaps between adjacent and all other sentences. These results suggest that Ali's and Arberry's translations have a greater incidence of lexical repetition at both the local and global levels than Abdel-Haleem's.

In addition to variables of lexical cohesion, another cohesive aspect that makes reading less difficult is the semantic relations or similarities between words and sentences (Foltz, 2007; McNamara et al., 2014). The results reveal that Abel-Haleem's translation includes

significantly fewer semantic similarities at both the local and global levels. Ali's and Arberry's translations are more semantically coherent than that of Abdel-Haleem. A comparison of all connective variables across the three translations reveals that Ali's translation has the highest scores. Abdel-Haleem has the lowest scores for most of the variables. Arberry's translation is closer to Ali's for most of the variables in terms of average values. In brief, Ali's and Arberry's translations contain more logical, temporal, additive, and contrastive/adversative connectives than Abdel-Haleem's.

The findings for these three categories (i.e., referential cohesion, semantic coherence, and connectives) reflect the strength of Coh-Metrix in discriminating between low-cohesion and high-cohesion translations. Statistical examination of the 16 chosen variables reveals that Ali's translation has a greater degree of referential cohesion, semantic coherence, and connectives than Abdel-Haleem's. Ali's overall averages in these three categories are not very different to those of Arberry. The software analysis investigated which of the translations was more statistically coherent. The Coh-Metrix results for referential cohesion, semantic coherence, and connectives yielded two major findings with regard to readability and literalism (see the following section).

From a readability perspective, another explanation for the cohesive ties results is that higher cohesion does not always indicate a higher-quality text, according to some studies. For instance, McNamara et al. (2014, p. 109) point out that 'cohesion indices such as content word overlap, LSA G/N, aspect repetition were negatively related to essay quality. Higher quality essays had lower cohesion'. Similarly, Crossley et al. (2011d) found that texts with a great incidence of content word overlap and positive logical connectors are indicative of less skilled writers. They add that these features of cohesion 'should make the text more readable, but, conversely, the text is assessed to be of a lower quality' (Crossley et al., 2011d, p. 302). In another study, Crossley and McNamara (2011c, p. 187) conclude that 'an increase in cohesion is a negative predictor of text quality'.

A comparison of the findings for cohesive ties with those of readability formulae revealed that those translations with high cohesion (i.e., Ali and Arberry) are more difficult than the translation with low cohesion (Abdel-Haleem). This aligns with the findings of McNamara et al. (2010). This finding is unsurprising: 'adding cohesion generally results in longer sentences and, consequently, higher grade-level texts according to Flesch-Kincaid Grade indexes […] lengthening the texts makes the texts more difficult' (McNamara et al., 2010, p.317).

The presence of connectives increases sentence length (Millis et al., 1993). The connectives also are associated with discourse complexity (Biber, 1992) and, therefore, result in higher reading difficulty (Millis et al., 1993; McNamara et al., 2014). The results show that both Ali's and Arberry's translations make greater use of connectives and thus have higher grade levels for reading difficulty. In contrast, Abdel-Haleem's translation uses fewer connectives and thus has a lower grade level for reading difficulty. The Millis et al. (1993) experiment, with its six-point scale of reading difficulty, examined the impact of connectives on expository texts and found that connectives did not facilitate recall, which was high in passages without connectives and low in passages with connectives. According to Millis et al. (1993), this result sounds counterintuitive and conflicts with the conclusions of other scholars, such as Carpenter and Just (1977) and Halliday and Hasan (1976), who claim that connectives facilitate the comprehensibility of a text.

5.4.4 Literalness

This section focuses on the degree of *literalness* as another possible factor influencing the readability of a Qur'ān translation. Literalness is viewed in this study as the act that transfers the ST stylistic and semantic norms to the TT (Abdul-Raof, 2018). According to Akbar (1978, cited in Abdul-Raof 2001), many English Qur'ān translations suffer from this translation style because it makes the message of the Qur'ān 'unintelligible'. It was demonstrated in Chapter 4 that literalness can be detected through the incidence of the connector 'and'. The findings of this study indicate that the variables of referential cohesion and semantic coherence are also indicators of literal translation.

As mentioned earlier, Ali's and Arberry's translations have similar patterns in respect of most variables of referential cohesion and semantic coherence and in some variables of connectives, such as the operator 'and' (additive connective). In fact, the translations show very little statistical difference at all. This resemblance in all types of lexical repetition and semantic similarities – and in their use of the operator 'and' – makes plain that Ali and Arberry have both attempted to maintain the ST language's cohesive and semantic patterns in their English translations. Abdel-Haleem's translation, in contrast, is less affected by the ST language's cohesive and semantic patterns and the Arabic operator و ('and'). These findings show that Abdel-Haleem's translation has significantly fewer cohesive and semantic ties and occurrences of the operator 'and', which usually make writing less readable. The results for cohesive ties provided by Coh-Metrix thus highlight the unnecessary closeness (i.e., literalness) of these two earlier translations to Qur'ān-specific cohesive ties.

To conclude, a higher proportion of these cohesive ties in Qur'ān translations is indicative of literalness. The combination of the variables of referential cohesion, semantic coherence, and the operator 'and' needs further testing on other Qur'ān translations. This would allow the production of a formula[11] that could predict the degree of literalness in Qur'ānic translations.

5.4.5 Genre

Another essential factor that contributes to the difficulty of a text is its genre in relation to the degree of narrativity (McNamara et al., 2014). Research findings have shown that 'narrative text is substantially easier to read, comprehend, and recall than other genres of text such as science, history, and other expository domains' (Bruner, 1986; Haberlandt and Graesser, 1985; Graesser and McNamara, 2011; cited in McNamara et al., 2014, p. 100). This is due to the fact that narrative discourse has features such as oral language and series of events that are likely to be familiar to readers and easy to understand (Biber 1988; McNamara et al., 2014). Section 5.4.5 provides a discussion on some characteristics of narrative style.

Narrativity is a salient generic characteristic of the Qur'ān. As Abdul-Raof (2019, p.23) writes, the Qur'ān is 'a text type (genre) in its own right', but it has legal, narrative, and instructional passages. An assessment of the degree of narrativity in these three translations (see section 5.2.5) shows that Abdel-Haleem's translation scored high for narrativity across the three types of chapter, whilst Ali's translation scored low and Arberry's fell between the two. This result suggests that the versions of the latter two translators contain fewer features of narrativity. One of the reasons for low narrativity might be the use of archaic forms that are not a feature of narrative style (see Biber, 1988; McNamara et al., 2014). The finding of high narrativity in Abdel-Haleem's translation does not contradict the findings of the word frequency analysis, which show that Abdel-Haleem's translation uses more familiar words than the other translations do.

All the short chapters across the three translations scored consistently low for narrativity, whereas the medium-length and longer chapters scored consistently high. Future studies should consider the degree of narrativity in relation to the different types of Qur'ānic chapter to understand whether the themes of short chapters are more likely to concern objects (such as the creation of mankind, the creation of universe) than theological notions (such as God's

[11] For more information on how to make a composite measure and how to create a formula, see Graesser et al. (2014) and H. Li and Graesser (2015), who combined automated variables to produce a formula that predicts the degree of formality in a text.

omnipotence, polytheism, monotheism, resurrection), in comparison to the themes of the longer and medium-length chapters, although Qur'ānic chapters 'interlock and get engaged with each other by overlapping themes' (Abdul-Raof, 2019, p.266).

To conclude, the specific characteristics of a genre require specific linguistic registers and rendering in the TT, and this could have an effect on translation readability and comprehension. The narrative style is a silent variation in the language of Qur'ān translations. The next section presents the textual dimensions of register variation in English Qur'ān translations.

5.4.6 Register

Register is another content factor influencing the comprehensibility of a Qur'ān translation. According to Biber (1992, p.134), discourse complexity is associated with 'differences across situational varieties or registers'. For example, his study found that written registers (such as biographies, editorials, academic prose, etc.) display more complexity than spoken registers (such as telephone conversation, debates, etc.). The register of texts can be 'characterized by a set of co-occurring linguistic features' (Biber, 1988, p.21). Biber's multi-dimensional analysis of textual variation was applied in this study within the complexity discourse of specific texts or registers (i.e., Qur'ān translations). The present study's findings show that the language of English Qur'ān translations is characterised by features of Qur'ān-specific cohesive ties, persuasive and imaginal style, integrative structure, elaborative structure, and narrative style.

These five textual dimensions not only describe the linguistic characteristics of the translations. They also reveal which characteristics make a translation more or less readable and comprehensible in the same register. The purpose of extracting these textual dimensions was to determine those that account for text difficulty and to compare the mean differences between the three translations in terms of the linguistic features that underly these dimensions.

Dimension 1, 'density of Qur'ān-specific cohesive ties', characterises some verses in the English Qur'ān translations that employ complex structure to achieve cohesion and coherence through lexical cohesion and semantic coherence. This dimension is very strong in our multidimensional model and represents a typical feature of (Arabic) Qur'ānic style, relying on lexical cohesion and semantic coherence in verses to express more explicitly the relationships between the meanings of various verses (see Abdul-Raof, 2019, 2001). The translations with high scores for the features of Dimension 1 score highly for Qur'ān-specific cohesive ties, reading difficulty, and syntactic complexity. In contrast, the translations with low scores for this dimension have the reverse characteristics. Abdel-Haleem's translation scores significantly low, Arberry's translation has a medium-level score, and Ali's translation is high on the

features of Dimension 1. As shown in the first part of the analysis, Abdel-Haleem's translation uses significantly shorter words, shorter sentences, simpler syntactic structures, and fewer Qur'ān-specific cohesive ties than the other two texts. This contributes to giving Abdel-Haleem's translation a lower grade level for reading difficulty.

Dimension 2, 'persuasive and imaginal style', describes the argumentative style of persuasion and imagination found in some verses of the translations where there are high levels of vivid language and imaginal details. Linguistic features associated with a persuasive and imaginative style are associated with sensory features such as concrete words, imageable words, hypernyms, and meaningful words, which are all major predictors of text difficulty or ease (McNamara et al., 2014). This dimension is another robust variation in the discourse of English Qur'ān translations; it represents a typical feature of the Qur'ānic style for use of artistic imagery properties (see Abdul-Raof, 2019). It is used as a discourse strategy to either convince the reader or provide evidence to the reader through tangible and sensory features of Qur'ānic matters (such as God's abilities, creation, power, etc.). The translations with high scores for the features of Dimension 2 employ a more persuasive and imaginal style, with more vivid language, imagery, and concrete details that are easy to understand and process. Those Qur'ān translations with low scores on this dimension are less persuasive and do not create vivid images in readers' minds. Hence, they score higher for reading difficulty. Abdel-Haleem's translation scores significantly high for the features of Dimension 2, whilst Arberry's is ranked medium and Ali's low, as shown in the first part of the analysis.

Dimension 3, 'integrative structure', explains how some verses in English Qur'ān translations integrate large quantities of information, presenting maximum content with minimal words (Biber, 1988; Tannen, 1982). Some of the co-occurring linguistic features in this dimension are identical to the features found by Biber (1988; 1992), associated with high informational density and integration of information into a text and are markers of discourse complexity (Biber, 1992). The English Qur'ān translations differ considerably in their respective degrees of informational density. Abdel-Haleem's translation seems to integrate less information than either Ali or Arberry's texts, as Abdel-Haleem uses shorter words and fewer noun phrases, modifiers per noun phrase, and stem overlaps (as shown in the first part of the analysis), resulting in lower informational density. Ali's translation seems to pack large amounts of information in a more integrative way than Arberry's does, as Ali uses longer words and more noun phrases, modifiers per noun phrase, and stem overlap, all resulting in higher informational density and greater structural complexity.

Dimension 4, ‘elaborative structure’, concerns how some verses in the translations are informationally and translationally elaborated with logical, clausal, contrastive, additive, and conditional connectives. Some of the co-occurring connectives in this dimension are similar to the elaborated structure of Biber (1992), which is associated with discourse complexity. Some of these connectives in Dimension 4 are associated with informational elaboration related to the speaker’s message, attitudes, and beliefs (Biber, 1988; Millis et al., 1993). However, the presence of connectives increases the length of sentences, which places a higher load on working memory and thus creates greater reading difficulty (Millis et al., 1993; McNamara et al., 2014). The translations with lower scores for the features of this dimension are likely to present information using the simple chaining of ideas, as in Abdel-Haleem’s translation, which results in greater reading ease. The Qur’ān translations with high scores, in contrast, tend to elaborate using complex chaining of ideas. Ali’s and Arberry’s translations tend to present information using more complex chaining of ideas because they score highly for all connective categories, which do not differ from each other in terms of most connectives, particularly the occurrences of logical and additive connectors. Ali’s and Arberry’s translations use significantly higher densities of logical operators and additive connectors – particularly ‘and’ – than Abdel-Haleem’s. High densities of logical operators (such as ‘and’, ‘or’, ‘if’, and ‘then’) reflect the complexity of a text and place a large load on working memory (Graesser et al., 2004). The overuse of the additive connector ‘and’ seems to be an indicator of literalism (literal translation), which might reduce the clarity and understanding of a translation.

Dimension 5, ‘narrative style’, distinguishes some verses of English Qur’ān translations that are expressed in some characteristics of narrative style. This style typically employs everyday language, with temporal and adverbial features that tell a story using characters, locations, and time references (Biber, 1988; McNamara et al., 2014). The features in Dimension 5 – such as temporal connectives, expanded temporal connectives, adverbial-phrase density, and word frequency (i.e., words more common in everyday language) – capture the characteristics of the narrative context (see Biber, 1988; Louwerse et al., 2004; Duran et al., 2007; McCarthy et al., 2006a; McNamara et al., 2014, 2011). The three English Qur’ān translations each employ large numbers of these stylistic markers of narration, but they differ in their respective degrees of narrativity. As reported in the results of the first part, the three translations are not significantly different in their respective uses of these stylistic markers, except word frequency, which was significantly lower in Ali’s and Arberry’s translations than in Abel-Haleem’s. This difference in word frequency explains the differences in their scores for Coh-Metrix’s narrativity. The

degree of narrativity is low in Ali's translation and medium-level in Arberry's, due to their frequent use of archaic words. In contrast, it is high in Abdel-Haleem's translation due to the frequent use of familiar words. Finally, and most importantly, the features underlying Dimension 5 in this research show some similarities with the features of Coh-Metrix's narrativity, thus pointing to a similar conclusion regarding narrativity as a textual dimension.

In conclusion, the Qur'ān translations are characterised by features of Qur'ān-specific cohesive ties, persuasive and imaginal style, integrative structure, elaborative structure, and narrative style. These five Qur'ānic textual dimensions can explain the linguistic features associated with complexity in English Qur'ān translations. Further research is needed to examine these dimensions in other English Qur'ān translations to enhance our multidimensional model, along with consideration of additional linguistic features, such as the incidence of all pronoun categories, passives, relative clauses, and so on, which are not examined in this research and which might underpin other dimensions affecting a Qur'ān translation comprehensibility.

5.4.7 Retranslation

Retranslation is commonly used to describe subsequent translations of the same ST into the same TL. The Qur'ān has been retranslated into English more than one hundred times, as discussed in Chapter 2. The first English Qur'ān version was an indirect translation, whilst all subsequent versions have been retranslations (see Chapter 2). Hence, the Qur'ān versions used in the analysis are classified into two types: two earlier translations and a recent retranslation. The next discussion focuses on the phenomenon of retranslation from two points of view: the validity of the retranslation hypothesis (RH) as proposed by Berman (1990) and Bensimon (1990) and the influence of retranslation on text readability and comprehensibility.

As discussed in Chapter 2, Bensimon and Berman's primary assumption in the RH concerns the translation strategies and linguistic 'ageing' of existing translations compared to subsequent retranslations. Bensimon (1990) argues that earlier translations adopt more domesticating or target-oriented strategies than retranslations, which are likely to be more foreignising or source-oriented. As discussed earlier in the 'literalness' section, the findings of the cohesion measures show that the earlier translations of the Qur'ān – namely, those by Arberry (1955) and Ali (1934) – have no substantial differences in their incidences of lexical repetition and semantic coherence or in their use of the operator 'and'. This makes clear that literality may be a feature of earlier translations that are very close to the stylistic and semantic features of Qur'ānic

discourse and therefore more ST-oriented. This translation strategy brings readers closer to the language of the ST. Such closeness may fail to clearly communicate the message of the Qur'ān. In contrast, Abdel-Haleem's recent retranslation scored significantly lower for all these cohesive variables. The lower incidences of cohesion in the recent retranslation might be due to an attempt to reduce the foreignness of Qur'ān-specific language. Recent retranslations are likely to adopt more audience-oriented strategies, making the TT language less strange to the target reader.

In none of these three versions is there evidence to support the RH claim of a domesticating strategy in earlier Qur'ān translations. However, this claim might be applicable to literacy translations or other types that have fewer translations, where the first translated versions of the same original are still available for use and were not published four centuries ago. This earlier hypothesis does not have the capacity to accommodate religious translations such as this, which has seen hundreds of English retranslations since the 17th century. It can be concluded that retranslation is universally accepted as a phenomenon in all fields of translation, but its hypothesis might not be compatible with certain types of translation.

Berman (1990) emphasises the issue of ageing translations as a justification for retranslation. He argues that existing translations 'age' with the passage of time, due to the language, thus allowing for the emergence of contemporary retranslations. This claim is true when the timespan from an initial translation to subsequent ones is long. When the findings of the automated measures for Abdel-Haleem's version are compared to those of the two earlier translations, Abdel-Haleem's work reflects substantial linguistic changes to the content factors (style, cohesion, literalness, genre, and register). The comparison between Arberry and Ali does not reveal substantial differences in these text factors. This result does not support the RH claim of ageing translations because the works of these two earlier translators were completed within a short period. Thus, a comparison between synchronic and diachronic retranslations might yield more information about linguistic and readability changes than a comparison between synchronic retranslations.

Another explanation for the non-significant differences is that the two earlier translations from the 20th century might be attempting to preserve the stylistic and translation norms of the time. They are rendered in a very high style that is characterised by more complex language and biblical style. This style usually makes the language of a translation more difficult to read and understand, particularly for those readers with less background knowledge. In contrast, the recent retranslation of the 21st century clearly avoids the style of earlier periods for the sake of

achieving higher readability. This finding supports earlier work on other translation types. For example, DuNour (1995) examined retranslations of children's literature into Hebrew to explore the 'linguistic and translational norms' of retranslations published in different periods of time, concluding that readability was found to be a central issue for recent retranslations, whilst earlier translations were identified as being less readable and having a more biblical style, which were the norm for translations in the early 20th century.

As a final word, it is noted that retranslation reveals changes in text readability and comprehensibility when the timespan between the translations is long. Retranslation allows for later works to achieve higher levels of text readability and to avoid the comprehensibility issues of earlier translations. Translation comprehensibility and the effect of readability changes on readers' comprehension – and upon readers with different levels of background knowledge – are explored in the following chapter.

5.5 Chapter Summary

This chapter presented the findings of two automated approaches to measuring text readability and comprehension in three corpora of English Qur'ān translations, allowing us to identify the text factors that make some translations of the Qur'ān more readable and comprehensible than others. This multimethod research used multiple measures of readability, cohesion, and language to determine differences between the three translations using a multivariate analysis and to explore a set of linguistic features and patterns that frequently co-occur in the language of Qur'ān translations. This multilevel theoretical framework of a textual analysis provided a description of the readability of a Qur'ān translation according to six major text factors: *style*, *cohesion*, *literalness*, *genre*, *register*, and *retranslation*. These might contribute to explaining the content factors affecting Qur'ānic translation comprehensibility.

Abdel-Haleem's version reduces the elements associated with difficult text factors, as compared to the earlier two translations. The earlier translations of the 20th century are rendered in more complex style of words and syntax; with higher rates of lexical repetition, semantic coherence, connectives, and literalness; lower rates of narrativity; and a high degree of linguistic features of register variation associated with text difficulty and discourse complexity. Retranslations published within similar periods might have smaller impacts on text readability because the timespan between them is short. The role of retranslation might be important for improving readability when timespans are long because this allows retranslators to avoid the comprehensibility problems of previous translations.

Our theoretical framework of automated measures and statistical data analysis is useful for two fields of study. In readability studies, it could be used for textual analysis to gauge the difficulty of a text according to content factors, such as the style of word and syntax, cohesion, genre, and register (and page layout, see Chapter 6). In translation studies, it could be used for textual analysis to assess the extent of the linguistic changes between retranslations and to explore the discourse registers of specific translations, such as literary and legal texts. To conclude, the findings in this chapter provide insights into the text factors that influence Qur'ān translation readability and cause difficulties for comprehension from the viewpoint of automated assessment. The following chapter presents the effect of readability changes on comprehension and upon those with different levels of background knowledge from the viewpoint of human judgements.

Chapter 6: Data Analysis of Human Judgements

6.1 Chapter Overview

This chapter presents the assessments of human readability and comprehension for two complete chapters, taken from the three corpora of English Qur'ān translations – namely, Abdel-Haleem (2004), Arberry (1955), and Ali (1934). The two chapters are Chapter 91, which comprises 15 verses, and Chapter 86, which comprises 17 verses, totalling 32 verses for evaluation. An individual web-based survey was designed for each translator (see Chapter 4). There were 20 participants for each survey, with a total of 60 participants for the three surveys. Each group of participants was gender-balanced, with 10 males and 10 females, all of whom were adults and native speakers of English.

This results chapter is divided into two parts. The first part presents the human judgements of the Qur'ān translations' readability, and the second explores human judgements of the Qur'ān translations' comprehensibility. A summary and discussion are provided at the end of the chapter, followed by a chapter recap. The readability and comprehension judgements are intended to respond to the last three questions of this study. The human judges interpreted readability and comprehension from differing perspectives and reader viewpoints, allowing comparisons of subjective and objective measures (discussed in the previous chapter).

6.2 Part One: Human Judgements of Translation Readability

Four semantic differential (SD) scales were outlined in the three surveys to allow ratings of the level of readability, based on four text factors: page layout, word usage, sentence usage, and cohesion usage. Six pairs of adjectives were provided for the page-layout scale, five pairs for the word-usage scale, four pairs for the sentence-usage scale, and three pairs for the cohesion-usage scale. The four SD scales had a total of 18 pairs of opposite adjectives (e.g., simple/complex words). Each pair in the scale was treated as a variable in the analysis to distinguish the mean values of human judgements across the three surveys or translators. Each factor was assessed by participants on a 5-point scale, with '5' being a negative attitude and '1' a positive attitude.

The results of the readability judgements were divided into two parts for each of the scales. The first analysis revealed the overall means for all participants for each pair on the scale. The second analysis revealed the overall means of all pairs of the scale as one group between the ratings of two groups of judges. The two groups of participants[12] were analysed: readers with 'low prior knowledge' were those who had never or who rarely read English translations of the Qur'ān, and readers with 'high prior knowledge' were those who frequently read English translations and interpretations of the Qur'ān.

Before reporting the results of the readability judgements, a Shapiro-Wilk test (normality assumption) was performed. Some ratings in each of the four scales across the three surveys falling 0.05 ($p < .05$) suggested a moderate deviation from normality. The mean values and median values of the ratings for each variable were almost equal in each survey (the mean and median values are reported in following sections). This equality between mean and median values indicates lack of significant departure from normality (Rasinger, 2013, p.142). However, the internal consistency of the four SD scales for each survey was checked by Cronbach's alpha reliability coefficients using the reliability analysis in SPSS. The Cronbach's alpha values for each scale were above the recommended value of.70 (Frey et al., 2000), indicating that the four SD scales for the three surveys had a high level of internal reliability.

[12] There were 11 participants with low knowledge in the Abdel-Haleem survey, 13 in the Arberry survey, and 12 in the Ali survey. There were 9 with high knowledge in the Abdel-Haleem survey, 7 in the Arberry survey, and 8 in the Ali survey.

6.2.1 Analysis of Page Layout

Each of the chosen translations has a different page-layout design, thus the influence of these different formats on readability was tested. The participants used the six pairs of adjectives (see Table 6.1) to rate the designs of three Qur'ānic page layouts. The table illustrates that the highest value of the participants' ratings in the Abdel-Haleem survey was for the adjectives *clear–unclear* in relation to text organisation, whereas the lowest value amongst all the pairs was for the adjectives *recognisable–unrecognisable* in relation to the title. Second, the highest value amongst the participants' ratings in the Arberry survey was for the adjectives *clear–unclear* in relation to verse numbers, whereas the lowest value amongst all the pairs was for the adjectives *recognisable–unrecognisable* in relation to title. Finally, the highest value amongst the participants' ratings in the Ali survey was for the adjectives *short–long* in relation to the verse numbers, whereas the lowest value amongst all pairs was for the adjectives *clear–unclear* in relation to verse numbers.

Table 6.1: Descriptive Statistics for the Pairs on the Page-Layout Scale

	Abdel-H. Survey			Arberry Survey			Ali Survey		
Pairs of Adjectives	MED	M	SD	MED	M	SD	MED	M	SD
Recognisable–Unrecognisable Title	1.00	1.05	0.22	1.00	1.25	0.43	1.00	1.65	0.91
Short–Long Version	1.00	1.35	0.57	2.00	1.90	0.62	2.00	1.90	0.89
Simple–Complex Text Layout	2.00	1.90	0.99	1.00	1.80	1.12	1.00	1.55	0.67
Familiar–Strange Paragraph Division	2.00	2.00	1.00	2.00	1.90	0.99	1.50	1.75	0.89
Clear–Unclear Verse Numbers	1.50	2.00	1.18	2.00	2.55	1.53	1.00	1.30	0.71
Clear–Unclear Text Organisation	2.00	2.10	1.18	2.00	1.85	0.77	2.00	1.75	0.85

Note: MED = Median; M= Mean; and SD = Standard Deviation.

An independent-samples t-test was conducted for the six pairs across the three translators. Only three pairs on the SD page-layout scale were found to reflect statistically significant differences between the three translations: *recognisable–unrecognisable* in relation to titles, *short–long* in relation to versions, and *clear–unclear* in relation to verse numbers. The mean values for the adjectives *clear–unclear* in relation to verse numbers indicate that Ali's translation was rated clearer ($M = 1.30$, $SD = 0.71$) than Abdel-Haleem's ($M = 2.00$, $SD = 1.18$) and Arberry's ($M = 2.55$, $SD = 1.53$). The result for Ali and Abdel-Haleem is statistically significant, with a mean difference of 0.70 ($t = 2.208$, $p = .033$), as is that for Ali and Arberry, with a mean difference of 1.25 ($t = 3.223$, $p = .000$), but that for Abdel-Haleem and Arberry is not. This means that Ali's numbering list for verses was clearer than that of Arberry's poetry line numbers and Abdel-Haleem's small superscript at the end of each verse (see samples of their page layouts in section 6.4.1). The responses for Arberry's verse numbers were much closer to the neutral side of the scale, as shown in Figure 6.1.

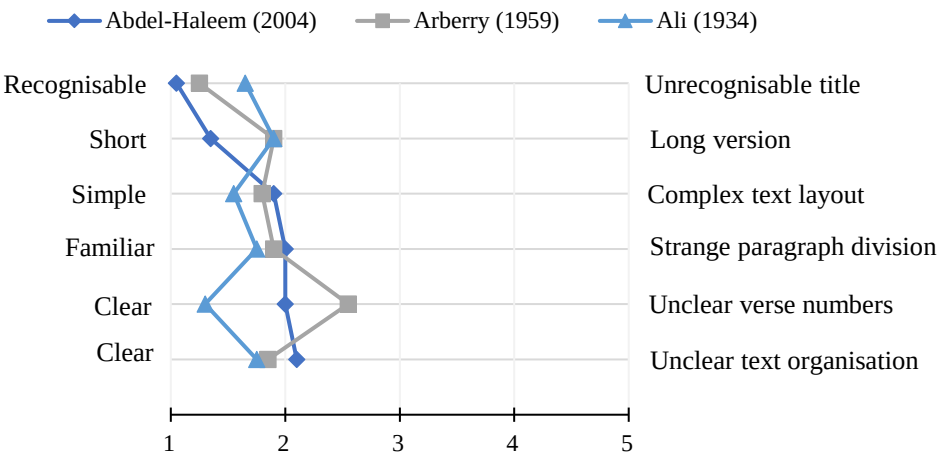


Figure 6.1: Mean Values for the Pairs on the Page-Layout Scale

The second significant result regarding page layout was in the ratings of the Qur'ānic chapter titles as *recognisable vs. unrecognisable*. Abdel-Haleem's titles were found to be more recognisable (M = 1.05, SD = 0.22) than Arberry's (M = 1.25, SD = 0.43) and Ali's (M = 1.65, SD = 0.91). This difference between Abdel-Haleem and Ali was statistically significant, with a mean of 0.60 (t = -2.796, p =.001), but no other differences were found to be statistically significant. Arberry and Abdel-Haleem translated the meanings of titles, using similar translation procedures, whilst Ali translated the titles but added a transliteration before each chapter title. According to the judgements, this procedure made his titles less recognisable than the others.

The final significant result concerning page layout was the ratings of the Qur'ānic chapter length as *short vs. long*. Abdel-Haleem's translation was perceived to be shorter (M = 1.35, SD = 0.57) than Ali's (M = 1.90, SD = 0.89) and Arberry's (M = 1.90, SD = 0.62). The (p) values for this pair were only statistically significant for Abdel-Haleem and Arberry, with a mean difference of 0.55 (t = -2.830, p =.007), whilst Abdel-Haleem and Ali had a mean difference of 0.55 (t = 2.267, p =.029), but Ali and Arberry did not show any significant difference in length. In summary, the page layouts of both Arberry and Abdel-Haleem left more white space than that of Ali, which is filled with Arabic script, an ornamental frame, and explanatory annotations.

On the other hand, the findings from the t-test reveal no statistically significant differences between the three translators for three pairs of adjectives, which were labelled *simple–complex* in relation to text layout, *familiar–strange* in relation to paragraph division, and *clear–unclear* in relation to text organization. In terms of descriptive statistics, Ali's page layout was much

closer to the positive side of the three pairs than that of either Arberry or Abdel-Haleem, as shown in Figure 6.1. In summary, the page layout for Ali's translation was rated (albeit insignificantly) more familiar in terms of paragraph division, clearer in terms of text organisation, and simpler in terms of text layout than the other page layouts. Ali's page layout was rated significantly clearer only in terms of the design of the verse numbers. Abdel-Haleem's page layout was rated significantly clearer in terms of the chapter titles and shorter in terms of its chapter lengths than the other page layouts.

6.2.1.1 Page Layout Judged by Two Groups

Table 6.2 displays the overall means for all six pairs on the page-layout scale, rated by low-knowledge and high-knowledge readers. An independent-samples t-test was conducted to identify the mean differences between the ratings of the two groups of judges for each translation. The responses show that those with low knowledge judged Ali's page layout (M = 1.69, SD = 0.16) to be generally better than those of both Arberry (M = 1.90, SD = 0.39) and Abdel-Haleem (M = 1.96, SD = 0.37). The t-test found no substantial statistical differences between the ratings of this group across the three translations.

Table 6.2: Results for Two Groups of Readers for the Page-Layout Scale

Surveys	Groups	N. of Readers	Mean	S. Deviation	Max.	Min.
Abdel-H.	Low-Knowledge Readers	11	1.96	0.37	2.50	1.41
	High-Knowledge Readers	9	1.56	0.15	1.83	1.39
Arberry	Low-Knowledge Readers	13	1.90	0.39	2.50	1.19
	High-Knowledge Readers	7	1.57	0.22	1.93	1.21
Ali	Low-Knowledge Readers	12	1.69	0.16	1.88	1.38
	High-Knowledge Readers	8	1.43	0.14	1.63	1.25

Similarly, the responses of those with high knowledge suggest that the general design of Ali's page layout was preferred (M = 1.43, SD = 0.14) to those of both Arberry (M = 1.57, SD = 0.22) and Abdel-Haleem (M = 1.56, SD = 0.15). The findings from the t-test illustrate that the (p) values for these readers in relation to the three page layouts were not statistically significantly different.

In summary, the two groups of judges agreed that the design of Ali's page layout was better than the others. The two groups showed similar patterns in terms of their ratings for each of the translations. Low-knowledge readers tended to be severe in their ratings of the translations, whilst high-knowledge readers were less so, as shown Figure 6.2.

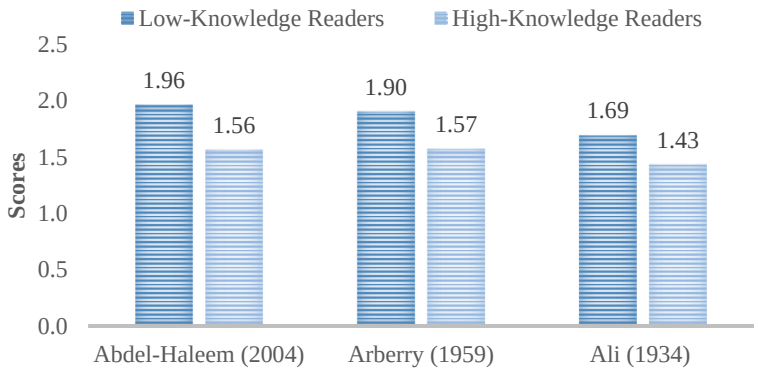


Figure 6.2: Mean Values for Two Groups of Readers for the Page-Layout Scale

6.2.2 Analysis of Word Usage

Table 6.3 displays the descriptive statistics for word usage in relation to the five pairs of adjectives. The overall findings suggest that the highest value given in the Abdel-Haleem survey in relation to words was for the adjectives *modern–archaic*, whereas the lowest value was for *short–long*. In the Arberry survey, the highest value given was for *modern–archaic* (in relation to words), whereas the lowest value was for *frequent–infrequent* (in relation to words). In the Ali survey, the highest value given was for *modern–archaic* (in relation to words), whereas the lowest value was for *frequent–infrequent* (in relation to English words).

Table 6.3: Descriptive Statistics for the Pairs on the Word-Usage Scale

	Abdel-H. Survey			Arberry Survey			Ali Survey		
Pairs of Adjectives	MED	M	SD	MED	M.	SD	MED	M.	SD
Familiar–Strange Words	1.00	1.60	0.73	2.00	2.08	0.98	2.00	2.25	1.04
Modern–Archaic Words	2.00	2.05	0.97	3.00	2.85	1.15	3.00	2.95	0.86
Frequent–Infrequent Words	1.50	1.75	0.94	2.00	1.80	0.75	2.00	2.25	1.04
Clear–Confusing Words	2.00	1.70	0.56	2.00	2.15	0.79	2.00	2.25	0.94
Short–Long Words	1.50	1.50	0.50	2.00	2.25	1.04	2.00	2.25	0.83

Note: MED = Median; M= Mean; and SD = Standard Deviation.

An independent-samples t-test was performed to compare the differences between the findings for the five pairs of adjectives. The t-test found statistically significant differences between the three translations for four pairs of adjectives: *familiar–strange*, *modern–archaic*, *short–long*, and *clear–confusing*. The mean values for *familiar–strange* reveal that the participants found Abdel-Haleem's words more familiar (M = 1.60, SD = 0.73) than Ali's (M = 2.25, SD = 1.04) or Arberry's (M = 2.08, SD = 0.98). The two latter translations were closer to the neutral side of the scale, as shown in Figure 6.3 below. This findings for this variable

were statistically significant for Abdel-Haleem and Ali, with a mean difference of 0.65 (t = -2.220, *p* =.032), but there were no significant differences between other groups.

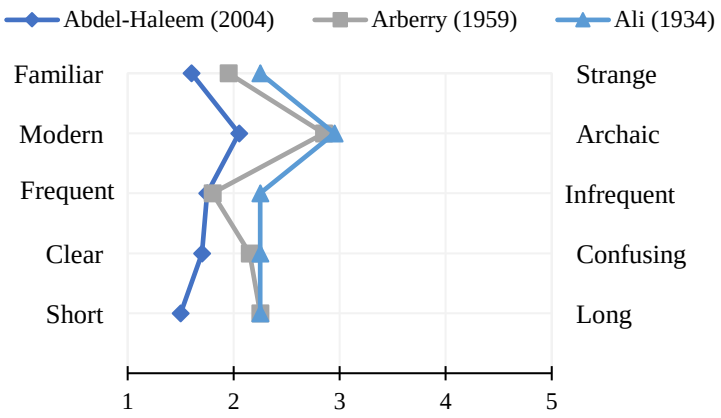


Figure 6.3: Mean Values for the Pairs on the Word-Usage Scale

The second significant result regarding word usage was in the ratings of the Qur'ānic words as *modern vs. archaic.* The participants found Abdel-Haleem's words to be more modern (*M* = 2.05, *SD* = 0.97), placing this translation close to the positive side of the scale, as shown in the figure below, compared to Ali's (*M* = 2.95, *SD* = 0.86) and Arberry's (M = 2.85, SD = 1.15). The two latter translations were somewhat inclined to the negative side of the scale. Responses were statistically significant for this pair for Abdel-Haleem and Arberry, with a mean difference of 0.80 (t = -2.312, *p* =.026), and for Abdel-Haleem and Ali, with a mean difference of 0.90 (t = -3.013, *p* =.005), but not for Ali and Arberry.

The mean values for the adjectives *short–long* (in relation to words) indicate that the words in the Abdel-Haleem survey were rated shorter (M = 1.50, SD = 0.50) than the words in the Ali text (M = 2.25, SD = 0.83) or those in the Arberry text (M = 2.25, SD = 1.04). As shown in Figure 6.3, the two latter translations were closer to the neutral side of the scale for this pair, whereas Abdel-Haleem was rated on the positive side of the scale. The findings from the t-test indicate significant differences between Abdel-Haleem's translation and Ali's, with a mean difference of 0.75 (*t* = -3.376, *p* = 0.02), and between Abdel-Haleem's and Arberry's, with a mean difference of 0.75 (*t* = -2.827, *p* = 0.07), but not between Ali's and Arberry's.

The final significant result for the three translations concerns the mean values for the adjectives *clear–confusing* (in relation to words). Based on the survey responses, Abdel-Haleem's words are significantly clearer (*M* = 1.70, *SD* = 0.56) than Ali's (*M* = 2.25, *SD* = 0.94) and Arberry's (*M* = 2.15, *SD* = 0.79). This variable was statistically significant for Abdel-

Haleem and Arberry, with a mean difference of 0.45 (t = -2.026, *p* =.050), and for Abdel-Haleem and Ali, with a mean difference of 0.55 (t = -2.190, *p* =.035), but not for Ali and Arberry.

As shown in the above figure, Abdel-Haleem's word usage was skewed to the positive side of the scale for all variables, more so than for the other translators. For both Ali and Arberry, word usage was generally closer to the neutral side of the scale. In summary, Abdel-Haleem's version was rated as using significantly shorter, more familiar, more modern, clearer, and using more common words than those of either Ali or Arberry.

6.2.2.1 Word Usage Judged by Two Groups

Table 6.4 provides another comparative analysis derived from the previous scale, showing the overall means for all five pairs on the word-usage scale, rated by both low-knowledge and high-knowledge readers. The findings show that those with low knowledge found Abdel-Haleem's word usage to be easier (*M* = 1.84, *SD* = 0.27) than Arberry's (*M* = 2.11, *SD* = 0.21) and Ali's (*M* = 2.44, *SD* = 0.29). The findings from the t-test reveal significant differences between Abdel-Haleem's and Ali's translations, with a mean difference of 0.60 (*t* = -2.447, *p* = 0.02), and between Abdel-Haleem's and Arberry's, with a mean difference of 0.27 (*t* = 1.819, *p* = 0.04), but not between Ali's and Arberry's.

Table 6.4: Results for Two Groups of Readers for the Word-Usage Scale

Surveys	Groups	N. of Readers	Mean	S. Deviation	Max.	Min.
Abdel-H.	Low-Knowledge Readers	11	1.84	0.27	1.91	1.50
	High-Knowledge Readers	9	1.80	0.06	1.89	1.72
Arberry	Low-Knowledge Readers	13	2.11	0.21	2.50	1.96
	High-Knowledge Readers	7	2.50	0.57	3.50	1.86
Ali	Low-Knowledge Readers	12	2.44	0.29	2.96	2.08
	High-Knowledge Readers	8	2.25	0.38	3.00	2.00

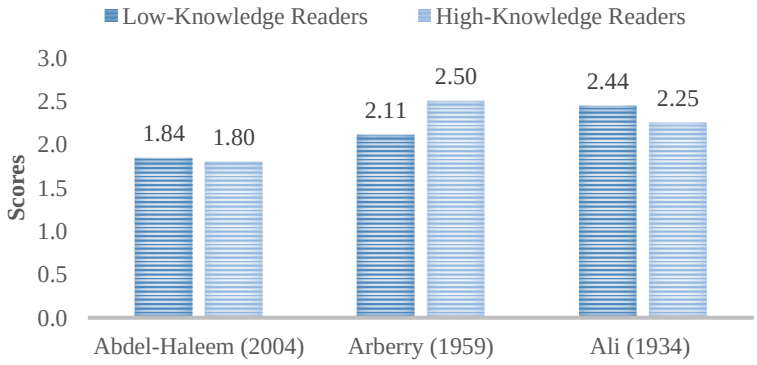


Figure 6.4: Mean Values for Two Groups of Readers on the Word-Usage Scale

In addition, those with high knowledge found the word usage in Abdel-Haleem's translation to be easier (M = 1.80, SD = 0.06) than that in Arberry's (M = 2.50, SD = 0.57) and in Ali's (M = 2.25, SD = 0.38). The t-test found significant differences between Abdel-Haleem's translation and Ali's, with a mean difference of 0.70 (t = 2.480, p = 0.03), and between Abdel-Haleem's and Arberry's, with a mean difference of 0.45 (t = 2.131, p = 0.03), but not between Ali's and Arberry's.

Figure 6.4 illustrates the differences between the responses from those with low knowledge and those with high knowledge for each translation, suggesting that both groups of readers agreed that Abdel-Haleem's words were less 'difficult' than those in the other versions. However, there were some disagreements between the two groups in their ratings of the other two versions' word usage: low-knowledge readers rated Arberry's words as less difficult than high-knowledge readers did, whilst low-knowledge readers rated Ali's words as more difficult than high-knowledge readers did.

6.2.3 Analysis of Sentence Usage

Table 6.5 summarises the judgements of all participants regarding sentence usage in the three translations. As shown in the table, the highest values given by participants in the Abdel-Haleem survey in relation to sentences were for the adjectives *clear–confusing*, with the lowest values for *short–long*. In the Arberry survey, the highest-value participant ratings for sentences were for the adjectives *short–long*, with the lowest for *familiar–strange*. In the Ali survey, the highest values were for *familiar–strange* (in relation to grammar) and the lowest were for *simple–complex* (in relation to sentences).

Table 6.5: Descriptive Statistics for the Pairs on the Sentence-Usage Scale

	Abdel-H. Survey			Arberry Survey			Ali Survey		
Pairs of Adjectives	MED	M	SD	MED	M	SD	MED	Mean	SD
Short–Long sentences	2.00	2.10	0.94	3.00	2.65	1.28	2.00	2.20	1.21
Familiar–Strange grammar	2.00	2.08	1.03	2.00	2.25	1.04	2.00	2.28	0.95
Simple–Complex sentences	2.00	1.90	0.94	2.00	2.60	1.07	2.00	2.05	0.86
Clear–Confusing sentences	2.00	2.05	0.97	2.00	2.50	0.97	2.50	2.45	0.86

Note: MED = Median; M= Mean; and SD = Standard Deviation.

An independent-samples t-test was conducted to compare the mean differences between these four pairs. The mean values for the adjectives *short–long* in relation to sentences reveal that Abdel-Haleem's sentences were rated shorter (M = 2.10, SD = 0.94) than Ali's (M = 2.20, SD = 1.21) and Arberry's (M = 2.65, SD = 1.28). As shown in Figure 6.5 below, for this pair,

the two latter translations were closer to the neutral side of the scale, whereas Abdel-Haleem fell on the neutral side. The result for Abdel-Haleem and Arberry was the only statistically significant difference (t = -2.040, p =.055).

On the *simple–complex* scale (in relation to sentences), Abdel-Haleem's words were rated simpler (M = 1.90, SD = 0.94) than Ali's (M = 2.05, SD = 0.86) and Arberry's (M = 2.60, SD = 1.07). Like to the previous finding, the t-test found the only statistically significant difference to be between Abdel-Haleem and Arberry, with a mean of 0.70 (t = 2.142, p =. 038).

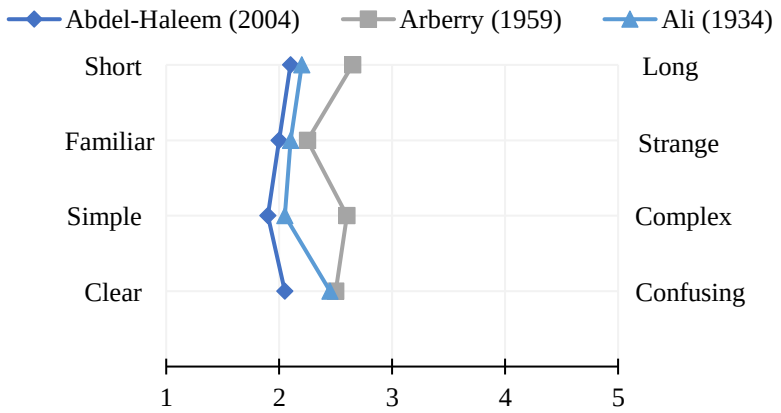


Figure 6.5: Mean Values for the Pairs on the Sentence-Usage Scale

The t-test found no statistically significant differences between the three translations in terms of the ratings on the *familiar–strange* scale in relation to grammar and the *clear–confusing* scale in relation to sentences. For the first pair, the grammatical structures of Abdel-Haleem's translation (M = 2.08, SD = 1.03) were insignificantly rated as more familiar than those of Ali's translation (M = 2.28, SD = 0.95) and Arberry's translation (M = 2.25, SD = 1.04). For the second pair, the sentences of Abdel-Haleem's translation (M = 2.05, SD = 0.97) were rated insignificantly clearer than those of Ali's translation (M = 2.45, SD = 0.86) and of Arberry's translation (M = 2.50, SD = 0.97). As shown in Figure 6.5, all the pairs in Abdel-Haleem's translation were more skewed to the positive side of the scale than those in either Ali's or Arberry's translations.

6.2.3.1 Sentence Usage Judged by Two Groups

Table 6.6 demonstrates the overall mean for all four pairs on the sentence-usage scale, as rated by low-knowledge readers and high-knowledge readers for all three translations. On average, those with low knowledge found Abdel-Haleem's sentence usage to be simpler and clearer (M

= 2.19, *SD* = 0.22) than that of Arberry (*M* = 2.54, *SD* = 0.21) or Ali (*M* = 2.33, *SD* = 0.18). A t-test found a significant difference between Abdel-Haleem's translation and Arberry's, with a mean difference of 0.35 (*t* = 1.987, *p* =.047). No significant differences were found for the ratings of low-knowledge readers between the other translations.

Table 6.6: Results for Two Groups of Readers for the Sentence-Usage Scale

Surveys	Groups	N. of Readers	Mean	S. Deviation	Max.	Min.
Abdel-H.	Low-Knowledge Readers	11	2.19	0.22	2.41	1.86
	High-Knowledge Readers	9	1.85	0.08	1.94	1.72
Arberry	Low-Knowledge Readers	13	2.54	0.21	2.85	2.27
	High-Knowledge Readers	7	2.13	0.14	2.29	1.93
Ali	Low-Knowledge Readers	12	2.33	0.18	2.54	2.04
	High-Knowledge Readers	8	1.91	0.19	2.19	1.69

Similarly, readers with high knowledge found the sentence usage of Abdel-Haleem to be simpler and clearer (*M* = 1.85, *SD* = 0.08) than that of Arberry (*M* = 2.13, *SD* = 0.14) or Ali (*M* = 1.91, *SD* = 0.19). A t-test found a significant difference between Abdel-Haleem's translation and Arberry's, with a mean difference of 0.27 (t=1.819, *p* =.024), but no significant differences were found elsewhere.

Figure 6.6 shows the differences for both groups of readers for all three translations. Those with low knowledge gave higher ratings for the difficulty of all the translations. The two groups of readers were in agreement about the simplicity of Abdel-Haleem's sentence usage, compared to that of Ali and Arberry, suggesting that prior knowledge plays an important role in reading comprehension, but the use of more complex language may not facilitate reading comprehension.

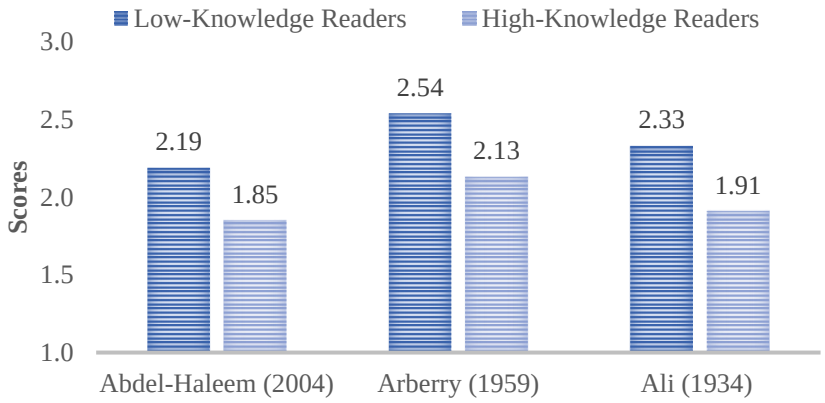


Figure 6.6: Mean Values for Two Groups of Readers on the Sentence-Usage Scale

6.2.4 Analysis of Cohesion Usage

Cohesion usage was the last scale used to assess translation readability. This scale included three pairs of adjectives to rate the cohesive elements of the Qur'ānic sentences. The overall judgements of all participants for these three pairs are summarised in Table 6.7. The overall means for the participants' ratings reveal that Abdel-Haleem's translation uses the most familiar conjunctions, coherent words, and coherent sentences. Ali's translation uses the least familiar conjunctions, whilst Arberry's translation has the least coherent sentences.

Table 6.7: Descriptive Statistics for the Pairs on the Cohesion-Usage Scale

Pairs of Adjectives	Abdel-H. Survey			Arberry Survey			Ali Survey		
	Med	M	SD	MED	M	SD	MED	M	SD
Coherent–Incoherent Words	1.00	1.30	0.56	2.00	1.90	0.70	2.00	1.95	0.92
Coherent–Incoherent Sentences	2.00	1.90	0.70	3.00	2.40	0.92	2.00	1.95	0.92
Familiar–Strange Conjunctions	1.00	1.75	0.94	2.00	2.05	0.67	2.00	2.25	0.83

Note: MED = Median; M= Mean; and SD = Standard Deviation.

An independent-samples t-test was conducted to further interpret the significant differences between the three translations for the three pairs of adjectives. The t-test showed statistical differences between the three translations for one of the three pairs: namely, *coherent–incoherent* in relation to words. Abdel-Haleem's words were rated more coherent (M = 1.30, SD = 0.56) than Arberry's (M = 1.90, SD = 0.70) and Ali's (M = 1.95, SD = 0.92). The t-test indicates significant differences between Abdel-Haleem and Arberry (t= -2.924, p =.006) and between Abdel-Haleem and Ali (t=-2.633, p =.012), but not between Ali and Arberry (p =.797). Abdel-Haleem's words were largely rated 'coherent', whilst Arberry's and Ali's ratings were inclined slightly towards 'incoherent', as shown in Figure 6.7.

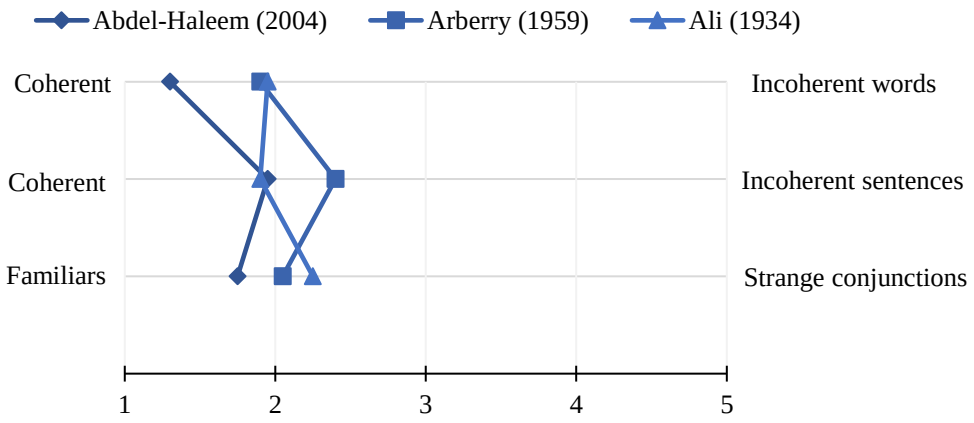


Figure 6.7: Mean Values for the Pairs on the Cohesion-Usage Scale

6.2.4.1 Cohesion Usage Judged by Two Groups

Table 6.8. below presents another comparative analysis of the previous scale, showing the overall mean for all three pairs for the cohesion-usage scale, as rated by low-knowledge readers and high-knowledge readers. An independent-samples t-test was conducted to identify the mean differences between the ratings of those with low knowledge and those with high knowledge for each translation.

Table 6.8: Results for Two Groups of Readers for the Cohesion Usages Scale

Surveys	Groups	N. of Readers	Mean	S. Deviation	Max.	Min.
Abdel-H.	Low-Knowledge Readers	11	1.88	0.31	2.14	1.45
	High-Knowledge Readers	9	1.63	0.07	1.72	1.56
Arberry	Low-Knowledge Readers	13	2.08	0.14	2.27	1.96
	High-Knowledge Readers	7	1.95	0.18	2.14	1.71
Ali	Low-Knowledge Readers	12	2.07	0.02	2.08	2.04
	High-Knowledge Readers	8	1.94	0.16	2.13	1.75

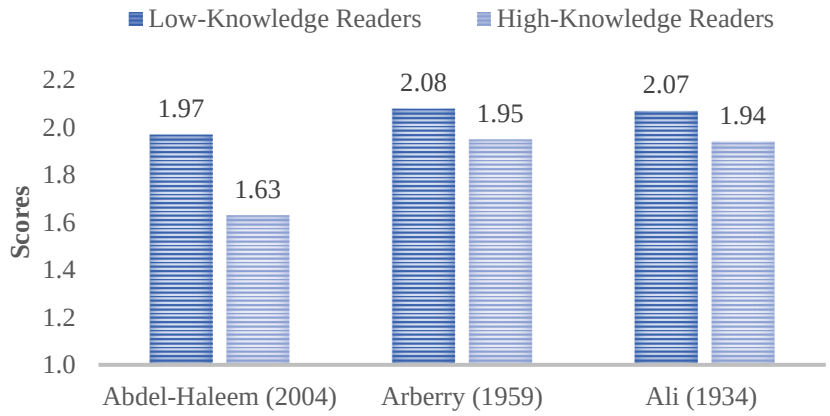


Figure 6.8: Mean Values for Two Groups of Readers for the Cohesion-Usage Scale

Abdel-Haleem's translation (*M* = 1.88, *SD* = 0.31) was rated by low-knowledge readers as more coherent than Arberry's translation (*M* = 2.08, *SD* = 0.14) or Ali's translation (*M* = 2.07, *SD* = 0.02). The findings from the t-test suggest no statistically significant differences between the ratings of those with low knowledge for any of the translations. In contrast, high-knowledge readers rated Abdel-Haleem's translation as more cohesive (*M* = 1.63, *SD* = 0.07) than Arberry's translation (*M* = 1.95, *SD* = 0.18) and Ali's translation (*M* = 1.94, *SD* = 0.16). As with the previous finding, the t-test found no statistically significant differences between the ratings of those with high knowledge for any of the translations. To conclude, the two groups of readers found no differences between the cohesive elements of the three Qur'ānic versions.

Low-knowledge readers tended to give higher ratings for all translations, whilst high-knowledge readers generally gave lower ratings, as shown Figure 6.8.

6.2.5 Overall Level of Readability

All human ratings of the four text factors were combined to obtain a composite variable for the readability score of each version. Composite scoring is a common procedure, extracting a total from several variables in a multi-item scale (Dörnyei, 2007). This composite score was taken to obtain both the overall level of readability for each translation (as shown in Table 6.9 below) and to correlate readability with comprehension (see section 6.3.3).

Table 6.9: Readability Scores for Each Translation

Translators	Mean	S. Deviation	Max.	Min.
Abdel-Haleem	1.85	0.27	2.25	1.05
Arberry	2.09	0.37	2.85	1.15
Ali	2.03	0.38	3.00	1.30

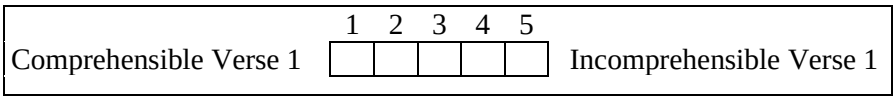
As can be seen from the table, the range of the readability scores for Abdel-Haleem's translation is narrow, with a mean of 1.85, a minimum of 1.05, and maximum of 2.25. Ali's scores have a mean of 2.03, a minimum of 1.30, and a maximum of 3.00, whilst Arberry's have a mean of 2.09, a minimum of 1.15, and a maximum of 2.85. A t-test suggests that the mean difference between the readability scores is significant for Abdel-Haleem and Arberry (t = -4.085, df = 35, p =.000) and for Abdel-Haleem and Ali (t = -2.646, p =.012), but not for Arberry and Ali (t = 1.035, p =.308). Abdel-Haleem's translation scored higher for readability than the other translations, as his translation was significantly rated better for most of the text factors (i.e., word usage, sentence usage, and cohesion usage), whilst four elements of the page layout were rated better for Ali's translation. However, three elements of the page layout of Ali's translation were not statistically significant. In short, Abdel-Haleem's translation was the most readable and Arberry's translation the least.

6.3 Part Two: Human Judgements of Translation Comprehensibility

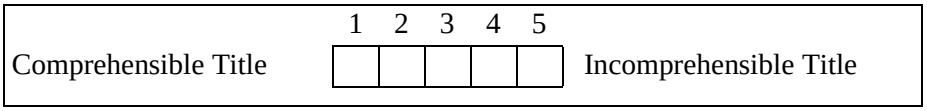
After readability had been judged, two complete Qur'ān chapters were evaluated based on four primary aspects of translation comprehension, as listed below. Each of the aspects was designed using a SD scale, from 1 to 5, and included in two of the surveys for evaluation – Arberry's was the exception, as no footnotes or introductions were originally provided for his

translated chapters. The results for each of the four aspects are reported in the following sections.

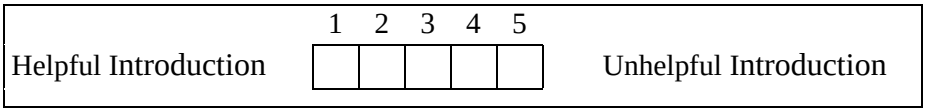
- Comprehension ease on a verse-by-verse basis was rated as follows:



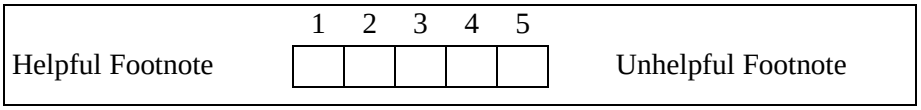
- The title of the Qur'ān chapter was rated as follows:



- The influence of a provided introduction on text comprehension was rated as follows:



- The influence of a given footnote on text comprehension was rated as follows:



6.3.1 Analysis of Verse Comprehension

The 20 participants in each survey were each asked to read two whole chapters, containing a total of 32 verses. The judges rated the difficulty of comprehension of each verse, based on the SD scale, from 1 to 5: with '1' being (fully) comprehensible and '5' incomprehensible (thus, a lower score reflects greater comprehensibility). Table 6.10 displays the mean scores for all 20 participants' ratings for each verse of the three translations.

Table 6.10: Mean Values of Each Verse in Each Translation

	Abdel-H. Survey	Arberry Survey	Ali Survey
Chapter 91	Mean	Mean	Mean
Verse 1	1.25	1.55	1.35
Verse 2	1.45	1.55	1.50
Verse 3	1.40	1.95	1.65
Verse 4	1.35	1.80	1.40
Verse 5	1.65	1.60	2.10
Verse 6	1.85	1.90	1.40
Verse 7	1.40	1.80	1.55

Verse 8	1.80	2.35	1.50
Verse 9	1.40	1.70	1.70
Verse 10	1.35	1.85	1.60
Verse 11	1.45	2.50	1.75
Verse 12	1.70	1.85	2.30
Verse 13	1.90	1.95	2.00
Verse 14	2.25	2.30	1.90
Verse 15	1.60	2.15	2.35
Chapter 86			
Verse 1	2.15	1.45	2.10
Verse 2	2.20	1.75	2.05
Verse 3	1.90	1.85	1.60
Verse 4	1.85	1.60	1.85
Verse 5	1.45	1.55	1.60
Verse 6	1.55	1.65	2.00
Verse 7	1.85	2.00	1.65
Verse 8	1.45	2.15	1.50
Verse 9	1.70	1.90	1.50
Verse 10	1.45	1.80	1.40
Verse 11	1.40	2.15	2.30
Verse 12	1.50	2.50	1.70
Verse 13	1.70	2.30	1.60
Verse 14	1.65	2.50	1.60
Verse 15	1.50	2.75	1.70
Verse 16	1.55	2.70	1.60
Verse 17	1.65	2.00	1.70
Overall Mean	1.63	1.98	1.73
S. Deviation	0.25	0.35	0.28

An initial reading of the data suggests that the comprehension difficulty of Abdel-Haleem's verses is of a narrower range, with an overall mean of 1.63, a minimum of 1.25, and a maximum of 2.25. In contrast, Ali's scores have an overall mean of 1.73, a minimum of 1.35, and a maximum of 2.35, whilst Arberry's have an overall mean of 1.98, a minimum of 1.45, and a maximum of 2.75.

Table 6.10 provides the overall mean and standard deviation of the 32 verses in each translation. Abdel-Haleem's translated verses were perceived to be more comprehensible (M = 1.63, SD = 0.25) than Ali's (M = 1.73, SD = 0.28) or Arberry's (M = 1.98, SD = 0.35). An independent-samples t-test was conducted to find the mean differences between the ratings for comprehension difficulty for all verses in all three translations. The independent t-test found significant differences between Abdel-Haleem's and Arberry's scores (t = -4.424, df = 62, p =.000) and between Arberry's and Ali's (t = 3.015, df = 62, p =.005), but not between Abdel-Haleem's and Ali's (t = -1.848, df = 62, p =.141). Figure 6.9 illustrates these differences and indicates no outliers.

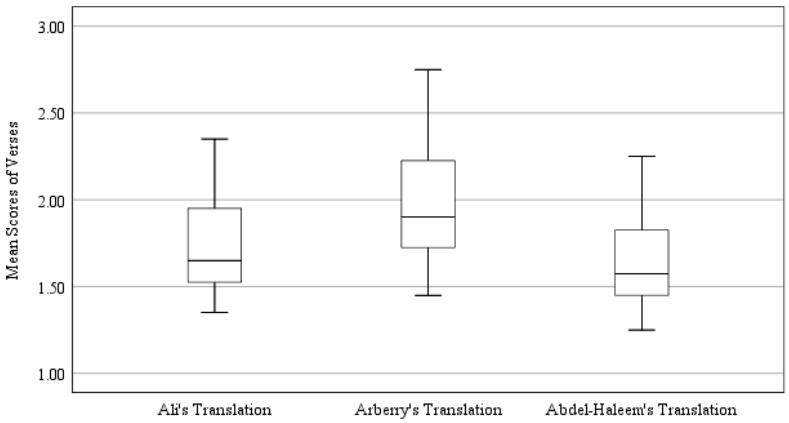


Figure 6.9: Overall Means for all Verses in Each Translation

6.3.2 Verse Comprehension Judged by Two Groups

Table 6.11 shows the overall means of the comprehension scores for the 32 verses, as rated by low-knowledge and high-knowledge readers. The readers with low knowledge gave Abdel-Haleem's translated verses higher scores for comprehensibility (M = 1.70, SD = 0.34) than Ali's (M = 1.99, SD = 0.37) or Arberry's (M = 2.04, SD = 0.14). The findings from the independent-samples t-test show that the (p) values were statistically significantly different for Abdel-Haleem and Arberry (t = -4.103, df = 62, p =.000) and for Abdel-Haleem and Ali (t = -4.163, df = 62, p =.000), but not for Ali and Arberry (t =.573, df = 62, p =.571). This indicates that those with low knowledge benefited significantly more from the higher readability of Abdel-Haleem's version than from the lower readability of other earlier versions.

Table 6.11: Results for Verse Comprehension Rated by Two Groups of Readers

Surveys	Groups	N. of Readers	Mean	S. Dev.	Max.	Min.
Abdel-H.	Low-Knowledge Readers	11	1.70	0.34	2.64	1.18
	High-Knowledge Readers	9	1.56	0.21	2.00	1.22
Arberry	Low-Knowledge Readers	13	2.04	0.32	2.64	1.54
	High-Knowledge Readers	7	1.87	0.50	3.00	1.00
Ali	Low-Knowledge Readers	12	1.99	0.37	3.00	1.50
	High-Knowledge Readers	8	1.35	0.21	1.88	1.13

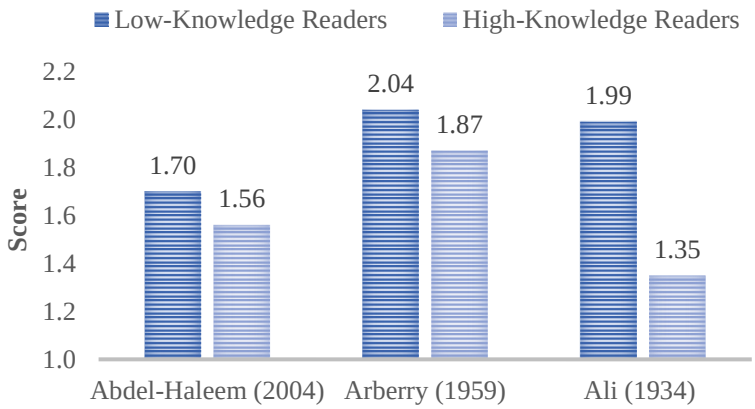


Figure 6.10: Overall Means for Verse Comprehension for Two Groups of Readers

In contrast, readers with high knowledge found Ali's translated verses to be more comprehensible (M = 1.35, SD = 0.21) than Abdel-Haleem's (M = 1.56, SD = 0.21) and Arberry's (M = 1.87, SD = 0.50). The findings of the t-test show that the (p) values were statistically significantly different for Abdel-Haleem and Arberry (t = -3.153, df = 31, p =.004), Abdel-Haleem and Ali (t = 3.968, df = 31, p =.000), and Ali and Arberry (t = 5.121, df = 31, p =.000). This indicates that those with high knowledge benefited more from Ali's translation.

Ali's translation, without doubt, is the most common and most widely distributed English translation of the Qur'ān in the world. Thus, those with high knowledge might be familiar with Ali's translation, though it is less readable than Abdel-Haleem's according to both the human ratings in this chapter and the automatic evaluation presented in Chapter 5. In this analysis, both high-knowledge and low-knowledge readers found Arberry's text to be the least comprehensible. Figure 6.10 displays the results for these two groups of readers for each translation, showing that those with high knowledge gave each of the translations lower scores, whilst those with low knowledge gave each translation higher scores.

6.3.3 Correlational Analysis Between Readability and Comprehension

A correlation analysis using Pearson's correlation coefficient was carried out. The results indicate a significant relationship between human readability scores and human comprehension scores for Abdel-Haleem's translation (r =.786, p =.000), Arberry's (r =.344, p =.030), and Ali's (r =.678, p =.000). The correlation coefficient shows a strong positive correlation for Abdel-Haleem, a moderate positive correlation for Ali, and a weak positive correlation for Arberry. The scatterplots in Figure 6.12 illustrate the linear relationship between readability and comprehension scores for all translations. All three suggest a significant relationship

between the two variables, indicating that when readability is increased, text comprehension is also increased, and vice versa. For example, the readability level of Abdel-Haleem's translation was shown to be high in the previous sections, and the ease of comprehension of his translated verses was higher than that of the other versions.

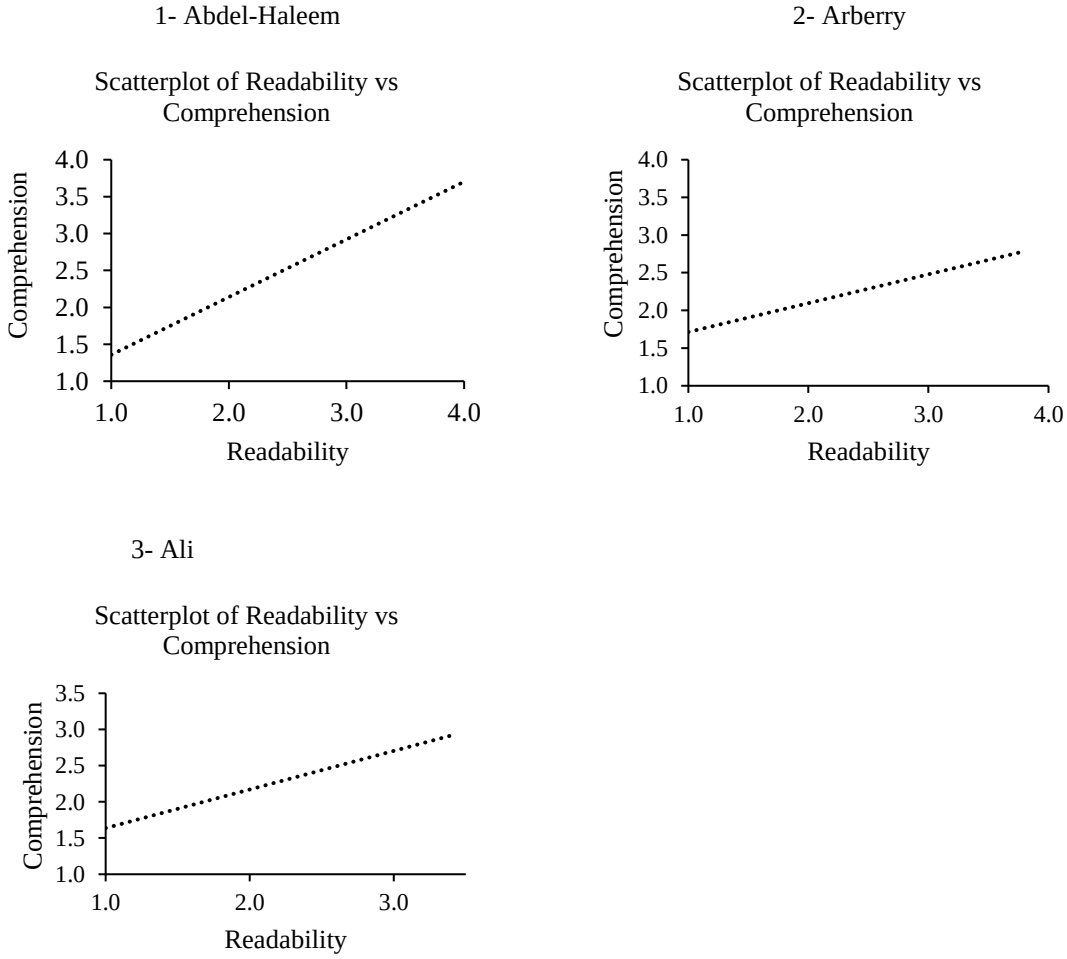


Figure 6.11: Scatterplots for Readability and Comprehension Scores for All Translations

6.3.4 Analysis of Comprehension Aids

The participants were asked to rate each translation in relation to three comprehension aids: chapter title, chapter introduction, and footnotes. All comprehension aids were included in Abdel-Haleem's and Ali's surveys, but not Arberry's, as this translation did not originally include any footnotes or introductions for the translated chapters. The primary purpose of this analysis was to assess the utility for text comprehension of providing supplementary

information. Table 6.12 presents the descriptive statistics for the comprehension aids, which are then discussed in turn.

Table 6.12: Results of the elements of Comprehension Aids

Translator	Elements	Mean	S. Deviation	Max.	Min.
Abdel-Haleem	Title	1.80	1.12	5.00	1.00
	Introduction	1.63	0.73	3.00	1.00
	Footnote	2.00	1.00	4.00	1.00
Ali	Title	1.73	1.07	4.00	1.00
	Introduction	1.78	1.04	5.00	1.00
	Footnote	1.88	0.93	4.00	1.00
Arberry	Title	1.40	0.73	4.00	1.00

An independent-samples t-test was used to determine whether there was a statistically significant difference between the means for comprehensibility of the chapter titles. There was a decrease in comprehensibility from Abdel-Haleem's titles (M = 1.80, SD = 1.12), to Ali's titles (M = 1.73, SD = 1.07), to Arberry's titles (M = 1.40, SD = 0.73). However, these decreases were not statistically significant, indicating that comprehensibility was similar for all translations.

A one-sample t-test was conducted to determine whether the two comprehension aids (the use of introductions and footnotes in Ali's and Abdel-Haleem's translations) enhanced text comprehension. For the first element, most participants found the use of introductions in Abdel-Haleem's translation (M = 1.63, SD = 0.73) and in Ali's translation (M = 1. 78, SD = 1.04) to be useful for text comprehension, with a statistically significant mean difference: t (39) = 13.882, p =.000 for Abdel-Haleem's introductions and t (39) = 10.694, p =.000 for Ali's introductions. For the second element, most participants found the use of footnotes in Abdel-Haleem's translation (M = 2.00, SD = 1.00) and in Ali's translation (M = 1.88, SD = 0.93) to be useful for text comprehension, with a statistically significant mean difference: t (39) = 12.490, p =.000 for Abdel-Haleem's footnotes, and t (39) = 12.631, p =.000 for Ali's footnotes. In summary, most participants suggested that these two comprehension aids were very useful for text comprehension. Thus, the lack of comprehension aids might be one of many factors reducing the comprehensibility of Arberry's translation.

6.3.5 Regression Analysis for Automatic Evaluation Measures

To determine which linguistic features predicted human judgements of verse comprehensibility, the final section of the analysis involved a multiple regression analysis. This analysis used all 29 Coh-Metrix indices and five readability formulae (in Chapter 5) as predicator variables to determine which linguistic features identified by these two automated measures were most predictive of comprehension difficulty. This study followed the common practice of recent studies in using a regression analysis as a statistical technique to identify which linguistic features are more predicative of text quality and comprehensibility (see Crossley et al., 2017; McNamara et al., 2010; Crossley and McNamara, 2011).

The reading ease of the chosen translations was measured using five readability formulae (i.e., the fog index, SMOG index, Dale-Chall formula, Coleman-Liau index, and Flesch-Kincaid grade level; see Chapter 5). A multiple regression analysis was conducted using the five formulae as independent variables to identify which of the five chosen readability formulae was the most predictive of comprehension difficulty. These five formulae were checked for outliers and multicollinearity. Collinearity and non-significant predictors were found in three formulae (the fog index, SMOG index and Flesch-Kincaid grade level), which were thus excluded from the regression model analysis. The regression analysis found that two formulae (the Coleman-Liau index and Dale-Chall formula) predicted Qur'ān translation comprehension difficulty with statistical significance: F (2, 117) = 25.662, p <.0005. These two formulae yielded a significant model, R =.305, R^2 =.293, indicating that the Coleman-Liau index and Dale-Chall formula together explained 29% of the variance in Qur'ān translation comprehension difficulty. The Coleman-Liau index was a significant predictor of Qur'ān translation comprehension difficulty (t = 2.702, p =.008), as was the Dale-Chall formula (t = 3.511, p =.001). The results of this first regression indicate that the Dale-Chall formula and Coleman-Liau index are important indicators of Qur'ān translation comprehension difficulty.

The word difficulty of the chosen translations was measured using six Coh-Metrix indices (see Chapter 5). A multiple regression analysis was conducted for the following six variables: word length, word hypernym, CELEX word frequency, word concreteness, word imageability, and word meaningfulness. These variables were checked for outliers and multicollinearity. The regression analysis found that the six variables statistically significantly predicted Qur'ān translation comprehension difficulty: F (6, 113) = 22.806, p <.0005. These six variables yielded a significant model, R =.548, R^2 =.524, indicating that the combined variables explained 52% of the variance in Qur'ān translation comprehension difficulty. Non-significant predictors were found in two variables (i.e., word concreteness and word imageability). The four most

significant predictive variables from the Coh-Metrix for comprehension difficulty were word length (t = 4.627, p =.000); word hypernym (t = -3.677, p =.000); CELEX word frequency (t = -1.992, p =.049); and word meaningfulness (t = -7.097, p =.000). Excluding the two non-significant predictors, these four variables yielded a significant model, F (4, 115) = 22.965, p <.0005, R =.444, $R2$ =.425, and accounted for 42% of the most predictive features of comprehension difficulty.

The syntactic complexity of the chosen translations was measured using seven Coh-Metrix indices (see Chapter 5). A multiple regression analysis was conducted using the following seven variables in the model: sentence length, number of words before the main verb, number of modifiers per noun phrase, density of logical operators, density of noun phrase, density of verb phrase, and density of adverbial phrase. These variables were checked for outliers and multicollinearity. The regression analysis using these seven variables yielded a significant model, F (7, 112) = 11.405, p <.0005, R =.416, R^2 =.380, indicating that the combined variables explained 38% of the variance in comprehension difficulty. Non-significant predictors were found in five variables. The two most significant predictive features were density of logical operators (t = 4.562, p =.000) and of noun phrases (t = 5.265, p=.000). Excluding five non-significant predictors, the two variables above yielded a significant model, F (2, 117) = 35.752, p <.0005, R =.379, $R2$ =.369, and accounted for 36% of the variance in comprehension difficulty. In this third regression model, the density of logical operators and noun phrases were significant indicators of Qur'ān translation comprehensibility.

Finally, the cohesive devices of the three translations were measured using 16 Coh-Metrix indices (see Chapter 5). A multiple regression analysis using the 16 variables as independent variables was conducted to distinguish which of these 16 cohesive variables was the most predictive of comprehension difficulty. The outliers and multicollinearity of these variables were checked. The regression analysis using the 16 variables yielded a significant model, F (16, 103) = 10.629, p <.0005, R =.623, R2 =.564, indicating that the combined variables explained 56% of the variance in comprehension difficulty. Non-significant predictors were found in 11 variables. The four most significant predictive variables of comprehension difficulty were density of noun overlap (t = 3.662, p=.000); of stem overlap (t = 8.086, p=.000); of argument overlap (t = 3.286, p=.001); and of 'all connectives' (t = 6.062, p=.000). Excluding the 11 non-significant predictors, these four latter variables yielded a significant model, F (4, 115) = 32.934, p <.0005, R =.534, $R2$ =.518, and accounted for 51% of the variance in

comprehension difficulty. Repetition aspects and connectives were the least strong indicators in our regression models.

6.4 Discussion of Quantitative Findings

The previous chapter responded to the first two of the five research questions. This section discusses the results of human judgements in relation to the final three. The five research questions are as follows:

1. What are the primary textual factors that could affect the readability and comprehensibility of Qur'ān translations, as indicated by automated measures?
2. Does retranslation result in higher levels of readability and comprehensibility?
3. Which of the three translations under study is the most readable and comprehensible, and which is the least readable and comprehensible, as indicated by human judgement?
4. Do human judgements of readability and comprehensibility correlate with the automated measures and yield consistent findings?
5. Do readers with low (or high) levels of prior knowledge of the Qur'ān gain more understanding from lower (or higher) levels of readability?

For the third question, the results suggested that Abdel-Haleem's version was judged the most readable and comprehensible, whilst Arberry's version was judged the least readable and comprehensible, according to the subjective judgements of readability and comprehension. In response to the forrth question, a strong correlation was observed between the human readability ratings and comprehension ratings. Comparisons of the subjective and objective measures of readability and comprehension levels produced convergent results (see Chapter 5). In response to the last question, it is noted that a reduction in the cohesion and complexity of the language might facilitate translation comprehension for readers, especially those with little background knowledge in the Qur'ān. These three questions are discussed further in the following sections in light of seven points emerging from analysis of the human judgements and automated measures presented in Chapter 5.

6.4.1 Judges of Translation Readability

The readability of each version was judged based on four SD ratings of the text factors: page layout, word, sentence, and cohesion. Each will be discussed in turn in this section.

1. Abdel-Haleem (2004)

91. THE SUN

A Meccan sura, the central theme of which is purifying or corrupting the soul, with the tribe of Thamud given as an example of corruption.

In the name of God, the Lord of Mercy, the Giver of Mercy

[1]By the sun in its morning brightness [2]and by the moon as it follows it, [3]by the day as it displays the sun's glory [4]and by the night as it conceals it, [5]by the sky and how He built it [6]and by the earth and how He spread it, [7]by the soul and how He formed it [8]and inspired it [to know] its own rebellion and piety! [9]The one who purifies his soul succeeds [10]and the one who corrupts it fails. [11]In their arrogant cruelty, the people of Thamud[a] called [their messenger] a liar, [12]when the most wicked man among them rose [against him].[b] [13]The messenger of God said to them, '[Leave] God's camel to drink,' [14]but they called him a liar and hamstrung her. Their Lord destroyed them for their crime and levelled them. [15]He did not hesitate[c] to punish[d] them.

[a] See e.g. 7: 73–9 (on the tribe of Thamud).
[b] Cf. 54: 29.
[c] Literally 'he does not fear'.
[d] One of the lexical meanings of *'uqba* is *jaza'*, here 'to punish'; or 'does not fear the

2. Arberry (1959)

XCI

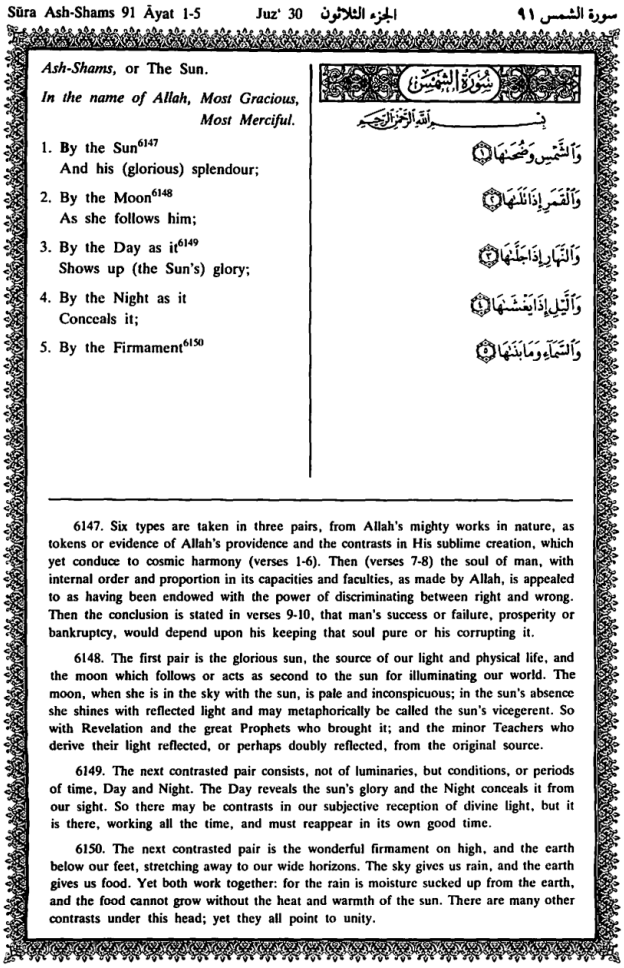
THE SUN

In the Name of God, the Merciful, the Compassionate

By the sun and his morning brightness
and by the moon when she follows him,
and by the day when it displays him
and by the night when it enshrouds him!
By the heaven and That which built it
and by the earth and That which extended it!
By the soul, and That which shaped it
and inspired it to lewdness and godfearing!
Prosperous is he who purifies it,
and failed has he who seduces it.

Thamood cried lies in their insolence
when the most wretched of them uprose,
then the Messenger of God said to them,
'The She-camel of God; let her drink!'
But they cried him lies, and hamstrung her,
so their Lord crushed them for their sin, and levelled them:
and He fears not the issue thereof.

3. Ali (1934-1989)



Sūra Ash-Shams 91 Āyat 1-5 Juz' 30 الجزء الثلاثون سورة الشمس ٩١

Ash-Shams, or The Sun.

In the name of Allah, Most Gracious, Most Merciful.

1. By the Sun[6147]
And his (glorious) splendour;
2. By the Moon[6148]
As she follows him;
3. By the Day as it[6149]
Shows up (the Sun's) glory;
4. By the Night as it
Conceals it;
5. By the Firmament[6150]

سُورَةُ الشَّمْسِ

بِسْمِ اللَّهِ الرَّحْمَٰنِ الرَّحِيمِ

وَالشَّمْسِ وَضُحَاهَا ﴿١﴾

وَالْقَمَرِ إِذَا تَلَاهَا ﴿٢﴾

وَالنَّهَارِ إِذَا جَلَّاهَا ﴿٣﴾

وَاللَّيْلِ إِذَا يَغْشَاهَا ﴿٤﴾

وَالسَّمَاءِ وَمَا بَنَاهَا ﴿٥﴾

6147. Six types are taken in three pairs, from Allah's mighty works in nature, as tokens or evidence of Allah's providence and the contrasts in His sublime creation, which yet conduce to cosmic harmony (verses 1-6). Then (verses 7-8) the soul of man, with internal order and proportion in its capacities and faculties, as made by Allah, is appealed to as having been endowed with the power of discriminating between right and wrong. Then the conclusion is stated in verses 9-10, that man's success or failure, prosperity or bankruptcy, would depend upon his keeping that soul pure or his corrupting it.

6148. The first pair is the glorious sun, the source of our light and physical life, and the moon which follows or acts as second to the sun for illuminating our world. The moon, when she is in the sky with the sun, is pale and inconspicuous; in the sun's absence she shines with reflected light and may metaphorically be called the sun's vicegerent. So with Revelation and the great Prophets who brought it; and the minor Teachers who derive their light reflected, or perhaps doubly reflected, from the original source.

6149. The next contrasted pair consists, not of luminaries, but conditions, or periods of time, Day and Night. The Day reveals the sun's glory and the Night conceals it from our sight. So there may be contrasts in our subjective reception of divine light, but it is there, working all the time, and must reappear in its own good time.

6150. The next contrasted pair is the wonderful firmament on high, and the earth below our feet, stretching away to our wide horizons. The sky gives us rain, and the earth gives us food. Yet both work together: for the rain is moisture sucked up from the earth, and the food cannot grow without the heat and warmth of the sun. There are many other contrasts under this head; yet they all point to unity.

Figure 6.12: Examples of the Page Layouts Presented in the Surveys

1. Page layout. The first SD scale employed six pairs of adjectives to obtain contrasting opinions from the judges regarding the organisation and clarity of the three Qur'ānic page layouts (see Figure 6.12 above). The analysis of page layout involved the assessment of six main elements: *chapter title*, *chapter length*, *paragraph division*, *verse numbers*, *text layout*, and *text organisation*. The results showed statistical differences in the ratings of the chapter title, chapter length, and verse numbers. The judges perceived the chapter titles in Abdel-Haleem's pages to be more easily recognisable than those in the other translations. This is because Abdel-Haleem's and Arberry's titles are placed in the centre of the pages. This position is more prominent than that in Ali's text, where the titles are positioned on the left side of the page, with an Arabic transliteration.

Another significant result regarding page layout was the design of verse numbers in the three versions. The judges' ratings indicated that Ali's numbering list of verses was clearer to readers than those of the other translations. Abdel-Haleem's verse numbers are written in small superscript at the end of each verse (similar to footnote numbering), whilst Arberry's verse numbers are presented in the format of poetry line numbers. The judges considered Arberry's design to be less clear than Abdel-Haleem's, although it might more confusing whether Abdel-Haleem's verse numbers refer to a footnote or a verse number.

The last significant result regarding page layout was for chapter length. Abdel-Haleem's chapter was judged to be shorter than that of Ali and Arberry. This might indicate that readers prefer shorter translations. However, no substantial differences were found between the other three elements of page layout (familiarity of paragraph division, clarity of text organisation, and simplicity of text layout). The judges generally preferred Ali's page layout in terms of its paragraph division, text layout, and text organisation. Further study is needed to examine these variables and other elements of page layout in other Qur'ān translations and identify their importance for readability.

2. Word usage. Significant differences between the three translations were found for four pairs of adjectives in the word usages scale, labelled *familiar-strange*, *modern-archaic*, *short-long*, and *clear-confusing*. All the judges described Abdel-Haleem's translation as using shorter, more familiar, more modern, and less confusing words than Ali's or Arberry's translations, which each had a tendency towards longer, stranger, more archaic, and confusing words. When the judges' ratings for these variables in Arberry's and Ali's versions were compared, no substantial differences were found.

3. Sentence usage. There were significant differences between the three translations for two variables of the SD sentence scale labelled *short-long* and *simple-complex*. For both variables, the sentences in Abdel-Haleem's version were judged to be shorter and simpler than those in the other versions. No substantial differences were found between Ali's and Arberry's sentences, with both judged to be using longer and more complex sentences than those of Abdel-Haleem. This similarity between Ali and Arberry was previously reported by the judges in relation to word usage.

4. Cohesion usage. The last SD scale employed three pairs of adjectives to obtain the participants opinions about the cohesion of the three translations. The results for one pair, *coherent–incoherent*, were found to be statistically significant for the three translations. In this pair, Abdel-Haleem's translation received the lowest ratings, indicating that the words were more coherent than those of other translations, according to the judges. This finding might suggest that the high cohesion in Ali's and Arberry's translations that was measured by all indices of Coh-Metrix in Chapter 5 was significantly affected by the cohesive patterns of the source text.

6.4.2 Prediction of Readability

All the previous human ratings of text factors were combined to reach a prediction of the overall readability of each version. Taken together, the results showed that Abdel-Haleem's version was the most readable and Arberry's the least. This is because Abdel-Haleem's translation was judged to be better on three text factors (i.e., word, sentence, and cohesion), whilst the page layout was rated better for Ali's translation on four variables (although three of these four variables were not significantly different). This result might suggest that the word and sentence variables contribute more than page layout variables to a prediction of readability. Gray and Leary (1935) support this finding. Gray and Leary's large survey examined possible contributors to text readability and found four text factors (in order of importance): content, style, format, and organisation. Finally, the word variables had higher weighting in readability levels than did the sentence and cohesion variables. This was due to the significant and larger differences between the three versions at the word level.

6.4.3 Judges of Translation Comprehensibility

The comprehensibility of each version was judged on that of the verses and titles in two chapters, as well as an assessment of the influence of the introduction and footnotes on text comprehensibility. Each will be discussed in turn in this section.

1. Verse comprehension. The comprehensibility of 32 verses in the three translations was compared. The human judgements indicated that the average ratings of most verses for all three translators were on the comprehensible side of the SD scale, but they differed in terms of the level of comprehensibility. The 32 verses of Abdel-Haleem's version were evaluated by all the judges to be the most comprehensible and Arberry's the least. The only statistically significant result was for Abel-Haleem and Arberry.

When the judgements of 32 verses were analysed according to the reader factor (previous knowledge of the Qur'ān), the judges – who differed in their levels of prior knowledge – gave different results for the comprehensibility of the versions. The results showed that low-knowledge readers gained significantly more understanding from Abdel-Haleem's version, which has higher readability, than either Ali or Arberry's versions, which have lower readability. By contrast, high-knowledge readers gained significantly more understanding from Ali's version than from the others. This finding is supported by previous research (Klare, 1976; Fass and Schumacher, 1978; Entin and Klare, 1980; 1985). For example, Entin and Klare (1985) found that texts with high readability are more beneficial for those with limited knowledge and low interest in the subject than for those with high knowledge and high interest.

The greater information-gain in Abdel-Haleem's version is likely to be affected by readability changes. The greater information-gain in Ali's version is likely to be influenced by the reader's prior knowledge, although Ali's version was assessed by automated measures and rated by two groups of judges to be higher in text difficulty than Abdel-Haleem's. The two groups of judges rated Arberry's version as less comprehensible than Ali's, although the average readability difference between Ali and Arberry was unchanged (see Chapter 5). The comprehensibility difference between Arberry and Ali might be affected by factors related to retranslation (see section 6.4.8) and use (or non-use) of comprehension aids, as explained below.

2. Comprehension aids. The human judgements showed that the comprehensibility of the three versions' chapter titles did not differ. This makes sense, as the three translators use similarly literal translations for these. With regards to comprehension aids, the two groups of judges suggested that the use of explanatory footnotes and introductions were very beneficial for text comprehensibility (both of which are used in Ali's and Abdel-Haleem's translations).

The absence of comprehension aids in Arberry's version is another factor associated with translation difficulty.

3. Correlation between readability and comprehension. The effect of verse readability on the judges' comprehension was measured by a correlational analysis (see section 6.3.3). The results suggest a linear relationship between the judges' ratings of the difficulty of the verses and the human ratings of verse comprehension for each version. These correlations were not alike for all of the versions. Abdel-Haleem's version showed a very high correlations between ratings of subjective readability and comprehension, whilst Arberry's version had very low correlations. This occurred because Abdel-Haleem's version was higher in readability and comprehension than the others. By contrast, Arberry's version was low in readability and comprehension. This provides evidence that, as the level of translation readability increases, the degree of translation comprehension also increases, and vice versa.

6.4.4 Agreement Among Groups of Judges

The human judgements of the judges were compared to obtain a consensus on the readability of each version for low-knowledge and high-knowledge readers. Low-knowledge readers and high-knowledge readers both rated the overall design of Ali's page higher than those of the other pages. Though this comparison did not reveal substantial differences, the decorative elements of Ali's page layout may draw more of the readers' attention. The page layouts of Abdel-Haleem and Arberry have more white space, no Arabic scripts, no Arabic frames, and no dense annotations (see Figure 6.12).

The two groups of judges also agreed that Abdel-Haleem's version had less difficult words, less complex sentences, and the most coherent words and sentences. This was a significant result when compared with the other versions, but not in the ratings for cohesion. Comparing Ali's and Arberry's versions, the two groups of judges gave them higher ratings for complex sentences and lower ratings for cohesion. However, they disagreed in their ratings of word usage. Low-knowledge readers judged Arberry's words to be of lower difficulty than Ali's words, whilst high-knowledge readers judged Arberry's words to be of higher difficulty than Ali's. This disagreement – and even the previous agreements of the judges between Ali and Arberry – did not show statistical significance. To conclude, human judgements might be affected by a translation's characteristics and the reader's prior knowledge. This is explained in the following discussion.

6.4.5 Direction of Ratings Between Groups of Judges

The comparison between the two groups of judges revealed a consistent direction in the ratings of text readability and comprehension. The following patterns of human judgements were found:

- low-knowledge readers frequently gave the translations higher scores for text difficulty and thus lower scores for text comprehension (see Figure 6.13 below), and
- high-knowledge readers frequently gave the translations lower scores for text difficulty and thus higher scores for text comprehension.

When the first group of judges compared between versions, they better understood the more readable version, namely Abdel-Haleem's. Readers with little background knowledge of the Qur'ān or Qur'ān translations might gain more information from a version with less difficult text factors.

On the other hand, the comparison made by the second group produced slightly varied results. High-knowledge readers gained more information from Abdel-Haleem's version because it was less difficult than Arberry's. This result was not consistent when Abdel-Haleem's version was compared with Ali's version: those with high knowledge better understood Ali's version than Abdel-Haleem's, although they judged Ali's version as having a higher level of text difficulty than Abdel-Haleem's.

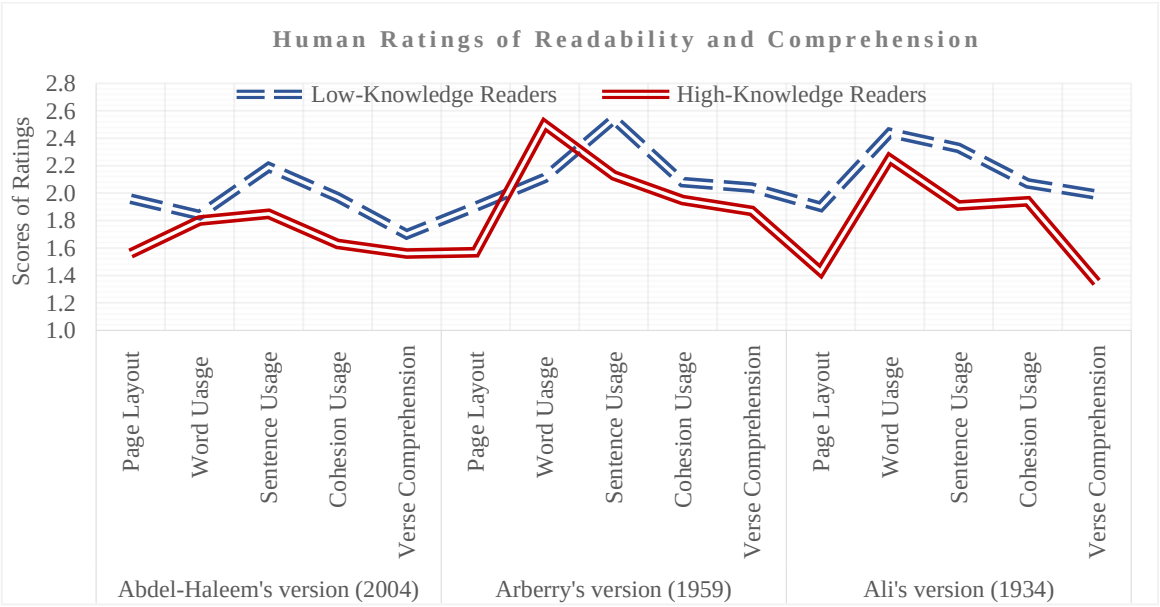


Figure 6.13: Judgements of Readability and Comprehension by Two Groups of Judges

To conclude, human judgements of translation readability and comprehension might be influenced by factors related to translation features and reader characteristics, such as prior knowledge. It is known that the readability and comprehensibility of a text do not depend on text features alone, but also on the characteristics of the reader (see Chapter 2). As shown in the results, those with previous knowledge found all three versions more comprehensible than did those with little knowledge. However, an increase in translation difficulty related to linguistic features (such as greater cohesion and more complex language) may not facilitate translation comprehension for readers, especially those with little prior knowledge.

6.4.6 Predictors of Readability and Comprehension

The findings of our regression models predicted the human judgements of comprehension, using Coh-Metrix indices and classic readability formulae. The computational indices of the lexical, syntactic, and cohesion features were better predictors of readability and comprehension than traditional readability formulae. A combination of these classic formulae explained 29% of the comprehensibility, whilst each of the linguistic indices related to lexical, syntactic, and cohesion features explained around 50%. The linguistic indices in the regression models describe reading comprehension difficulty at multiple levels of language and discourse. This model was designed in alignment with theories of text and discourse comprehension (Graesser and McNamara, 2011; McNamara et al., 2014).

The regression models in this study found that the translations were judged to be less comprehensible when the words were less common, less meaningful, and longer; when the words included fewer hypernyms; and when the sentences included more noun phrases, logical operators, cohesion (repetition aspects), and connectives. Specifically, cohesion and connectives were found to be the strongest predictor variables of readability and comprehensibility in the Qur'ān translations. High cohesion was not found to increase comprehensibility. This result supports the notion that an increase in cohesion does not make texts more readable or comprehensible (Crossley and McNamara, 2011c; Crossley et al., 2017).

The predictive power of the regression models supports the findings of previous studies that regressed automated measures with human judgements. For example, Pitler and Nenkova (2008) combined lexical, syntactic, and discourse features to predict human readers' judgements of text readability. They found that several linguistic features at the word, syntactic, and discourse levels were more significant predictors of readability than traditional readability formulae. In their experiment, 'discourse relations' were found to be the strongest predictors

of readability. This finding does not contradict the findings of our regression analyses of cohesion indices, which we found to be the strongest predictors of translation comprehensibility. In a similar study, Crossley et al. (2017) found that several lexical and discourse features were more predictive of human judgements of text comprehension than traditional readability formulae. They found no syntactic complexity indices that were predictive of reading comprehension.

6.4.7 Subjective Judgements and Objective Measures

This study used three approaches to text readability and comprehension: two types of objective and automated measures (the classic readability formulae and the computational tool Coh-Metrix) and subjective judgements of translation readability and comprehension. The three approaches are discussed here in terms of their accuracy, validity, efficiency, and limitations.

The classic approach involved estimating the reading difficulty of the three Qur'ān versions, based on two lower levels of text features: some aspects of word difficulty and sentence complexity. This estimation from one traditional formula into another was not consistent in giving similar grade levels for reading difficulty – although this variation was not out of the range of the predicted levels (see Chapter 5). All classic formulae used in this study predicted that Arberry's version would have the lowest comprehensibility and Abdel-Haleem's version the highest. This order did not differ for either the subjective judgements of verse comprehensibility or the objective Coh-Metrix measures of word difficulty and sentence complexity (see Chapter 5). However, the classic approach to readability analysis might be inefficient and inapplicable for estimating the reading difficulty of an English Qur'ān version. All five formulae overestimated the reading level of the three versions, which were graded at children's levels. Moreover, this approach did not holistically evaluate the language and all linguistic features affecting text readability and comprehension.

Text readability and comprehension is a multilevel process, thus it should be not limited to a single linguistic (or non-linguistic) predictor or variable. The computational tool Coh-Metrix enabled this study to distinguish between the linguistic features of the three Qur'ān versions, using several linguistic indices or measures of the lexicon, the syntax, the connectives, referential cohesion, and semantic coherence. This multilevel theoretical framework for analysing language or discourse was used in this study to describe and differentiate between the readability of Qur'ān translations in terms of six text factors: *style*, *cohesion, literalness, genre*, *register*, and *retranslation*. The subjective readability ratings were consistent with the

objective measures in terms of the difficulty of each writing style (see Chapter 5). Ali's and Arberry's versions were judged to contain more difficult words and complex sentences than Abdel-Haleem's. This overall human rating does not differ from the estimation of the automated style variables in Chapter 5.

In assessments of translation cohesion, Ali's and Arberry's versions were judged to be less lexically coherent than Abdel-Haleem's, although Ali's and Arberry's versions were measured by the Coh-Metrix cohesion indices as high in all aspects of language repetition, connectives, and semantic similarities. The human ratings of cohesion might appear contradictory, but they may support the findings of the automated measures of cohesion (presented in Chapter 5). High cohesion in the two former translations might be more affected by Qur'ān-specific cohesion patterns, literalness, and complex language. The readability of a Qur'ān translation could be influenced by other non-textual factors related to retranslation, such as the role of powerful institutions and publishing houses in subsidising the retranslations (discussed in the following section). The computational and classic approaches appeared to focus primarily on analysing text features, ignoring the organisation of the page layouts and the readers' background characteristics (such as prior knowledge), and their respective contributions to translation understanding.

The human judgement approach considered how readers with different levels of background knowledge engaged with Qur'ān versions of different levels of readability, making subjective ratings of translation readability and comprehensibility. This approach allowed for the derivation of more practical findings, when human ratings of verse comprehension were regressed with automated measures to predict Qur'ān translation comprehension difficulty. Several automated indices of lexicon, syntax, connectives, and particularly cohesion (repetition aspects) appeared to be the strongest predictors of readability and comprehension. This result increased the validity of the automated indices and research findings.

A human approach – especially using a rating scale – would have been convenient for obtaining judgements of text comprehension or an overall rating of translation difficulty (including word difficulty and sentence complexity). This procedure, however, might require the careful selection of judges for the assessments, especially when multi-item measures are used in a survey. This is because ratings of multiple linguistic features might be more subtle, time-consuming, and more challenging for an inexperienced judge – although the multi-item constructs allowed this study to cover many aspects of translation readability and comprehension. Such a construct might require an expert judge. The current computational

tools – such as the Coh-Metrix – might be an alternative to human judgement in text analysis, as they do the same job efficiently and reliably. Reader judgements produced more skewed, varying (i.e., between those with different levels of prior knowledge), and less generalisable results due to the relatively small sample of judges and other unexamined reader variables, such as age, religion, Arabic knowledge, and being a non-native speaker of English, whilst objective measures of readability and cohesion generated more consistent, reliable, and generalisable findings.

Despite these limitations, the three approaches and statistical techniques used in this study provide a new approach to gaining a full understanding of readability and textual changes in re/translations, especially when several multiple linguistic indices and statistical tests are analysed using a theoretical framework for textual analysis. The findings of automated measures might be more important on the theoretical level, whilst the results of subjective measures might be more useful on the practical level. The combination of approaches here has enabled this research to provide a richer understanding of the text and the factors affecting the readability of the Qur'ān translations. Other non-textual factors related to retranslation may also have influenced comprehensibility, as discussed in the following section.

6.4.8 Retranslation

As mentioned in the findings presented in Chapter 5, retranslation provides an opportunity to improve text readability and comprehensibility when the timespan between the translations is long. It allows for later works to achieve higher levels of text readability and to avoid the comprehensibility issues of earlier translations. The human judgements of translation comprehension support the findings of the automated measures, showing that all readers with either low or high knowledge judged Abdel-Haleem's version to be higher in verse comprehensibility than Arberry's version. This judgement was probably due to Abdel-Haleem's higher levels of readability.

However, Abdel-Haleem's comprehension result was slightly different when this version was compared to Ali's. High-knowledge readers judged Ali's version to be higher in verse comprehensibility than the other versions, but they also rated Ali's version as having a lower level of readability. A number of factors associated with retranslation – such as the policies of the publisher when selecting a retranslation, as well as the translator's own background characteristics – might explain why high-knowledge readers rated Ali's version higher for comprehensibility. Venuti (2004, p.30) writes as follows:

> A commercially oriented publisher may decide to issue retranslations of foreign canonical texts that have fallen into the public domain simply because their canonicity ensures a market demand … Hence, an ideology of commercialism will govern the selection of a foreign text for retranslation and dictate a discursive strategy that enhances the readability of the translation to ensure sales. A publisher driven by a profit motive may in fact wish to save the expense of commissioning a retranslation by reprinting a previous translation that has proven itself in the marketplace, even if in a revised version.

The judgements of high-knowledge readers were not only influenced by Qur'ānic *'text familiarity'* or knowledge of the text topic (Klare, 1978; 1988; Zakaluk and Samuels, 1988), but also by other non-textual factors, such as the role of powerful institutions and publishing houses in sponsoring and spreading certain translations to a wider readership. Ali's version is one of the most frequently reprinted, popular, and widely distributed Qur'ān translations in America, India, the United Kingdom, and the Middle East. Kidwai (2007) has traced more than 200 reprints of Ali's work. It is also subsidised by most well-known publishers, including King Fahd Complex for the Printing of the Holy Qur'ān in Saudi Arabia and Amana Publication in the United States. The role of powerful institutions and publishers in replacing retranslations and subsidising new ones is an important non-textual influence on readability – for translations of high and low levels of readability.

Another factor influencing translation readability and publisher selections is the translator's background characteristics, such as his level of knowledge about the culture, language, and history of a source text, as well as his religion, particularly when translating sacred texts. A translator's belief in and faithfulness to the meanings and interpretations of a source text might contribute to the success and spread of a retranslation (see Chapter 2). Publishing houses compete to sponsor translations by Muslims to meet the needs of Muslim readers. For example, Arberry's version has undergone approximately 30 reprints, but the reissuing frequency of his translation has been in decline in the last decade (Kidwai, 2007). Qur'ān translations by Muslims usually have higher numbers of reprints than versions by non-Muslims. Pickthall's version (1934), for example, has seen more than 80 reprints, whilst Bell's (1937) has been reprinted just seven times (for extensive reviews, see Kidwai, 2007; 2018) – even though these translations were published in the same decade.

To conclude, textual factors, readers' prior knowledge, and retranslations may not be the only influences on readability. Non-textual factors such as the importance of institutions and

publishing houses and the translator's general background might also contribute to the comprehensibility of a Qur'ān translation.

6.5 Chapter Summary

This chapter presented the findings of human judgements. It provided subjective ratings of several textual factors of translation readability (including vocabulary, syntax, cohesion, and page-layout organisation) and translation comprehensibility (including the comprehensibility of the chapter verses and titles and the use/non-use of comprehension aids such as footnotes and introductions). The human judgements allowed for comparisons between the findings of the subjective and objective measures.

The reader judgements of readability and comprehensibility were compared for two groups of judges: low-knowledge readers and high-knowledge readers. Prior knowledge of the Qur'ānic texts interacted with the readability and comprehension variables, suggesting that readers – especially those with low knowledge – understood more from the more readable texts (those with less complex styles of word usage and syntax and lower rates of lexical repetition, semantic coherence, connectives, and literalness). In contrast, high-knowledge readers understood more from the less readable Qur'ān versions with more complex linguistic features. The judgements of the high-knowledge readers were not only influenced by their familiarity with the Qur'ānic text, but also by other non-textual factors related to retranslation – such as the role of powerful institutions and publishing houses and the translator's general background. The less comprehensible versions (such as Arberry's) were found to use no comprehension aids. It was also found that style variables – such as word difficulty and syntactic complexity – showed greater weighting in predictions of translation readability than the variables of cohesion and page layout. There was a strong relationship between judges' ratings of the difficulty of verses and the human ratings of verse comprehension for each version, suggesting that as the level of translation readability increases, the degree of translation comprehension also increases, and vice versa.

This multimethod study also compared the findings of reader judgements with the results of two objective measures of readability, cohesion, and language: the classic and computational approaches. The classic approach accurately predicted the order of comprehension difficulty for the three versions, but its predictions of reading level tended to be at the child level for all versions. This finding suggests the need to develop readability formulae that provide more accurate estimates of language difficulty in texts other than (school) textbooks, such as Qur'ānic and literary translations that usually have variations in language and difficulty. The findings from our regression models indicate that a number of individual indices from the Coh-Metrix computational tool related to lexical, syntactic, connective, and specifically cohesion (repetition) measures were more strongly predictive of comprehension difficulty than the

classic readability formulae, which only explained approximately 29% of the verse comprehensibility.

Using a principal components analysis and statistical pairwise comparisons of the translations, the Coh-Metrix indices allowed the description and differentiation of the readability of the translations, according to six major text factors: *style*, *cohesion*, *literalness*, *genre*, *register*, and *retranslation*. The human judgements of readability aligned with the automated measures of the writing style of the two earlier versions, which were judged to use more complex styles of vocabulary and syntax than Abdel-Haleem's text, the most recent version. The human judgements and automated measures suggest that high cohesion in Qur'ān translations is not an indicator of greater readability, but rather of complex language and literalness.

The computational approaches (e.g., the Coh-Metrix approach) appeared to be more efficient, faster, cheaper, and more reliable, providing a richer measure of readability, cohesion, language, and multiple sources of comprehension difficulty, than other approaches to text analysis –such as human judgement and the classic formulae. This does not mean that the human approach was not an important procedure, but rather it focused primarily on providing more insights into differences between readers' views and behaviours and the interactions between prior knowledge of a subject and readability and comprehension measures. The human judgements might be more useful on the practical level, whilst the findings of automated measures might be more important on the theoretical level. All three approaches considered either text or reader factors that make translations more readable and comprehensible, but they ignored non-textual factors related to retranslation – such as the role of institutions and publishing houses and the translator's general background, as well as their influence on the comprehensibility of a translation.

To conclude, the reader judgements produced more skewed and varying and less generalisable findings than the objective measures of readability and cohesion and language (such as the classic readability and Coh-Metrix indices). The number of judges in each survey was not large enough (20 participants) to highlight robust statistical differences or large comprehension and readability differences. Other reader variables – such as age, religion, knowledge of Arabic, and being a non-native speaker of English, and the effects of these on text comprehension – were not considered in this study. The limitations and other challenges of this study are discussed further in the following conclusion chapter.

Chapter 7: Conclusion

7.1 Introduction

This chapter presents the conclusions of this multimethod study, which has explored the primary text factors and reader factors that contribute to the readability and comprehension of a Qur'ān translation. The chapter first provides a summary of the main study findings emerging from three approaches to assessing readability and comprehension – two objective and automated measures (the classic readability formulae [CRFs] and the computational tool Coh-Metrix) and the subjective-quantitative judgements of translation readability and comprehension. It also discusses the contributions and limitations of the current study. Finally, suggestions for future studies are presented.

7.2 Summary of Research Findings

This section reviews the findings of the three approaches used in the previous two chapters. This multimethod study has described and differentiated the readability and comprehensibility of various Qur'ān translations according to the following text and reader factors: *style*, *cohesion*, *literalness*, *genre*, *register*, *page layout*, *comprehension aids*, *reader's prior knowledge*, and *retranslation.* Before turning to these possible contributors to Qur'ān translation readability and comprehension, however, the research utilised a classic and simple approach as a preliminary analysis to objectively estimate the reading difficulty and reading grade level of the Qur'ān versions, based on five well-known CRFs. The five CRFs accurately estimated the relative levels of comprehensibility of the three versions, and these predictions matched the human ratings of the three versions' readability.

However, the classic approach to readability measurement may encounter two issues of validity. First, all five CRFs used in this thesis estimated the reading grade level of the three versions to be appropriate for elementary-age children. This estimated grade level explained about 20% of the human judgements of rating-verse comprehension. This traditional approach was found to be less predictive of verse comprehension when regressed with a number of individual Coh-Metrix indices related to language and discourse. Although the classic approach is a reliable tool for replication, the study findings provide evidence that CRFs explained little of the text comprehension and performed poorly in estimating reading grade levels of Qur'ān translations. This result provides support for similar findings from a recent study on the failure of CRFs to predict text comprehension by adult readers (see Crossley et al., 2017); thus, we suggest that a new readability formula is needed to more accurately estimate the appropriate reading grade levels of written materials other than school textbooks.

Second, another issue of validity in the classic approach is that it does not assess all text and reader factors when measuring text comprehension. Thus, this multimethod study analysed the readability of Qur'ān versions on a multiple number of computational linguistic indices, including lexicon, syntax, connectives, referential cohesion, sematic coherence, and narrativity. This multilevel theoretical framework of text characteristics was able to provide more reliable and valid findings regarding the readability of the Qur'ān translations. The human judgements of translation readability and comprehension used in the study may not only reveal differences between approaches in assessing readability, but also help us to converge the findings of each approach. Put differently, this use of multiple approaches to the same issue enabled this research to validate the findings of subjective and objective measures.

The primary text factors and some of the reader characteristics that might contribute to the readability and comprehensibility of Qur'ān translations are summarised below.

Style. The style of the vocabulary and syntax used in written materials is one of the most significant and basic contributors to readability. This is because it can make the process of reading a Qur'ānic thought simpler or more difficult. As different translators have their own styles of writing, this study used six automated indices of lexical difficulty and six elements of syntactic complexity to determine style difficulty. In the study of the six measures of lexical-psychological and semantic features, it was found that a lower incidence of shorter words, word frequency, hypernyms, concreteness, imageability, and meaningfulness in the writing style of a translation might not facilitate text comprehension for readers, especially low-knowledge readers. A higher incidence of these lexical characteristics might make Qur'ānic meanings clearer and simpler, creating more vivid language, which could keep the reader engaged. Arberry's and Ali's versions are low in these features and not considerably different from each other in their style variables. Abdel-Haleem's version, in contrast, is very high in these lexical features.

In addition, other syntactic measures can explain the difficulty of writing style. According to the findings of the Coh-Metrix syntactic indices, Ali's and Arberry's versions use more complex syntactic structures, which appears to have an effect on ease of reading. The findings of the automated lexical and syntactic measures align with the subjective judgements of translation readability and comprehension. Overall, the vocabulary and syntax of Ali's and Arberry's versions tend to be more complex across short, medium, and long chapters than those of Abdel-Haleem's translation, which adopts a more contemporary and simpler style of language across all chapters.

Cohesion. The readability of a Qur'ān translation is not only assessed through a single text factor or by the limited variables of style used in CRFs, but also through the connections between ideas and sentences in the text. This study discriminated between low-cohesion and high-cohesion Qur'ān versions, based on 16 cohesion indices of connectives, referential cohesion (lexical repetition), and sematic coherence (semantic similarities between words). Ali's and Arberry's versions have a greater degree of lexical repetition, semantic similarities, and all types of connectives than Abdel-Haleem's translation. Ali and Arberry are not significantly different in most measures of cohesion. The findings for the automated cohesion measures and subjective judgements of text readability provide evidence that high cohesion in Qur'ān translations is not an indicator of high readability and comprehensibility, but rather of complex language and literalness.

Literalness. A significant influence on the readability of a Qur'ān translation is the degree of literalness. This is because literalness can contribute to making a TT more or less comprehensible to the reader. This study focused on the act of transferring the SL stylistic and semantic norms to the TL. Ali's and Arberry's versions were found to have similar patterns in their respective amounts of lexical repetition, semantic similarities, and all types of connectives, especially the connector 'and' (additive connective). This resemblance in these cohesion measures indicates that Ali and Arberry have preserved the semantic and cohesive patterns of the SL in their translated versions. In contrast, Abdel-Haleem's version appears to avoid closeness to Qur'ān-specific language and cohesive ties, making his translation more readable and comprehensible.

Genre. Another textual factor found to have an effect on ease of reading comprehension is text genre or type of text – narrative, expository, persuasive, or descriptive (McNamara et al., 2014). A text types such as the Qur'ān might be a combination of discourse types. According to previous research findings, narrative texts are easier to understand and recall than other types (Haberlandt and Graesser, 1985; Graesser and McNamara, 2011). Although the Qur'ān is its own text type and combines legal, narrative, and instructional passages (Abdul-Raof, 2019), this study measured the degree of narrativity in Qur'ān translations. The findings of this research shows that Abdel-Haleem's version contains more features of narrativity than other versions. This type of discourse is not difficult to understand, but some linguistic features of text register may be chosen by a translator that reduce the narrativity. Narrative style appears to be a salient generic characteristic in English Qur'ān translations, as discussed in the next text factor.

Register. The register of a text is usually determined by its linguistic patterns. This research used a multi-dimensional approach to textual variation to determine the characteristics of the register of the translations and the sources of difficulty within the register. The findings show that the registers of the translations are characterised by features of Qur'ān-specific cohesive ties, persuasive and imaginal style, integrative structure, elaborative structure, and narrative style. Ali's and Arberry's versions are characterised by higher occurrences of Qur'ān-specific cohesive ties, integrative structure, and elaborative structure than Abdel-Haleem's. The linguistic features of these dimensions are associated with text difficulty and discourse complexity (see Chapter 5). Abdel-Haleem's version uses fewer of the linguistic characteristics of these three dimensions and is characterised by a higher presence of narrative style and persuasive and imaginal style. Abdel-Haleem's version employs more linguistic features of persuasive and imaginal style, with more vivid language, imagery, and concrete details that are easy to understand. Abdel-Haleem's version uses fewer Qur'ān-specific cohesive ties and integrates less information than either Ali's or Arberry's texts. Ali's and Arberry's texts tend to employ more complex syntactic structures than Abdel-Haleem's version, which uses simpler syntactic structures. The final register to discuss has characteristics of narrative style, with frequent occurrences of everyday language, temporal, and adverbial features. The three versions do not differ in their use of this style, except in their use of everyday language. The two earlier versions employ a narrative style with a higher lexical register (i.e., archaisms), whereas Abdel-Haleem employs a narrativity with a lower lexical register. These language variations are not difficult to process, but some linguistic choices made by translators are known to be associated with text difficulty and discourse complexity (see Chapter 5).

Page layout. Page layout is another crucial factor in readability. This is because the page layout guides the reader through the physical elements and structure of the text, making the content clearer and easier to process. The page layout of a translated text usually follows that of the ST, but this can be unsuitable for the genre norms of the TL and a different design of the page elements may be required. This study asked participates to judge different Qur'ānic page layouts on six elements: *chapter title*, *chapter length*, *verse number*, *paragraph division*, and *overall clarity* and *organisation of page layout*. The human judgements differed significantly only in terms of the first three elements of page layout. According to the subjective judgements, the use of a list of verse numbers is likely to be clearer to readers than other formats such as numbering verses with a small superscript or the style of poetry line numbers. Second, the location of a chapter title tends to be more obvious when it is centred. Another finding is that

a Qur'ānic chapter is likely to be easier to read when it is not long. While measures of other elements of the page layout – such as paragraph division – did not differ substantially from the human ratings, paragraph division plays a role in readability. This is because paragraph division shows the reader how verses and thoughts are connected. A version with prose paragraphs is likely to flow more smoothly and logically for the reader than a version with stanza paragraphs. A separate empirical study is required to test whether the division of paragraphs in Qur'ān translations creates difficulty for reading comprehension.

Comprehension aids. Comprehension aids such as introductory information and footnotes are vital page layout elements, guiding the reader through the structure of the content and potentially enhancing the comprehensibility of the translation (see Chapter 3). Comprehension aids can give the readers a summary of the main themes of each chapter and clarify linguistic and cultural elements in the footnotes. According to the subjective judgements, the inclusion of introductory information and footnotes in Qur'ānic chapters is useful for text comprehension. The translations with low comprehensibility lacked these comprehension aids. This low comprehensibility was particularly apparent in Arberry's version. At the same time, when a chapter is full of long, extended commentaries or footnotes, this might risk the clarity and comprehensibility of the Qur'ānic text, although this needs further empirical testing.

Reader's prior knowledge. Thus far, the summary findings have focused on textual factors contributing to the readability of the Qur'ānic texts. However, the characteristics of the reader – such as their prior knowledge – also influence the comprehensibility of a translation. Since a translation of the Qur'ān is usually intended for the public in general, its readability and language use should be appropriate for most adult readers, including those who are less educated and who have little background knowledge about the Qur'ān. These study findings show that high-knowledge readers rated Qur'ān translations as lower in text difficulty and higher in text comprehensibility than did low-knowledge readers. When the judgments of these two groups of readers were compared between Abdel-Haleem's and Arberry's versions, they understood more of the more readable Qur'ān version (i.e., Abdel-Haleem's), which they found to use a simpler style of words and syntax and have lower rates of lexical repetition, semantic coherence, connectives, and literalness; higher rates of narrativity; and a lower degree of the linguistic features of register variation associated with text difficulty and discourse complexity. When the judgments of these two groups of readers were compared between Abdel-Haleem's and Ali's versions, high-knowledge readers understood more of the less readable Qur'ān version (i.e., Ali's), which uses more complex linguistic features. In summary, the judgements

of the high-knowledge readers were not only influenced by their familiarity with the Qur'ānic text, but also by other non-textual factors related to retranslation.

Retranslation. Language and socio-cultural norms (see Chapter 2) in retranslation are not universal: they change, as time and generations pass. The two earlier translations of the Qur'ān represent the stylistic flavour and translational norms of an age in which it was common to use more biblical style, complex language, and more ST-oriented features. With little time having elapsed between the publication of the two works, the texts reveal no substantial linguistic changes. They might be influenced by other non-linguistic motives (see Chapter 2). When the time span between translations is longer, retranslation allows the later works to avoid readability and comprehensibility issues. For example, the study findings show that Abdel-Haleem's most recent version avoids the style and language of the earlier period of Qur'ān translations, thereby achieving higher levels of readability and comprehensibility. However, the role of powerful institutions and publishers in replacing retranslations and subsidising new ones is an important non-textual influence on readability – for translations of high and low levels of readability.

To conclude, the findings of this multimethod research present a multilevel theoretical framework for analysing and assessing the readability of Qur'ān translations based on the following primary text factors and reader characteristics: *style*, *cohesion*, *literalness*, *genre*, *register*, *page layout*, *verse comprehensibility*, *comprehension aids*, *reader's prior knowledge*, and *retranslation*. This theoretical framework for objective and subjective measurements of textual and reader factors has theoretical and practical value for estimating readability in several fields of study. The contributions of this study are presented in the following section.

7.3 Research Contributions

The findings of this multimethod research have theoretical, practical, and methodological implications for three fields of study: readability research, Qur'ānic studies, and translation studies. At the theoretical level, this study contributes to readability studies because it presents a multilevel theoretical framework for objectively estimating text readability from different angles. It introduces a new approach by showing how a number of individual computational indices related to lexical characteristics, syntactic structure of sentences, measures of referential cohesion, semantic coherence, connectives, and narrativity are utilised to describe and differentiate readability according to key text factors, including style of vocabulary and syntax, cohesion, genre, and register. This combination of automated measures to objectively

estimate readability and analyse register in readability studies is a novel approach that could be used in future studies to gauge the language and characteristics of a text or translation at various levels of difficulty. The current study provides a theoretical framework based on automated measures for assessing readability. Future research could test this framework with other texts and develop it further to include all reader characteristics (see the following section).

The methodology of this multimethod study reveals the importance of computational tools for translation studies and especially translation quality assessment. More specifically, computational linguistic indices such as the Coh-Metrix provide a more objective, faster, more reliable, and less complex approach to detecting textual differences than the traditional linguistic-textual approaches commonly used in translation quality assessment (see House, 1997) to determine the quality of a translation. Coh-Metrix indices are built on theories of language and discourse comprehension (McNamara et al., 2014) that enable researchers to lexically, syntactically, and textually distinguish between characteristics of a translation.

In addition, the methodology of this quantitative research was useful for not only estimating translation at various levels of difficulty, but also for converging the findings of subjective and objective measures. The multiple approaches allowed for the investigation of readability from different perspectives and ensured more valid and accurate findings, revealing differences between the three approaches. While the classic approach did not appear to accurately estimate the reading grade levels of the Qur'ān translations, the findings of the classic approach were generally consistent with those of other approaches when comparing the versions' respective levels of readability and comprehensibility. The computational linguistic indices approach was found to be less problematic, more reliable, and more valid as a measure of readability and more valuable for theoretical knowledge than other approaches to readability analysis. The human judgements of text readability and comprehension were useful for obtaining more practical findings and for identifying the overall reactions of the readers concerning text difficulty and organisation of page layout, as well as showing how the readability of the different versions was perceived and understood by readers with different levels of background knowledge.

In relation to quantitative analysis, this study sheds light on the value of simple and advanced statistical techniques in readability research and translation studies for analysing language patterns and textual-linguistic changes between retranslations. The findings indicate that longer timespans between retranslations can reveal robust statistical differences and broad textual-linguistic changes. However, other advanced statistical analyses – such as regression

analysis – are helpful for deriving practical findings from automated measures and human ratings of verse comprehension. Regression analysis revealed the most robust predictors of readability and comprehension in the Qur'ān translations. Finally, a multi-dimensional approach to textual variation was useful for determining not only the register of the translations, but also the sources of difficulty in the register. The statistical approach utilised in this thesis to identify register or linguistic variations is new in *English* Qur'ān translation studies.

The combination of statistical techniques with several predictors of readability and comprehensibility taken from different quantitative approaches was beneficial for presenting evidence of validity and reliability. The multiple approaches in this quantitative research were shown to be valuable not only for exploring the differences between subjective and objective measures, but also for providing more objective, reliable, valid, and generalisable findings.

The present study has practical value for Qur'ānic studies, with its exploration of the text and reader factors that contribute to comprehensibility. This research has shown that human judgements of translation readability and comprehensibility are influenced by text and reader factors. The significance of reducing text difficulties has been demonstrated in assessing the three versions of the Qur'ān. The results show that the Qur'ān versions with more complex styles of words and syntax; higher rates of lexical repetition, semantic coherence, connectives, and literalness; lower rates of narrativity; and a higher degree of linguistic features of register variation associated with text difficulty may not necessarily facilitate text comprehension for readers, especially low-knowledge readers. This result provides support for the findings and claims of several readability and comprehension researchers, such as Klare (1976:1988), Chall (1996), Kintsch and Miller (1981), O'Reilly and McNamara (2007), and McNamara et al. (2011). This practical value of language use for some readers can contribute to Qur'ānic translators by considering text and reader factors when the Qur'ān is retranslated. Another practical value of this research is that it sheds light on the importance of text readability in conjunction with other non-textual factors, such as the background characteristics of the translator and the role of powerful institutions and publishers, both of which could contribute to the comprehensibility of a Qur'ān translation.

7.4 Research Limitations

While this multimethod study provides some important insights into the text and reader factors that contribute to the readability of a Qur'ān translation, it does have limitations that need to

be acknowledged. The limitations concern five parts of this thesis: text factors, reader characteristics, sample size, study methodology, and the number of linguistic indices.

First, the theoretical framework used to assess readability focuses primarily on the key text factors that make Qur'ān translations easier (or more difficult) to read and understand. This research is limited to measures of readability and comprehension, investigating how readers with low or high levels of prior knowledge interact with Qur'ān versions with low or high levels of readability, without considering the relationship between the original text and the translation. The results of the study thus do not provide any evidence of whether a Qur'ān translation is semantically (i.e., in terms of the accuracy of the message transfer), pragmatically, or culturally equivalent to the original text. These translation factors could affect the comprehensibility of a text.

Another limitation is that, although the study included adult readers from various age groups (18–60 years old), men and women, and Muslims and non-Muslims, the study only considered the readers' background knowledge in relation to their subjective judgements of readability and comprehension. The study did not examine the role of other reader characteristics, such as gender, beliefs, interest, or reading ability in textual understanding. These characteristics might interact with readability measures and influence what readers bring to the text (see Klare, 1976, 1988; Entin and Klare, 1985; McNamara and Kintsch 1996; Bell et al., 2012; Kendeou et al., 2011; Logan and Johnston, 2009; Hannon, 2014). Finally, the study presents no empirical evidence about the readability or comprehensibility of the Qur'ān translations for other age categories (such as children and teenagers).

Another reader-related limitation is the sample size used to obtain human judgements. There were 20 participants in each survey. This seemed to be sufficient for the purpose of this study (see Chapter 4), but it is not adequately representative for the purposes of generalisation. It is too small to reveal more robust and greater statistical differences between human judgements of translation readability and comprehension.

The methodology of the study presents a fourth limitation, although the merit of this multimethod study is that it involved direct comparisons between the findings of subjective and objective measures. This study used three quantitative approaches to assess readability and comprehension: two objective procedures (the CRF and the computational tool Coh-Metrix) and subjective-quantitative judgements. The study did not include any qualitative methods, such as interviews with readers or Qur'ān translators, which would provide verbal outputs regarding textual difficulties in Qur'ān versions. In addition, other advanced qualitative

methods – such as eye tracking in psychology – could be useful for observing human behaviours and cognitive processes during reading Qur'ānic extracts (for more on eye tracking studies, see Duchowski, 2007; O'Brien, 2009; Hvelplund, 2017).

Finally, the number of linguistic indices used in this study to analyse the language variation in Qur'ān translations is another limitation. The number of Coh-Metrix indices was limited to 29. The linguistic features used in the register analysis captured the main language varieties but did not cover all linguistic features or variations in Qur'ān translations. These are the limitations of this study that future studies should address.

7.5 Suggestions for Future Research

The previous limitations could be useful suggestions for future studies. For example, future research could include larger samples of adult readers to consider the interactions between readability and comprehension measures and reader characteristics, including prior knowledge, gender, beliefs, interest, and reading ability. Another important suggestion for future research is to investigate the readability of Qur'ān translations amongst other age groups, such as children and teenagers. This could provide insights into the language and text factors of a Qur'ān version that make it appropriate for readers in different age groups. As Crossly et al. (2017, p.340) observe, their study 'has identified a number of linguistic features that influence the reading comprehension of young readers; yet, less is known about whether and how these findings extend to adult readers'.

The register analysis used in this study to computationally analyse linguistic variation opens the door to new avenues in two fields. First, future research into Qur'ān translations could involve a corpus-based study to test our multi-dimensional approach to register analysis in more translations, developing the approach by adding more linguistic features to uncover more linguistic patterns and variation in the language of Qur'ān translations. This study used 29 linguistic features – at the word, sentence, and discourse levels – and future Qur'ānic researchers could include a wider range of linguistic characteristics, such as tenses, modals, reduced forms, passive forms, negations, pronouns, and so on (for more, see Biber, 1988, 1992; Louwerse et al., 2004). In addition, future research in translation studies could use computational linguistic tools and quantitative techniques to explore and identify register characteristics in other translation types (e.g., literary, legal, scientific, and media). Multi-dimensional approaches have the power to not only identify register across translations, but also recognise sources of difficulty within the language of the same register.

A further suggestion for future studies is to utilise the theoretical framework of readability analysis used in this thesis to analyse more recent Qur'ān translations, combining other research paradigms and methods such as eye tracking and interview. Eye tracking would allow a researcher to measure how participants' comprehension is affected by the cognitive efforts associated with different aspects of translated Qur'ānic vocabulary, syntax, cohesion, and page layout. Eye tracking is an unexplored instrument in Qur'ān translation studies, and it could provide valuable insights into real-life scenarios of comprehension difficulty. In addition, the combination of eye tracking and computational indices with qualitative interviews, for example, would enable in-depth observations of people's experiences and opinions. Open-ended interviews could be conducted with Qur'ān readers and translators, and their responses could be compared. This mixed methodology could produce rich results concerning questions of cognitive load and the translation elements that facilitate or hinder readers' comprehension.

Finally, this multimethod research has provided useful evidence about the key textual factors and reader characteristics that make some Qur'ān translations more readable and comprehensible than others. The findings of this research can be considered practical and suggestive, and a starting-point for future research and for Qur'ān translators rethinking about Qur'ān readers' differing levels of background knowledge. Less difficult English Qur'ān versions are highly important for a large percentage of the population. Qur'ān versions with more complex styles in terms of words and syntax; higher rates of lexical repetition, semantic coherence, connectives, and literalness; lower rates of narrativity; and more linguistic features of register variation associated with text difficulty and discourse complexity do not facilitate text comprehension for many readers, especially those with little prior knowledge of the content. Abdel-Haleem's version, which was found here to be the most readable and comprehensible of the translations for this less-knowledgeable reader, should be tested in future research and compared with other recent, prominent Qur'ān translations.

Bibliography

Abdel-Haleem, M. 2011. *Understanding the Qur'an: Themes and Style*. 3rd ed. London: I. B. Tauris.

Abdel-Haleem, M. 2004. *The Qur'ān: A New Translation*. London: Oxford University Press.

Abdul-Raof, H. 2001. *Qur'an Translation: Discourse, Texture and Exegesis*. London and New York: Routledge.

Abdul Raof, H. 2005. Cultural aspects in Qur'ān translation. In L. Long ed. *Translation and religion: Holy untranslatable?* Clevedon, Buffalo, Toronto: Multilingual Matters, LTD. pp. 162-172.

Abdul-Raof, H. 2004. The Qur'an: Limits of translatability. In: S. Faiq ed. *Cultural encounters in translation from Arabic*. Clevedon: Multilingual Matters, pp. 91–106.

Abdul-Raof, H. 2007. On the Stylistic Variation in the Qur'ānic Genre, *Journal of Semitic Studies*, **52** (1), pp. 79-111.

Abdul-Raof, H. 2018. *New Horizons in Qur'anic Linguistics: A Syntactic, Semantic and Stylistic Analysis*. London and New York: Routledge.

Abdul-Raof, H. 2019. *Text Linguistics of Qur'anic Discourse: An Analysis*. London and New York: Routledge.

Adelberg, A.H. 1983. The accounting syntactic complexity formula: a new instrument for predicting the readability of selected accounting communications. *Accounting and Business Research*, **13** (51), pp.163-175.

Al-Tamimi, A.K., Jaradat, M., Al-Jarrah, N., and Ghanem, S. 2014. AARI: Automatic Arabic readability index. *Int. Arab J. Inf. Technol*. **11**(4), 370-378.

Al-Ajlan A., Al-Khalifa H., and Al-Salman A. 2008. Towards the Development of an Automatic Readability Measurements for Arabic Language. *In Proceedings of the 3rd International Conference on Digital Information Management, 13/16 November 2008*, London, United Kingdom, pp. 506-511.

Al-Amri, W. B. 2008. Qur'ān translatability at the phonic level. *Perspectives: Studies in Translatology*, **15**(3), pp. 159-176.

Al-Amri, W. B. 2010. Qur'ān Translation and Commentary: An Uncharted Relationship? *Islam and Science*. **8**(2), pp.81-110.

Al-Attar, M. 2010. *Islamic Ethics: Divine Command Theory in Arabo-Islamic Thought*. New York: Routledge.

Al-Asqlānī, I. 1959. *fatḥ al-bārī*. Vol 3. Beirut: Dār Al-maʿrifa.

Aldahesh, A.Y. 2014. (Un)translatability of the Qur'ān: A theoretical perspective. *International Journal of Linguistics*. **6**(6), pp. 23.

Al-Heeti, K. N. 1984. *Judgment analysis technique applied to readability prediction of Arabic reading material*. Ph.D. thesis, University of Northern Colorado.

Al-Halālī, M. T. 1979. *al-sarrāj al-munīr fī tanbīh jamāʿa al-tablīġ ʿalā ʾaḳṭāʾihim*. Morocco: maṭabʿa al-najāḥ al-jadīda.

Ali, A.Y., 1934/1937. *The Holy Qur'ān: Arabic Text with an English Translation and Commentary*. Lahore: Sh. Muhammed Ashraf.

Ali, Y.A. 1989. *The Holy Qur'ān: English Translation of the Meanings and Commentary*. Madinah: King Fahd Complex for the Printing of the Holy Qur'ān.

Al- Zarkashi, B. M. 1957. *al-Burhān fi 'ulūm al-Qur'ān*. Vol 1. Cairo: Dār Al-turāṯ.

Anderson, R. C., and Davison, A. 1988. Conceptual and empirical bases of readability formulas. In A. Davison and G. M. Green eds. *Linguistic complexity and text comprehension: Readability issues reconsidered*, Lawrence Erlbaum Associates: Inc, pp. 23-53.

Anguera, M.T., Blanco-Villaseñor, A., Losada, J.L., Sánchez-Algarra, P. and Onwuegbuzie, A.J. 2018. Revisiting the difference between mixed methods and multimethods: Is it all in the name? *Quality & Quantity*, **52**(6), pp.2757-2770.

Arberry, A.J. 1983. *The Koran Interpreted*. Oxford: Oxford University Press.

Arberry, A. J. 1996. *The Koran Interpreted*. New York: Touchstone.

Armbruster, B. B. 1984. The problem of inconsiderate text. In: Duffey, G. ed. *Comprehension instruction*. New York: Longman, pp. 202-217.

Bailin, A. and Grafstein, A. 2016. *Readability: Text and context*. London: Palgrave Macmillan.

Beaman, K. 1984. Coordination and subordination revisited: Syntactic complexity in spoken and written narrative discourse. In Tannen, D. ed. *Coherence in spoken and written discourse*. Norwood, NJ: Ablex.

Bell, R. 1937. *The Qur'ān, Translated, with a Critical Re-arrangement of the Surahs*. Edinburgh: T. and T. Clark.

Bensimon, P. 1990. Présentation. *Palimpsestes*. **4**, pp. ix–xiii.

Berman, A. 1990. La retraduction comme espace de la traduction. *Palimpsestes*. **4**, pp. 1-7.

Biber, D. 1988. *Variation across Speech and Writing*. Cambridge University Press: Cambridge.

Biber, D. 1992. On the complexity of discourse complexity: A multidimensional analysis. *Discourse Processes*. **15**(2), pp.133-163.

Biber, D. 1990. Methodological issues regarding corpus-based analyses of linguistic variation. *Literary and linguistic computing*. **5**(4), pp.257-269.

Binkley, M.R. 1988. New ways of assessing text difficulty. In: Zakaluk, B.L. and Samuels, S.J. *Readability: Its past, present, and future*. Newark, DE: International Reading Association, pp.98-120.

Boullata, I.J. 1988. The rhetorical interpretation of the Qur'ān: i'jaz and related topics. In: Rippin, A. ed. *Approaches to the History of the Interpretation of the Qur'ān*. Oxford: Clarendon Press, pp.139-157.

Brewer, J., and Hunter, A. 2006. *Foundations of Multimethod Research: Synthesizing Styles*. 2nd ed. London: Sage Publications, Inc.

Broadwell, G., Boz, U., Cases, I., Strzalkowski, T., Feldman, L., Taylor, S.M., Shaikh, S., Liu, T., Cho, K., and Webb, N. 2013. Using Imageability and Topic Chaining to Locate Metaphors in Linguistic Corpora. In: Greenberg A.M., Kennedy W.G., Bos N.D. eds. *International Conference on Social Computing, Behavioral-Cultural Modelling, and Prediction*, Aprill 2013, Springer, Berlin [Online]. Springer: Heidelberg, pp.102-110. [Accessed 1 May 2019]. Available from: https://doi.org/10.1007/978-3-642-.

Bruce, B., Rubin, A. and Starr, K. 1981. Why readability formulas fail. *IEEE Transactions on Professional Communication*, **24** (1), pp.50-52.

Bryman, A. 2012. *Social research methods*. 4th ed. Oxford: Oxford University Press.

Bu, H., Connor-Linton, J. and Wang, L. 2020. *Linguistic variation in the discourse of corporate annual reports: A multi-dimensional analysis*. *Discourse Studies*, **22**(6), pp. 647-677.

Cambridge Advanced Learner's Dictionary. [Online]. 4th ed. 2013. [Accessed 7 October 2020]. Available from: www.dictionary.cambridge.org.

Chafe, W.1982. Integration and involvement in speaking, writing, and oral literature. In Tannen D. ed. *Spoken and written language: Exploring orality and literacy*. Norwood, NJ: Ablex, pp3-40.

Chall, J. 1988. The Beginning Years. In: Zakaluk, B.L. and Samuels, S.J. *Readability: Its past, present, and future*. Newark, DE: International Reading Association, pp.2-13.

Chall, J. 1947. This Business of Readability. *Educational Research Bulletin*, **XXVI** (1) pp.1-13.

Chall, J. 1996. Varying approaches to readability measurement. *Revue québécoise de linguistique*, **25**(1), pp.23-40.

Charrow, V. 1988. Readability vs. comprehensibility: A case study in improving a real document. In: Davison, A. and Green G. eds. *Linguistic Complexity and Text Comprehension*. Hillsdale, NJ: Lawrence Erlbaum, pp.85-114.

Chebat, J.C., Gelinas-Chebat, C., Hombourger, S. and Woodside, A.G. 2003. Testing consumers' motivation and linguistic ability as moderators of advertising readability. *Psychology & Marketing*, **20**(7), pp.599-624.

Chesterman, A. 2004. Hypotheses about Translation Universals. In: Hansen, G. Malmkjær, K., and Gile, D. eds. *Claims, Changes and Challenges in Translation Studies*. Amsterdam: John Benjamins, pp.1-14.

Chesterman, A. 2010. Why study translation universals? *Acta Translatologica Helsingiensia*, 1, pp. 38-48.

Chesterman, A. 2000. A Causal Model for Translation Studies. In: Olohan, M. ed. *Intercultural Faultlines. Research Models in Translation Studies: Textual and Cognitive Aspects*, Manchester: St Jerome, pp. 15–27.

Choi, J. and Kim, K.H. 2021. Translating Islamic Knowledge for the New Soil: Two Korean Translations of the Qur'ān in South Korea. *Interventions,* **23**(7), pp.973-996.

Coleman, M. and Liau, T.L. 1975. A computer readability formula designed for machine scoring. *Journal of Applied Psychology*, **60**(2), p.283.

Collins-Thompson, K. 2014. Computational assessment of text readability: A survey of current and future research. *ITL-International Journal of Applied Linguistics*, *165*(2), pp.97-135.

Connell, L. and Lynott, D. 2012. Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition*, **125**(3), pp.452-465.

Creswell, J.W. 2014. *A concise introduction to mixed methods research*. London: SAGE Publications.

Crossley, S.A., Allen, D. and McNamara, D.S. 2012. Text simplification and comprehensible input: A case for an intuitive approach. *Language Teaching Research*, ***16***(1), pp.89-108.

Crossley, S.A., Allen, L.K., and McNamara, D.S. 2014. A multi-dimensional analysis of essay writing. In: Sardinha, T.B. and Pinto, M.V. eds. *Multi-dimensional analysis, 25 years on: A tribute to Douglas Biber*, Amsterdam-Philadelphia: John Benjamins Publishing Company, pp. 197–237.

Crossley, S.A., Salsbury, T., McNamara, D.S. and Jarvis, S. 2011a. What is lexical proficiency? Some answers from computational models of speech data. *TESOL Quarterly*, **45**(1), pp.182-193.

Crossley, S.A., Allen, D.B. and McNamara, D.S. 2011b. Text readability and intuitive simplification: A comparison of readability formulas. *Reading in a foreign language*, **23**(1), pp.84-101.

Crossley, S.A. and McNamara, D.S. 2011c. Understanding expert ratings of essay quality: Coh-Metrix analyses of first and second language writing. *International Journal of Continuing Engineering Education and Life Long Learning*, **21**(2-3), pp.170-191.

Crossley, S.A., Weston, J.L., McLain Sullivan, S.T. and McNamara, D.S. 2011d. The development of writing proficiency as a function of grade level: A linguistic analysis. *Written Communication*, **28**(3), pp.282-311.

Crossley, S.A., Louwerse, M.M., McCarthy, P.M. and McNamara, D.S. 2007. A linguistic analysis of simplified and authentic texts. *The Modern Language Journal*, **91**(1), pp.15-30.

Crossley, S.A., Skalicky, S., Dascalu, M., McNamara, D.S. and Kyle, K. 2017. Predicting text comprehension, processing, and familiarity in adult readers: New approaches to readability formulas. *Discourse Processes*, **54**(5-6), pp.340-359.

Dale, E. and Tyler, R. W. 1934. A study of the factors influencing the difficulty of reading materials for adults of limited reading ability. *The Library Quarterly*, **4**(3), pp.384-412.

Dale, E. and Chall, J. 1949. The concept of readability. *Elementary English*, ***26***(1), pp.19-26.

Dale, E. and Chall, J. 1948. A Formula for Predicting Readability. *Educational Research Bulletin*, **27**(1), 11-28.

Dale, E., and Chall. J. 1995. *Readability revisited, the new Dale-Chall readability formula*. Cambridge, MA: Brookline Books.

Dastjerdi, H. V., and Jamshidian, E. 2011. A Sacrament Wordplay: An Investigation of Pun Translatability in the two English Translations of the Qur'ān. *Asian Social Science*, **7**(1), pp. 131-144.

Davison, A. 1984. Readability formulas and comprehension. In: Duffy, G. G., Roehler, L. R. and Mason, J. eds. *Comprehension instruction: Perspectives and suggestions*. New York: Longman, pp. 128-144.

Davison, A. 1986. Readability: The situation today. *Reading education report No. 70*, University of Illinois.

Dawood, B. 1977. *The relationship between readability and selected language variables*. MA thesis, University of Bagdad.

Dawood, N. J. 1990. *The Koran: Translated with Notes*. 5th ed. London: Penguin Books.

Deane-Cox, S. 2014. *Retranslation: Translation, Literature and Reinterpretation*. London: Bloomsbury Publishing.

Dey, A. 2012. Bengali Translation of the Qur'ān and the Impact of Print Culture on Muslim Society in the Nineteenth Century. *Societal Studies*, **4**(4), pp.1299-1315.

Dickins, J., Hervey, S. and Higgins, I. 2017. *Thinking Arabic translation: A course in translation method: Arabic to English*. 2nd ed. New York: Routledge.

Dickins J. 2017. The pervasiveness of coordination in Arabic, with reference to Arabic>English translation. *Languages in Contrast*. **17**(2), pp. 229-254.

Dörnyei, Z. 2007. *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. London: Oxford University Press.

Dörnyei, Z. and Taguchi, T. 2009. *Questionnaires in second language research: Construction, administration, and processing*. New York: Routledge.

Dowell, N.M., Graesser, A.C. and Cai, Z. 2016. Language and discourse analysis with Coh-Metrix: Applications from educational material to learning environments at scale. *Journal of Learning Analytics*, **3**(3), pp.72-95.

DuBay, W.H. 2004. The Principles of Readability. [Online]. [Accessed 25 August 2017]. Available from: http://www.impact-information.com/impactinfo/readability02.pdf.

DuBay, W.H. 2007. The Classic Readability Studies. [Online]. [Accessed 25 August 2017]. Available from: https://files.eric.ed.gov/fulltext/ED506404.pdf.

Duran, N. D., Bellissens, C., Taylor, R., and McNamara, D. S. 2007. Quantifying text difficulty with automated indices of cohesion and semantics. In: McNamara, D. S., Trafton, G. eds. *Proceedings of the 29th Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society, pp. 233–238.

Duran, N.D., McCarthy, P.M., Graesser, A.C. and McNamara, D.S. 2007. Using temporal cohesion to predict temporal coherence in narrative and expository texts. *Behavior Research Methods*, **39**(2), pp.212-223.

El Damanhoury, Y. 2015. *Translation of the Qur'ān from Arabic to Japanese: A Study of Translation Techniques Usage in Translating Cultural References*. MA dissertation, Leiden University.

El-Haj, M. and Rayson, P. 2016. Osman – a novel Arabic readability metric. In: *10th Language Resources and Evaluation Conference*. [Online]. Slovenia, pp. 250–255. [Accessed 17 May 2017]. Available from: OSMAN: a novel Arabic readability metric - Lancaster EPrints (lancs.ac.uk).

Elimam, A.S. and Fletcher, A.S. 2021. *The Qur'an, Translation and the Media: A Narrative Account*. London and New York: Routledge.

Elliott, A. C., and Woodward, W. A. 2007. *Statistical Analysis Quick Reference Guidebook.* London: Sage Publications, Inc.

El-Zeiny, I., 2011. Criteria for the translation and assessment of Qur'ānic metaphor: A contrastive analytic approach. *Babel*, ***57***(3), pp.247-268.

Entin, E. B. and Klare, G. R. 1985. Relationships of measures of interest, prior knowledge, and readability to comprehension of expository passages. *Advances in Reading and Language Research*, **3**, pp. 9-38.

Fatani, A. 2006. Translation and the Qur'ān. In: Leaman, O. ed. *The Qur'ān: an encyclopedia*. London and New York: Routledge. pp. 657–669.

Fiederer, R. and O'Brien, S. 2009. Quality and machine translation: A realistic objective. *The Journal of Specialised Translation*, **11**, pp.52-74.

Flesch, R. 1948. A new readability yardstick. *Journal of Applied Psychology*, **32**(3), 221–233.

Flesch, R. 1949. *The art of readable writing*. New York: Harper.

Foltz, P.W. 2007. Discourse coherence and LSA. In: Landauer, T.K., McNamara, D.S., Dennis, S. and Kintsch, W. eds. *Handbook of latent semantic analysis*. Mahwah, NJ: Erlbaum, pp. 167–184.

François, T. 2015. When readability meets computational linguistics: a new paradigm in readability. *Revue française de linguistique appliquée*, ***20***(2), pp.79-97.

Frey, L.R., Botan, C.H., and Kreps, G.L. 2000. *Investigating communication: An introduction to research methods.* 2nd ed. Boston: Allyn and Bacon.

Fry, E. B. 1988. Writeability: the principles of writing for increased comprehension. In: Zakaluk, B.L. and Samuels, S.J. *Readability: Its past, present, and future*. Newark, DE: International Reading Association, pp.77-97.

Gilliland, J. 1972. *Readability*. London: University of London Press Ltd for the United Kingdom Reading Association.

Gokkır, B. 2006. Origins, Organization and Influences on the Qur'ān. In: Leaman, O. ed. *The Qur'ān: an encyclopedia*. London and New York: Routledge. pp. 474–480.

Göpferich, S. 2009. Comprehensibility assessment using the Karlsruhe comprehensibility concept. *The Journal of Specialised Translation*, **11**, pp.31–52.

Graesser, A.C., Cai, Z., Louwerse, M., and Daniel, F. 2006. Question Understanding Aid (QUAID): A web facility that helps survey methodologists improve the comprehensibility of questions. *Public Opinion Quarterly*, *70*(1), 3–22.

Graesser, A. C., McNamara, D.S., Cai, Z., Conley, M., Li, H., and Pennebaker, J. 2014. Coh-Metrix measures text characteristics at multiple levels of language and discourse. *Elementary School Journal*, **115**(2), pp.210-229.

Graesser, A.C., McNamara, D.S., Louwerse, M.M., and Cai, Z. 2004. Coh-Metrix: Analysis of text on cohesion and language. Behavior Research Methods, Instruments, & Computers: *A Journal of the Psychonomic Society*, **36**(2), 193–202.

Graesser, A.C. and McNamara, D.S. 2011. Computational analyses of multilevel discourse comprehension. *Topics in cognitive science*, **3**(2), pp.371-398.

Graesser, A.C., McNamara, D.S. and Louwerse, M.M. 2011. Methods of automated text analysis. In Kamil, M. L., Pearson, D., Moje, E. B., Afflerbach, P. eds. *Handbook of reading research*. London and New York: Routledge, pp. 34-53.

Gray, W.S., and Leary, B.E. 1935. *What makes a book readable*? Chicago, IL: University of Chicago Press.

Green, M.C., and Brock, T.C. 2002. In the mind's eye: Transportation-imagery model of narrative persuasion. In: Green, M.C., Strange, JJ., and Brock T.C. eds., *Narrative impact: Social and cognitive foundations*. New York: Lawrence Erlbaum Associates Publishers, pp. 315-341.

Gürçağlar, Ş.T. 2009. Retranslation. In: Baker, M. and Saldanha, G. eds. *Routledge Encyclopaedia of Translation Studies*. London: Routledge, pp, 233–235.

Halliday, T.C., and Briss E.A. 1977. The Evaluation and Systems Analysis of the Systran Machine Translation System. Report RADC-TR-76-399, Rome Air Development Center, Griffiss Air Force Base, NY.

Hansen-Schirra, S. and Gutermuth, S., 2015. Approaching comprehensibility in translation studies. In: Maksymski, K., Gutermuth, G., and Hansen-Schirra, S. Eds. *Translation and comprehensibility*. Berlin: Frank and Timme, pp. 53-76.

Harrison, C. 1980. *Readability in the classroom*. London: Cambridge University Press.

Harrison, C. 1986. *Readability in the United Kingdom*. *Journal of Reading*, **29**(6), pp.521-529.

Hartley, J. and Trueman, M. 1983. The effects of headings in text on recall, search and retrieval. *British journal of educational psychology*, **53**(2), pp.205-214.

Hartley, J. and Trueman, M. 1982. The effects of summaries on the recall of information from prose: Five experimental studies. *Human Learning*, **1**(1), 63–82.

Hesse-Biber, S.N., Rodriguez, D. Frost, N.A. 2015. A qualitatively driven approach to multimethod and mixed methods research. In: Hesse-Biber, S.N. and Johnson, R.B. eds. *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*, New York: Oxford University Press, pp. 3-20.

House, J.1997. *Translation quality assessment: A model revisited*. Germany: Gunter Narr Verlag.

Hussain, M. 2020. *The Majestic Qur'an: A plain English translation*. Nottingham: Invitation Publishing.

Irving, T.B. 1992. *The Noble* Qur'ān*: The First American Translation and Commentary*. Brattleboro: Amana Books.

Jassem, Z.A. 2014. The Noble Qur'ān: A Critical Evaluation of Al-Hilali and Khan's Translation. *International Journal of English and Education*, **3**(2), pp. 237- 273.

Jianzhong, X. 2003. Retranslation: necessary or unnecessary. *Babel*, **49**(3), pp.193-202.

Johnstone, B. 1990. 'Orality' and discourse structure in Modern Standard Arabic. In: Mushira E. ed. *Perspectives on Arabic linguistics*. Amsterdam: John Benjamins Publishing Co, pp.215-233.

Jones, M.J. 1997. Methodological themes: Critical appraisal of the Cloze procedure's use in the accounting domain. *Accounting, Auditing and Accountability Journal*, **10**(1), pp.105-128.

Kamalski, J., Sanders, T. and Lentz, L. 2008. Coherence marking, prior knowledge, and comprehension of informative and persuasive texts: Sorting things out. *Discourse Processes*, **45**(4-5), pp.323-345.

Kandula, S. and Zeng-Treitler, Q. 2008. Creating a gold standard for the readability measurement of health texts. In: *Proceedings of the AMIA Annual Symposium Proceedings, 8/12 November 2008, Washington, USA*, pp. 353-357.

Keenan, J.M., Baillet, S.D. and Brown, P. 1984. The effects of causal cohesion on comprehension and memory. *Journal of verbal learning and verbal behavior*, **23**(2), pp.115-126.

Kidwai, A.R. 1987. Translating the untranslatable: a survey of English translations of the Qur'ān. *The Muslim World Book Review*, **7**(4), pp.66-71.

Kidwai, A.R. 2007. *Bibliography of the Translations of the Meanings of the Glorious* Qur'ān *into English, 1649-2002: A Critical Study*. Madinah: King Fahd Qur'ān Printing Complex.

Kidwai, A.R. 2018. *God's Word, Man's Interpretations: A Critical Study of the 21st Century English Translations of the Qur'ān*. New Delhi: Viva Books.

Kintsch, W., & Vipond, D. 1979. Reading comprehension and readability in educational practice and psychological theory. In: Nilsson L.G. ed. *Perspectives of memory research*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc, pp. 325-366.

Kintsch, W. and Miller, J.R. 1981. Readability: A view from cognitive psychology. In: Flood, J. ed. *Understanding reading comprehension*. Newark, DE: International Reading, pp.220-232.

Klare, G.R. 1988. The formative years. In: Zakaluk, B.L. and Samuels, S.J. *Readability: Its past, present, and future*. Newark, DE: International Reading Association, pp.14-35.

Klare, G.R. 1963. *The measurement of readability*. Ames: Iowa State Press.

Klare, G.R. 1974. Assessing readability. *Reading research quarterly*,10 (1), pp.62-102.

Klare, G.R. 1976. A second look at the validity of readability formulas. *Journal of reading behavior*, **8**(2), pp.129-152.

Klare, G.R. 1968. The Role of Word Frequency in Readability. *Elementary English*, **45**(1), 12-22.

Kolahi, S., Khanmohammad, H. and Shirvani, E. 2013. A Comparison of the Application of Readability Formulas in English Translation Textbooks and Their Translations. *International Journal of English Language Education*, **1**(1), pp.140-161.

Koskinen, K. and Paloposki, O. 2004. Thousand and One Translations: Retranslation Revisited. In: Gyde, H., Kirsten, M. and Daniel G. eds. *Claims, Changes and Challenges in Translation Studies*. Amsterdam/Philadelphia: John Benjamins, pp. 27-38.

Koskinen, K. and Paloposki, O. 2010. Retranslation. In: Gambier Y., and Doorslaer L. eds. *Handbook of Translation Studies*. Amsterdam: John Benjamins, pp. 294–298.

Koskinen, K. and Paloposki, O. 2003. Retranslations in the Age of Digital Reproduction. *Cadernos de tradução*, **1**(11), pp.19-38.

Kuhn, T.S. 1961. The function of measurement in modern physical science. *Isis*, **52**(2), pp. 161–193.

Lacunza-Balda, J. 1997. Translations of the Qur'ān into Swahili, and Contemporary Islamic Revival in East Africa. In: Westerland, D. and Rosander, E. eds. *African Islam and Islam in Africa: Encounters Between Sufis and Islamists*. London: Hurst, pp 95-126.

Laerd Statistics. 2015. Statistical tutorials and software guides. [Online]. [Accessed 4 April 2020]. Available from: https://statistics.laerd.com.

Li, H., Graesser, A. and Cai., Z. 2014. Comparison of Google translation with human translation. In Eberle, W. and Boonthum-Denecke C. eds. *Proceedings of the Twenty-seventh International Florida Artificial Intelligence Research Society Conference.* Palo Alto, CA: AAAI Press, pp. 190-195.

Li, H., Cheng, C., Graesser, A.C. 2015. A measure of text formality as a human construct. In: Russel, I. and Eberle, B. eds. *Proceedings of the Twenty-Eighth International Florida Artificial Intelligence Research Society Conference*. Palo Alto: AAAI Press, pp. 175–180.

Lightman, E.J., McCarthy, P.M., Dufty, D.F. and McNamara, D.S. 2007. Cohesion and Structural Organization in High School Texts. *In Proceedings of the Twentieth International Florida Artificial Intelligence Research Society Conference, 7/9 May 2007*, Florida. pp. 235-240.

Lively, B.A. and Pressey, S.L. 1923. A method for measuring the vocabulary burden of textbooks. *Educational administration and supervision*, **9**(7), pp.389-398.

Loimeier, R. 2005. Translating the Qur'ān in Sub-Saharan Africa: Dynamics and Disputes. *Journal of Religion in Africa*, **35**(4), 403-423.

Lorge, I. 1939. Predicting reading difficulty of selections for children. *The Elementary English Review*, **16**(6), pp.229-233.

Louwerse, M.M., McCarthy, P.M., McNamara, D.S. and Graesser, A.C. 2004. Variation in language and cohesion across written and spoken registers. In: Forbus, K., Gentner, D. and Regier T., eds. *Proceedings of the 26th Annual Conference of the Cognitive Science Society* Mahwah, NJ: Erlbaum, pp. 843-848.

Maksymski, K., Gutermuth, S., and Hansen-Schirra, S. 2015. Preface. In: Maksymski, K., Gutermuth, G., and Hansen-Schirra, S. eds. *Translation and comprehensibility*. Berlin: Frank and Timme, pp. 1-7.

Mark, M.M., 2015. Mixed and multimethods in predominantly quantitative studies, especially experiments and quasi-experiments. In: Hesse-Biber, S. and Johnson, R.B. eds. *The Oxford handbook of multimethod and mixed methods research inquiry*. London: Oxford University Press, pp. 21-41.

Mark, M.M., and Shotland, R.L. 1987. Alternative models for the use of multiple methods. In: Mark, M.M. and Shotland, R.L. eds. *Multiple methods in program evaluation*. San Francisco, CA: Jossey-Bass, pp. 95-100.

Mateen, M.A. and Ali, S. 2015. New Trends in Japanese Scholarship on Islam. *Ma'arif Research Journal*, **1**(10), pp. 53-60.

McCarthy, P.M., Graesser, A.C., and McNamara, D.S. 2006a. Distinguishing genre using Coh-metrix indices of cohesion. [Poster]. In 16th Annual Meeting of the Society for Text and Discourse. Minneapolis, MN.

McCarthy, P.M., Lewis, G.A., Dufty, D.F., and McNamara, D. S. 2006b. Analyzing writing styles with Coh-Metrix. In Sutcliffe, G.C.J. and Goebel, R. G. eds. *Proceedings of the 19th annual Florida Artificial Intelligence Research Society International conference*. Melbourne Beach, FL: AAAI Press, pp. 764–770.

McCarthy, P.M., Lehenbauer, B.M., Hall, C., Duran, N.D., Fujiwara, Y. and McNamara, D.S. 2007. A Coh-Metrix analysis of discourse variation in the texts of Japanese, American, and British Scientists. *Foreign Languages for Specific Purposes*, **6**, pp.46-77.

McLaughlin, G.H. 1969. SMOG grading – a new readability formula. *Journal of reading*, **12**(8), pp.639-646.

McNamara, D.S., Graesser, A.C., McCarthy, P.M. and Cai, Z. 2014. *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge: Cambridge University Press.

McNamara, D.S., Graesser, A.C., Cai, Z. and Kulikowich, J.M. 2011. Coh-Metrix easability components: Aligning text difficulty with theories of text comprehension. In *annual meeting of the American Educational Research Association,* New Orleans, pp. 1-14.

McNamara, D.S., Louwerse, M.M., McCarthy, P.M. and Graesser, A.C. 2010. Coh-Metrix: Capturing linguistic features of cohesion. *Discourse Processes,* **47**(4), pp.292-330.

Mellinger, C. and Hanson, T. 2016. *Quantitative research methods in translation and interpreting studies*. New York: Routledge.

Millis, K., Graesser, A.C., and Haberlandt, K. 1993. The impact of connectives on memory for expository texts. *Applied Cognitive Psychology*, ***7***(4), pp.317-339.

Mohammed, K. 2005. Assessing English Translations of the Qur'an. *Middle East Quarterly Forum,* **12**(2), pp. 58-71.

Moravcsik, J.E., and Kintsch, W. 1993. Writing quality, reading skills, and domain knowledge as factors in text comprehension. *Canadian Journal of Experimental Psychology*, ***47***(2), 370-374

Mujiyanto, Y. 2016. The comprehensibility of readable English texts and their back-translations. *International journal of English linguistics*, **6**(2), p.21-33.

Mustapha H. 2009. Qur'an (Koran). In: Baker, M. and Saldanha, G. eds. *Routledge Encyclopaedia of Translation Studies*. London: Routledge, 225–230.

Naudé, J.A. and Miller-Naudé, C.L. 2011. Colonial and postcolonial encounters with the indigenous: The case of religious translation in Africa. *Southern African Linguistics and Applied Language Studies*, **29**(3), pp.313-326.

Newmark, P. 1981. *Approaches to translation.* Oxford: Pergamon Press.

Newmark, P. 1988. *A Textbook of Translation*. New York: London: Prentice Hall.

Nisbeth Jensen, M. 2015. Optimising comprehensibility in interlingual translation: The need for intralingual translation. In: Maksymski, K., Gutermuth, G., and Hansen-Schirra, S. Eds. *Translation and comprehensibility*. Berlin: Frank and Timme, pp.163-194.

O'Brien, S. 2010. Controlled language and readability. *Translation and cognition*, **28**(*15*), pp.143-168.

Osgood, C.E., Suci, G.J. and Tannenbaum, P.H. 1957. *The measurement of meaning*. Illinois: Ninth Printing.

O'Reilly, T. and McNamara, D. S. 2007. Reversing the reverse cohesion effect: good texts can be better for strategic, high-knowledge readers. *Discourse Processes*, **43**(2), pp.121-152.

Oxford Advanced Learner's Dictionary. [Online]. 10th ed. 2020. [Accessed 7 October 2020]. Available from: www.oxfordlearnersdictionaries.com.

Ozuru, Y., Dempsey, K. and McNamara, D.S. 2009. Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and instruction*, **19**(3), pp. 228-242.

Paivio, A., Yuille, J.C. and Madigan, S.A., 1968. Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of experimental psychology*, **76**(1), pp. 1-25.

Pickthall, M. 1930. *The Meaning of the Glorious Koran: An Explanatory Translation*. London: George Allen and Unwin Ltd.

Pikulski, J. 2002. Readability. [Online]. [Accessed 12 October 2020]. Available from: http://www.eduplace.com/state/author/pikulski.pdf.

Pitler, E. and Nenkova, A. 2008. Revisiting readability: A unified framework for predicting text quality. *In Proceedings of the 2008 conference on empirical methods in natural language processing*, October2008. [Online]. Honolulu, pp. 186–195. [Accessed 25 May 2019]. Available from: Revisiting Readability: A Unified Framework for Predicting Text Quality (aclanthology.org).

Poonawala, I.K. 1990. Translatability of the Qur'ān: Theological and Literary Considerations. In: David, M.G. ed. *Translation of Scripture*. Philadelphia, PA: Annenberg Research Institute, pp. 161–192.

Preissle, J., Glover-Kudon, R., Rohan, E. A., Boehm, J. E. and DeGroff, A. 2015. Putting ethics on the mixed methods map. In Hesse-Biber, S. and Johnson, R.B. eds. *The Oxford handbook of multimethod and mixed methods research inquiry*. Oxford: Oxford University Press, pp.144-166.

Pym, A. 1998. *Method in Translation History*. Manchester: St. Jerome.

Rabin, A.T. 1988. Determining Difficulty Levels of Text Written in Languages Other than English. In: Zakaluk, B.L. and Samuels, S.J. *Readability: Its past, present, and future*. Newark, DE: International Reading Association, pp.46-76.

Rahayu, E.T. and Fauzan, A. 2020. The Language Choice as A Reflection of Islamic Communication in The Qur'ān-Japanese Translation. *Madania: Jurnal Kajian Keislaman,* **24**(1), 73 - 82.

Rasinger, S.M. 2013. *Quantitative research in linguistics: An introduction*. 2nd ed. London: Continuum International Publishing Group.

Redish, J. 2000. Readability formulas have even more limitations than Klare discusses. *ACM Journal of Computer Documentation*, **24**(3), pp.132-137.

Richardson, J.T. 1975. Concreteness and imageability. *The Quarterly Journal of Experimental Psychology*, **27**(2), pp.235-249.

Riddell, P.G. 2014. Translating the Qur'an into Indonesian languages. *Al-Bayan: Journal of Qur'an and Hadith Studies*, **12**(1), pp.1-27.

Robinson, N. 1997. Sectarian and ideological bias in Muslim translations of the Qur'an. *Islam and Christian-Muslim Relations*, **8**(3), pp.261-278.

Ross, A. 1649. *The Alcoran of Mahomet: Translated out of Arabique into French*. London: Anno Dom.

Saddiki, H., Bouzoubaa, K. and Cavalli-Sforza, V. 2015. Text readability for Arabic as a foreign language. *In 2015 IEEE/ACS 12th International Conference of Computer Systems and Applications, 17/20 Nov*, Marrakech, Morocco. [Online] IEEE, pp. 1-8. [Accessed 17 May 2017]. Available from: 10.1109/AICCSA.2015.7507232.

Saheeh International, 1997. *The Qur'an: English Meanings*. Jeddah: Abul-Qasim Publishing House.

Saldanha, G. and O'Brien, S. 2014. *Research methodologies in translation studies*. Manchester: St. Jerome.

Sale, G. 1888. *The Koran*. London: Frederick Warne.

Schriver, K.A. 1989. Evaluating text quality: The continuum from text-focused to reader-focused methods. *IEEE Transactions on professional communication*, **32**(4), pp.238-255.

Schwanenflugel, P.J. and Stowe, R.W. 1989. Context availability and the processing of abstract and concrete words in sentences. *Reading Research Quarterly*, **24**(1) pp.114-126.

Shah, M.S. 2017. Arthur John Arberry as an Interpreter of the Holy Qur'an. *Quarterly Research Journal of the Faculty of Oriental Learning*. **30**, pp.3-18.

Sheikh, M. 2020. The Qadizadelis and Sufism'. In L. Ridgeon, ed. *Handbook of Sufism*. London: Routledge, pp.418-432.

Shuttleworth, M. and Cowie, M. 1997. *Dictionary of Translation Studies*. Manchester: St Jerome.

Sinaiko, H.W. and Klare, G.R. 1972. Further experiments in language translation: readability of computer translations. *ITL-International Journal of Applied Linguistics*, **15**(1), pp.1-29.

Singer, H. 1988. Foreword. In: Zakaluk, B.L. and Samuels, S.J. *Readability: Its past, present, and future*. Newark, DE: International Reading Association, vii-ix.

Smith, M. and Taffler, R. 1992. Readability and understandability: Different measures of the textual complexity of accounting narrative. *Accounting, Auditing & Accountability Journal*, **5**(4), pp. 84-98.

Solihu, A.K.H. 2015. The earliest Yoruba translation of the Qur'an: missionary engagement with Islam in Yorubaland. *Journal of Qur'anic Studies*, **17**(3), pp.10-37.

Spira, I. 2005. *Chinese translations of the Qur'an: a close reading of selected passages.* MA thesis, Oslo University.

Stenner, A., Horabin, I., Smith, D., and Smith, M.1988. *The Lexile Framework*. Durham, NC; Metametrics, Inc.

Suojanen, T., Koskinen, K., Tuominen, T. 2014. *User-centred translation*. London: Routledge.

Susam-Saraeva, S. 2003. Multiple-Entry Visa for Travelling Theory: Retranslations of Literary and Cultural Theories, *Target: International Journal of Translation Studies*, **15**(1), pp. 1-36.

Taivalkoski-Shilov, K. 2015. Friday in Finnish: A character's and (re) translators' voices in six Finnish retranslations of Daniel Defoe's Robinson Crusoe. *Target. International Journal of Translation Studies*, **27**(1), pp.58-74.

Tannen, D.1982. Oral and literate strategies in spoken and written narratives. *Language*, **58**(1), pp.1-21.

Thomas, D. and Chesworth, J. 2017. Works on Christian-Muslim relations 1600-1700. In: Thomas, D. and Chesworth, J. eds. *Christian-Muslim Relations a Bibliographical History*. Boston: Brill, pp. 27–389.

Thorndike, E.L. 1921. *The Teacher's Word Book*. New York: Teacher's College, Columbia University.

Tiselius, E. 2009. Revisiting Carroll's scales. Testing and assessment in translation and interpreting studies, pp.95-121.

Tiselius, E. 2009. Revisiting Carroll's Scales. In Angelelli, C. and Jacobson, H.E. eds. *Testing and assessment in translation and interpreting studies*. Amsterdam: John Benjamins, pp. 95–121.

Venuti, L. 2004. Retranslations: The Creation of Value. In: Faull, K.M. ed. *Bucknell Review: Translation and Culture*. Cranbury: Associated University Presses, pp. 25-38.

Vogel, M. and Washburne, C. 1928. An objective method of determining grade placement of children's reading material. *The Elementary School Journal*, **28**(5), pp.373-381.

Wild, S. 2015. Muslim Translators and Translations of the Qur'an into English. *Journal of Qur'anic Studies*, **17**(3), pp.158-182.

Wilson, M. 2009. The First Translations of the Qur'an in Modern Turkey (1924-38). *International Journal of Middle East Studies*, **41**(3), 419-435.

Wolfer, S. 2015. *Comprehension and comprehensibility*. In: Maksymski, K., Gutermuth, G., and Hansen-Schirra, S. Eds. *Translation and comprehensibility*. Berlin: Frank and Timme, pp. 33-51.

Yacoub, M. 2005. *Asbāb Al'akhṭā' Fī Altafsīr: Dirāsah Taṣlīḥiyyah*. Damam: Dar Ibn Aljawzi.

Yang Huaizhong, M.U. 1996. The four upsurgences of Islamic culture in Chinese history. *Journal of Muslim Minority Affairs*, **16**(1), pp.9-20.

Zakaluk, B.L. and Samuels, S.J. 1988. Toward a new approach to predicting text comprehensibility. In: Zakaluk, B.L. and Samuels, S.J. *Readability: Its past, present, and future*. Newark, DE: International Reading Association, pp.121-144.

Zwaan, R.A. 1994. Effect of genre expectations on text comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **20**(4), 920–933.

Appendices

Table of contents

Appendix A: Participant Information Sheet

You are being invited to take part in a research project.
Please take time to read the following information carefully to help you decide whether or not you would like to participate in the study. It is important to understand why the research is being done and what it will involve.

What is the purpose of the project?
This project aims to evaluate two English translated chapters of the Qur'ān based on ease (easiness) of (i) comprehension, (ii) their linguistic elements such as words, grammar and cohesion, and (iii) their non-linguistic elements such as page layout.

Why have I been invited to take part?
You have been chosen as you are a native speaker of English. This project aims to examine the readability of the chosen translations for English readers.

What will happen to me if I take part?
You will be asked to fill in an online survey evaluating two English translated chapters of the Qur'ān. The survey will take around 15-20 minutes to complete and all responses are anonymous.

What happens to the information I give?
The answers you give will be stored anonymously on a secure University of Leeds server and will be used in the analysis of this research. After this research has been completed, the data will be kept securely for use in future research by the University of Leeds and their academic collaborators. All of the data obtained will be treated as confidential and stored securely as is required by the Data Protection Act. The data collected will be used as part of a doctoral thesis and may be written up for publication. No identifying information about you will be included in the report.

Do I have to take part in the study?
Taking part in this study is voluntary, and participants can withdraw at any time without giving a reason. However, any responses already provided will be retained due to anonymity of responses.

Will I be contacted as a result of anything I answer in the survey?
The survey is anonymous. Therefore you cannot be contacted regardless of what you answer in the survey.

How do I take part?
You can take part in this study by completing the consent form on the following page and answering the survey questions.

If I have questions about the study, who can I ask?
If you would like further information please contact us on the following details:

Abdulrahman Fahad Albalawi
Ph.D. Student
School of Languages, Cultures, and Societies
Woodhouse Lane,
Leeds LS2 9JT
Tel: 07455534533
e-mail: Ml14afga@leeds.ac.uk

My supervisor's details:
Prof. James Dickins
School of Languages, Cultures and Societies
Michael Sadler Building,
Leeds LS2 9JT
Tel: 0113 34 31311
e-mail: J.Dickins@leeds.ac.uk

To take part in the research please click 'next'.

Appendix B: Participant Consent Form

I confirm that I have read the information on the previous page and understood what I am being asked to do in this research.

I understand that my responses to the survey will remain confidential.

I understand that once my responses have been captured by Online Survey I will be unable to withdraw them.

I understand that my participation is voluntary and I can withdraw at any time without giving a reason.

I give consent to take part in this research and for my anonymised data to be stored and used in the analysis of this and future research.

If you understand the information provided and consent to taking part in the study please click on the next button and go on to complete the survey.

Appendix C: Participant Background Information

1.What is your gender? *Required*

- Male
- Female

2.What is your age? *Required*

- 18-24
- 25-34
- 35-44
- 45-54
- 55-64
- 65 or over

3.What is the highest level of formal education you have completed? *Required*

- High school graduate
- Bachelor's degree
- Master's degree
- Doctoral degree

4.What is your major or specialism? *Required*

5.What is your religion? *Required*

- Muslim
- Non-Muslim

6.To what extent do you have knowledge of the Qur'ān? *Required*

- Not at all
- To a small extent
- To a moderate extent
- To a great extent
- To a very great extent

If you have not chosen "not at all" for Q. 6, how do you evaluate your knowledge in the Qur'ān according to the options below?

	Very frequently	Frequently	Sometimes	Never
I read the Arabic Qur'ān	☐	☐	☐	☐
I read English translations	☐	☐	☐	☐
I read interpretations of the Qur'ān	☐	☐	☐	☐

7.What amount of knowledge of the Arabic language would you assess yourself as having? *Required*

- ○ None at all
- ○ A small amount
- ○ A moderate amount
- ○ A great amount
- ○ A very great amount

8.Is Arabic your native language?

- ○ Yes
- ○ No

9.What amount of knowledge of the English language would you assess yourself as having?

- ○ None at all
- ○ A small amount
- ○ A moderate amount
- ○ A great amount
- ○ A very great amount

10.Is English your native language?

- ○ Yes
- ○ No

Appendix D: Instructions for Completing the Survey

Before you start the survey, I will explain what you will need to do in the following pages:

1. You will be asked to read two chapters of the Qur'ān in English translation. These chapters are labelled Chapter 1 and Chapter 2 and they deal with different topics. Each chapter is followed by a separate table for evaluation. Some of these Qur'anic chapters have footnotes and an introduction for those who do not have background knowledge about the Qur'ān.
2. After you finish reading, you will be asked to choose on a scale from 1-5, with 1 being the best and 5 the worst, to rate the linguistic and non-linguistic elements.

E. Evaluating the readability of an English Qur'ān translation (Arberry's version)

Sample 1 (Chapter 91)

Read the following translation of a Qur'anic chapter completely and then move on to the next section to evaluate it. The numbers on the left are verse numbers.

XCI

THE SUN

In the Name of God, the Merciful, the Compassionate

By the sun and his morning brightness
and by the moon when she follows him,
and by the day when it displays him
and by the night when it enshrouds him!
By the heaven and That which built it
and by the earth and That which extended it!
By the soul, and That which shaped it
and inspired it to lewdness and godfearing!
Prosperous is he who purifies it,
and failed has he who seduces it.

Thamood cried lies in their insolence
when the most wretched of them uprose,
then the Messenger of God said to them,
'The She-camel of God; let her drink!'
But they cried him lies, and hamstrung her,
so their Lord crushed them for their sin, and levelled them:
and He fears not the issue thereof.

Q1. Overall, how would you rate the comprehension ease of Chapter 1 on the following criteria?

	1	2	3	4	5	
Comprehensible title	☐	☐	☐	☐	☐	Incomprehensible title
Comprehensible Verse 1	☐	☐	☐	☐	☐	Incomprehensible Verse 1
Comprehensible Verse 2	☐	☐	☐	☐	☐	Incomprehensible Verse 2
Comprehensible Verse 3	☐	☐	☐	☐	☐	Incomprehensible Verse 3
Comprehensible Verse 4	☐	☐	☐	☐	☐	Incomprehensible Verse 4
Comprehensible Verse 5	☐	☐	☐	☐	☐	Incomprehensible Verse 5
Comprehensible Verse 6	☐	☐	☐	☐	☐	Incomprehensible Verse 6
Comprehensible Verse 7	☐	☐	☐	☐	☐	Incomprehensible Verse 7
Comprehensible Verse 8	☐	☐	☐	☐	☐	Incomprehensible Verse 8
Comprehensible Verse 9	☐	☐	☐	☐	☐	Incomprehensible Verse 9
Comprehensible Verse 10	☐	☐	☐	☐	☐	Incomprehensible Verse 10
Comprehensible Verse 11	☐	☐	☐	☐	☐	Incomprehensible Verse 11
Comprehensible Verse 12	☐	☐	☐	☐	☐	Incomprehensible Verse 12
Comprehensible Verse 13	☐	☐	☐	☐	☐	Incomprehensible Verse 13
Comprehensible Verse 14	☐	☐	☐	☐	☐	Incomprehensible Verse 14
Comprehensible Verse 15	☐	☐	☐	☐	☐	Incomprehensible Verse 15

Q2. Overall, how would you rate the page layout of Chapter 1 on the following criteria?

	1	2	3	4	5	
Recognizable title	☐	☐	☐	☐	☐	Unrecognizable title
Short version	☐	☐	☐	☐	☐	Long version
Simple text layout	☐	☐	☐	☐	☐	Complex text layout
Familiar paragraph division	☐	☐	☐	☐	☐	Strange paragraph division
Clear verse numbers	☐	☐	☐	☐	☐	Unclear verse numbers
Clear text organization	☐	☐	☐	☐	☐	Unclear text organization

Q3. Overall, how would you rate the word usages of Chapter 1 on the following criteria?

	1	2	3	4	5	
Familiar words	☐	☐	☐	☐	☐	Strange words
Modern words	☐	☐	☐	☐	☐	Archaic words
Frequent English words	☐	☐	☐	☐	☐	Infrequent English words
Clear words	☐	☐	☐	☐	☐	Confusing words
Short words	☐	☐	☐	☐	☐	Long words

Q4. Overall, how would you rate the grammatical usages of Chapter 1 on the following criteria?

	1	2	3	4	5	
Short sentences	☐	☐	☐	☐	☐	Long sentences
Familiar English grammar	☐	☐	☐	☐	☐	Strange English grammar
Simple sentences	☐	☐	☐	☐	☐	Complex sentences
Clear sentences	☐	☐	☐	☐	☐	Confusing sentences

Q5. Overall, how would you rate the cohesion usages (links between elements) of Chapter 1 on the following criteria?

	1	2	3	4	5	
Coherent words	☐	☐	☐	☐	☐	Incoherent words
Coherent sentences	☐	☐	☐	☐	☐	Incoherent sentences
Familiar conjunctions	☐	☐	☐	☐	☐	Strange conjunctions

Sample 2 (Chapter 86)

Read the following translation of a Qur'anic chapter completely and then move on to the next section to evaluate it. The numbers on the left are verse numbers.

LXXXVI

THE NIGHT-STAR

In the Name of God, the Merciful, the Compassionate

By heaven and the night-star!
And what shall teach thee what is the night-star?
The piercing star!
Over every soul there is a watcher.

So let man consider of what he was created;
he was created of gushing water
issuing between the loins and the breast-bones.
Surely He is able to bring him back
upon the day when the secrets are tried,
and he shall have no strength, no helper.

By heaven of the returning rain,
by earth splitting with verdure,
surely it is a decisive word;
it is no merriment.

They are devising guile,
and I am devising guile.
So respite the unbelievers; delay with them awhile.

Q1. Overall, how would you rate the comprehension ease of Chapter 2 on the following criteria?

	1	2	3	4	5	
Comprehensible title	☐	☐	☐	☐	☐	Incomprehensible title
Comprehensible Verse 1	☐	☐	☐	☐	☐	Incomprehensible Verse 1
Comprehensible Verse 2	☐	☐	☐	☐	☐	Incomprehensible Verse 2
Comprehensible Verse 3	☐	☐	☐	☐	☐	Incomprehensible Verse 3
Comprehensible Verse 4	☐	☐	☐	☐	☐	Incomprehensible Verse 4
Comprehensible Verse 5	☐	☐	☐	☐	☐	Incomprehensible Verse 5
Comprehensible Verse 6	☐	☐	☐	☐	☐	Incomprehensible Verse 6
Comprehensible Verse 7	☐	☐	☐	☐	☐	Incomprehensible Verse 7
Comprehensible Verse 8	☐	☐	☐	☐	☐	Incomprehensible Verse 8
Comprehensible Verse 9	☐	☐	☐	☐	☐	Incomprehensible Verse 9
Comprehensible Verse 10	☐	☐	☐	☐	☐	Incomprehensible Verse 10
Comprehensible Verse 11	☐	☐	☐	☐	☐	Incomprehensible Verse 11
Comprehensible Verse 12	☐	☐	☐	☐	☐	Incomprehensible Verse 12
Comprehensible Verse 13	☐	☐	☐	☐	☐	Incomprehensible Verse 13
Comprehensible Verse 14	☐	☐	☐	☐	☐	Incomprehensible Verse 14
Comprehensible Verse 15	☐	☐	☐	☐	☐	Incomprehensible Verse 15
Comprehensible Verse 16	☐	☐	☐	☐	☐	Incomprehensible Verse 16
Comprehensible Verse 17	☐	☐	☐	☐	☐	Incomprehensible Verse 17

Q2. Overall, how would you rate the page layout of Chapter 2 on the following criteria?

	1	2	3	4	5	
Recognizable title	☐	☐	☐	☐	☐	Unrecognizable title
Short version	☐	☐	☐	☐	☐	Long version
Simple text layout	☐	☐	☐	☐	☐	Complex text layout
Familiar paragraph division	☐	☐	☐	☐	☐	Strange paragraph division
Clear verse numbers	☐	☐	☐	☐	☐	Unclear verse numbers
Clear text organization	☐	☐	☐	☐	☐	Unclear text organization

Q3. Overall, how would you rate the word usages of Chapter 2 on the following criteria?

	1	2	3	4	5	
Familiar words	☐	☐	☐	☐	☐	Strange words
Modern words	☐	☐	☐	☐	☐	Archaic words
Frequent English words	☐	☐	☐	☐	☐	Infrequent English words
Clear words	☐	☐	☐	☐	☐	Confusing words
Short words	☐	☐	☐	☐	☐	Long words

Q4. Overall, how would you rate the grammatical usages of Chapter 2 on the following criteria?

	1	2	3	4	5	
Short sentences	☐	☐	☐	☐	☐	Long sentences
Familiar English grammar	☐	☐	☐	☐	☐	Strange English grammar
Simple sentences	☐	☐	☐	☐	☐	Complex sentences
Clear sentences	☐	☐	☐	☐	☐	Confusing sentences

Q5. Overall, how would you rate the cohesion usages (links between elements) of Chapter 2 on the following criteria?

	1	2	3	4	5	
Coherent words	☐	☐	☐	☐	☐	Incoherent words
Coherent sentences	☐	☐	☐	☐	☐	Incoherent sentences
Familiar conjunctions	☐	☐	☐	☐	☐	Strange conjunctions

F. Evaluating the readability of an English Qur'ān translation (Abdel-Haleem's version)

Sample 1 (Chapter 91)

Read the following translation of a Qur'anic chapter completely and then move on to the next section to evaluate it. The numbers on the left are verse numbers.

91. THE SUN

A Meccan sura, the central theme of which is purifying or corrupting the soul, with the tribe of Thamud given as an example of corruption.

In the name of God, the Lord of Mercy, the Giver of Mercy

[1] By the sun in its morning brightness [2] and by the moon as it follows it, [3] by the day as it displays the sun's glory [4] and by the night as it conceals it, [5] by the sky and how He built it [6] and by the earth and how He spread it, [7] by the soul and how He formed it [8] and inspired it [to know] its own rebellion and piety! [9] The one who purifies his soul succeeds [10] and the one who corrupts it fails. [11] In their arrogant cruelty, the people of Thamud[a] called [their messenger] a liar, [12] when the most wicked man among them rose [against him].[b] [13] The messenger of God said to them, '[Leave] God's camel to drink,' [14] but they called him a liar and hamstrung her. Their Lord destroyed them for their crime and levelled them. [15] He did not hesitate[c] to punish[d] them.

[a] See e.g. 7: 73–9 (on the tribe of Thamud).
[b] Cf. 54: 29.
[c] Literally 'he does not fear'.
[d] One of the lexical meanings of '*uqba* is *jaza*', here 'to punish'; or 'does not fear the consequences'.

Q1. Overall, how would you rate the comprehension ease of Chapter 1 on the following criteria?

	1	2	3	4	5	
Comprehensible title	☐	☐	☐	☐	☐	Incomprehensible title
Helpful introduction for text comprehension	☐	☐	☐	☐	☐	Unhelpful introduction for text comprehension
Comprehensible Verse 1	☐	☐	☐	☐	☐	Incomprehensible Verse 1
Comprehensible Verse 2	☐	☐	☐	☐	☐	Incomprehensible Verse 2
Comprehensible Verse 3	☐	☐	☐	☐	☐	Incomprehensible Verse 3
Comprehensible Verse 4	☐	☐	☐	☐	☐	Incomprehensible Verse 4
Comprehensible Verse 5	☐	☐	☐	☐	☐	Incomprehensible Verse 5
Comprehensible Verse 6	☐	☐	☐	☐	☐	Incomprehensible Verse 6
Comprehensible Verse 7	☐	☐	☐	☐	☐	Incomprehensible Verse 7
Comprehensible Verse 8	☐	☐	☐	☐	☐	Incomprehensible Verse 8
Comprehensible Verse 9	☐	☐	☐	☐	☐	Incomprehensible Verse 9
Comprehensible Verse 10	☐	☐	☐	☐	☐	Incomprehensible Verse 10
Comprehensible Verse 11	☐	☐	☐	☐	☐	Incomprehensible Verse 11
Comprehensible Verse 12	☐	☐	☐	☐	☐	Incomprehensible Verse 12
Comprehensible Verse 13	☐	☐	☐	☐	☐	Incomprehensible Verse 13
Comprehensible Verse 14	☐	☐	☐	☐	☐	Incomprehensible Verse 14
Comprehensible Verse 15	☐	☐	☐	☐	☐	Incomprehensible Verse 15
Helpful footnotes for text comprehension	☐	☐	☐	☐	☐	Unhelpful footnotes for text comprehension

Q2. Overall, how would you rate the page layout of Chapter 1 on the following criteria?

	1	2	3	4	5	
Recognizable title	☐	☐	☐	☐	☐	Unrecognizable title
Short version	☐	☐	☐	☐	☐	Long version
Simple text layout	☐	☐	☐	☐	☐	Complex text layout
Familiar paragraph division	☐	☐	☐	☐	☐	Strange paragraph division
Clear verse numbers	☐	☐	☐	☐	☐	Unclear verse numbers
Clear text organization	☐	☐	☐	☐	☐	Unclear text organization

Q3. Overall, how would you rate the word usages of Chapter 1 on the following criteria?

	1	2	3	4	5	
Familiar words	☐	☐	☐	☐	☐	Strange words
Modern words	☐	☐	☐	☐	☐	Archaic words
Frequent English words	☐	☐	☐	☐	☐	Infrequent English words
Clear words	☐	☐	☐	☐	☐	Confusing words
Short words	☐	☐	☐	☐	☐	Long words

Q4. Overall, how would you rate the grammatical usages of Chapter 1 on the following criteria?

	1	2	3	4	5	
Short sentences	☐	☐	☐	☐	☐	Long sentences
Familiar English grammar	☐	☐	☐	☐	☐	Strange English grammar
Simple sentences	☐	☐	☐	☐	☐	Complex sentences
Clear sentences	☐	☐	☐	☐	☐	Confusing sentences

Q5. Overall, how would you rate the cohesion usages (links between elements) of Chapter 1 on the following criteria?

	1	2	3	4	5	
Coherent words	☐	☐	☐	☐	☐	Incoherent words
Coherent sentences	☐	☐	☐	☐	☐	Incoherent sentences
Familiar conjunctions	☐	☐	☐	☐	☐	Strange conjunctions

Sample 2 (Chapter 86)

Read the following translation of a Qur'anic chapter completely and then move on to the next section to evaluate it. The numbers on the left are verse numbers.

86. THE NIGHT-COMER

A Meccan sura that focuses on a series of examples of things coming out: the piercing night-star, spurting semen, the baby that bursts out of the womb, and plants that sprout out of the ground. All of these are used to illustrate resurrection from the grave.

In the name of God, the Lord of Mercy, the Giver of Mercy

[1]By the sky and the night-comer—[2]What will explain to you what the night-comer is? [3]The piercing star—[4]there is a watcher over every soul.

[5]Man should reflect on what he was created from. [6]He is created from spurting fluid, [7]then he[a] emerges from between the backbone and breastbone:[b] [8]God is certainly able to bring him back to life. [9]On the Day when secrets are laid bare [10]he will have no power and no one to help him.

[11]By the sky and its recurring rain, [12]by the earth that cracks open![c] [13]This is truly a decisive statement; [14]it is not something to be taken lightly. [15]They plot and scheme, [16]but so do I: [17][Prophet], let the disbelievers be, let them be for a while.

[a] The pronoun here is taken to refer to the person rather than the fluid.

[b] Of the mother, where she carries the baby. He emerges from the womb as he will emerge from the grave.

[c] For plants to come out, and humans on the Day of Resurrection.

Q1. Overall, how would you rate the comprehension ease of Chapter 2 on the following criteria?

	1	2	3	4	5	
Comprehensible title	☐	☐	☐	☐	☐	Incomprehensible title
Helpful introduction for text comprehension	☐	☐	☐	☐	☐	Unhelpful introduction for text comprehension
Comprehensible Verse 1	☐	☐	☐	☐	☐	Incomprehensible Verse 1
Comprehensible Verse 2	☐	☐	☐	☐	☐	Incomprehensible Verse 2
Comprehensible Verse 3	☐	☐	☐	☐	☐	Incomprehensible Verse 3
Comprehensible Verse 4	☐	☐	☐	☐	☐	Incomprehensible Verse 4
Comprehensible Verse 5	☐	☐	☐	☐	☐	Incomprehensible Verse 5
Comprehensible Verse 6	☐	☐	☐	☐	☐	Incomprehensible Verse 6
Comprehensible Verse 7	☐	☐	☐	☐	☐	Incomprehensible Verse 7
Comprehensible Verse 8	☐	☐	☐	☐	☐	Incomprehensible Verse 8
Comprehensible Verse 9	☐	☐	☐	☐	☐	Incomprehensible Verse 9
Comprehensible Verse 10	☐	☐	☐	☐	☐	Incomprehensible Verse 10
Comprehensible Verse 11	☐	☐	☐	☐	☐	Incomprehensible Verse 11
Comprehensible Verse 12	☐	☐	☐	☐	☐	Incomprehensible Verse 12
Comprehensible Verse 13	☐	☐	☐	☐	☐	Incomprehensible Verse 13
Comprehensible Verse 14	☐	☐	☐	☐	☐	Incomprehensible Verse 14
Comprehensible Verse 15	☐	☐	☐	☐	☐	Incomprehensible Verse 15
Comprehensible Verse 16	☐	☐	☐	☐	☐	Incomprehensible Verse 16
Comprehensible Verse 17	☐	☐	☐	☐	☐	Incomprehensible Verse 17
Helpful footnotes for text comprehension	☐	☐	☐	☐	☐	Unhelpful footnotes for text comprehension

Q2. Overall, how would you rate the page layout of Chapter 2 on the following criteria?

	1	2	3	4	5	
Recognizable title	☐	☐	☐	☐	☐	Unrecognizable title
Short version	☐	☐	☐	☐	☐	Long version
Simple text layout	☐	☐	☐	☐	☐	Complex text layout
Familiar paragraph division	☐	☐	☐	☐	☐	Strange paragraph division
Clear verse numbers	☐	☐	☐	☐	☐	Unclear verse numbers
Clear text organization	☐	☐	☐	☐	☐	Unclear text organization

Q3. Overall, how would you rate the word usages of Chapter 2 on the following criteria?

	1	2	3	4	5	
Familiar words	☐	☐	☐	☐	☐	Strange words
Modern words	☐	☐	☐	☐	☐	Archaic words
Frequent English words	☐	☐	☐	☐	☐	Infrequent English words
Clear words	☐	☐	☐	☐	☐	Confusing words
Short words	☐	☐	☐	☐	☐	Long words

Q4. Overall, how would you rate the grammatical usages of Chapter 2 on the following criteria?

	1	2	3	4	5	
Short sentences	☐	☐	☐	☐	☐	Long sentences
Familiar English grammar	☐	☐	☐	☐	☐	Strange English grammar
Simple sentences	☐	☐	☐	☐	☐	Complex sentences
Clear sentences	☐	☐	☐	☐	☐	Confusing sentences

Q5. Overall, how would you rate the cohesion usages (links between elements) of Chapter 2 on the following criteria?

	1	2	3	4	5	
Coherent words	☐	☐	☐	☐	☐	Incoherent words
Coherent sentences	☐	☐	☐	☐	☐	Incoherent sentences
Familiar conjunctions	☐	☐	☐	☐	☐	Strange conjunctions

Appendix G: Evaluating the readability of an English Qur'ān translation (Ali's Version)

Sample 1 (Chapter 91)

Read the following translation of a Qur'anic chapter completely and then move on to the next section to evaluate it. The numbers on the left are verse numbers.

Intro. to S. 91.

INTRODUCTION AND SUMMARY: SŪRAT *Ash-Shams*, 91.

This is one of the early Makkan revelations. Beginning with a fine nature passage, and leading up to man's need of realising his spiritual responsibility, it ends with a warning of the terrible consequences for those who fear not the Hereafter.

- 1959 -

Ash-Shams, or The Sun.

In the name of Allah, Most Gracious, Most Merciful.

1. By the Sun[6147]
   And his (glorious) splendour;
2. By the Moon[6148]
   As she follows him;
3. By the Day as it[6149]
   Shows up (the Sun's) glory;
4. By the Night as it
   Conceals it;
5. By the Firmament[6150]

سُورَةُ الشَّمْسِ

بِسْمِ اللَّهِ الرَّحْمَٰنِ الرَّحِيمِ

وَالشَّمْسِ وَضُحَاهَا ﴿١﴾

وَالْقَمَرِ إِذَا تَلَاهَا ﴿٢﴾

وَالنَّهَارِ إِذَا جَلَّاهَا ﴿٣﴾

وَاللَّيْلِ إِذَا يَغْشَاهَا ﴿٤﴾

وَالسَّمَاءِ وَمَا بَنَاهَا ﴿٥﴾

6147. Six types are taken in three pairs, from Allah's mighty works in nature, as tokens or evidence of Allah's providence and the contrasts in His sublime creation, which yet conduce to cosmic harmony (verses 1-6). Then (verses 7-8) the soul of man, with internal order and proportion in its capacities and faculties, as made by Allah, is appealed to as having been endowed with the power of discriminating between right and wrong. Then the conclusion is stated in verses 9-10, that man's success or failure, prosperity or bankruptcy, would depend upon his keeping that soul pure or his corrupting it.

6148. The first pair is the glorious sun, the source of our light and physical life, and the moon which follows or acts as second to the sun for illuminating our world. The moon, when she is in the sky with the sun, is pale and inconspicuous; in the sun's absence she shines with reflected light and may metaphorically be called the sun's vicegerent. So with Revelation and the great Prophets who brought it; and the minor Teachers who derive their light reflected, or perhaps doubly reflected, from the original source.

6149. The next contrasted pair consists, not of luminaries, but conditions, or periods of time, Day and Night. The Day reveals the sun's glory and the Night conceals it from our sight. So there may be contrasts in our subjective reception of divine light, but it is there, working all the time, and must reappear in its own good time.

6150. The next contrasted pair is the wonderful firmament on high, and the earth below our feet, stretching away to our wide horizons. The sky gives us rain, and the earth gives us food. Yet both work together: for the rain is moisture sucked up from the earth, and the food cannot grow without the heat and warmth of the sun. There are many other contrasts under this head; yet they all point to unity.

And its (wonderful) structure;[6151]

6. By the Earth
And its (wide) expanse:

7. By the Soul,
And the proportion and order
Given to it;[6152]

8. And its inspiration
As to its wrong
And its right;–

9. Truly he succeeds
That purifies it,

10. And he fails
That corrupts it![6153]

11. The Thamūd (people)
Rejected (their prophet)
Through their inordinate
Wrong-doing.[6154]

وَٱلْأَرْضِ وَمَا طَحَىٰهَا ﴿٦﴾

وَنَفْسٍ وَمَا سَوَّىٰهَا ﴿٧﴾

فَأَلْهَمَهَا فُجُورَهَا وَتَقْوَىٰهَا ﴿٨﴾

قَدْ أَفْلَحَ مَن زَكَّىٰهَا ﴿٩﴾

وَقَدْ خَابَ مَن دَسَّىٰهَا ﴿١٠﴾

كَذَّبَتْ ثَمُودُ بِطَغْوَىٰهَآ ﴿١١﴾

6151. The *mâ maṣdariya* in Arabic, in this and the subsequent clauses, is best translated in English by nouns. Thus what would literally be "and the (wonderful) making or construction of it" or "the fact of its (wonderful) construction" is, idiomatically, "its (wonderful) structure." "The (wide) spreading out" of the earth is rendered "its (wide) expanse," and so on.

6152. Allah makes the soul, and gives it order, proportion, and relative perfection, in order to adapt it for the particular circumstances in which it has to live its life. *Cf.* xxxii. 9. See also n. 120 to ii. 117. He breathes into it an understanding of what is sin, impiety, wrong-doing and what is piety and right conduct, in the special circumstances in which it may be placed. This is the most precious gift of all to man, the faculty of distinguishing between right and wrong. After the six external evidences mentioned in verses 1-6 above, this internal evidence of Allah's goodness is mentioned as the greatest of all. By these various tokens man should learn that his success, his prosperity, his salvation depends on himself,–on his keeping his soul pure as Allah made it; and his failure, his decline, his perdition depends on his soiling his soul by choosing evil.

6153. This is the core of the Sūra, and it is illustrated by a reference to the story of the Thamūd in the following verses.

6154. The allusion to the story of the Thamūd will be understood by a reference to vii. 73-79; see specially n. 1044. Their prophet was Ṣāliḥ, but he had to deal with an arrogant people, who oppressed the poor and denied them their rights of watering and pasture for their cattle.

إِذِ ٱنۢبَعَثَ أَشۡقَىٰهَا ﴿١٢﴾

12. Behold, the most wicked
Man among them was
Deputed (for impiety).[6155]

فَقَالَ لَهُمۡ رَسُولُ ٱللَّهِ نَاقَةَ ٱللَّهِ وَسُقۡيَٰهَا ﴿١٣﴾

13. But the messenger of Allah[6156]
Said to them: "It is
A She-camel of Allah!
And (bar her not
From) having her drink!"

فَكَذَّبُوهُ فَعَقَرُوهَا فَدَمۡدَمَ
عَلَيۡهِمۡ رَبُّهُم بِذَنۢبِهِمۡ فَسَوَّىٰهَا ﴿١٤﴾

14. Then they rejected him
(As a false prophet),
And they hamstrung her[6157]
So their Lord, crushed them
For their sin and
Levelled them.

وَلَا يَخَافُ عُقۡبَٰهَا ﴿١٥﴾

15. And for Him[6158]
Is no fear
Of its consequences.

6155. The prophet Ṣāliḥ was given a certain she-camel as a Sign, a test case, "This she-camel of Allah is Sign unto you: so leave her to graze in Allah's earth and let her come to no harm, or ye shall be seized with a grievous punishment" (vii. 73). But they plotted to kill her and sent the most wicked man among them to dare and do that deed of impiety. It was probably when she came to drink at the stream that she was hamstrung and killed. See xxvi. 155, and liv. 27.

6156. That is, Ṣāliḥ: see last note.

6157. The man who was deputed to do the impious deed of hamstringing the she-camel had of course the sympathy and co-operation of the whole people. Only he was more daring than the rest.

6158. This verse has been variously construed. I follow the general opinion in referring the pronoun "Him" to "their Lord" in the last verse and the pronoun "its" to the Punishment that was meted out to all, high and low, equally. In that case the meaning would be: God decreed the total destruction of the Thamūd; in the case of creatures any such destruction might cause a loss to them, and they might fear the consequences of such loss or destruction, but Allah has created and can create at will, and there can be no question of any such apprehension in His case. An alternative view is that "him" refers to the prophet Ṣāliḥ, mentioned in verse 13. Then the interpretation would be: Ṣāliḥ had no fear of the consequences for himself; he had warned the wicked according to his commission; he was saved by Allah's mercy as a just and righteous man, and he left them with regrets (vii. 79). Yet another alternative refers "him" to the wicked man (mentioned in verse 12) who hamstrung the she-camel: he feared not the consequences of his deed.

Q1. Overall, how would you rate the comprehension ease of Chapter 1 on the following criteria?

	1	2	3	4	5	
Comprehensible title	☐	☐	☐	☐	☐	Incomprehensible title
Helpful introduction for text comprehension	☐	☐	☐	☐	☐	Unhelpful introduction for text comprehension
Comprehensible Verse 1	☐	☐	☐	☐	☐	Incomprehensible Verse 1
Comprehensible Verse 2	☐	☐	☐	☐	☐	Incomprehensible Verse 2
Comprehensible Verse 3	☐	☐	☐	☐	☐	Incomprehensible Verse 3
Comprehensible Verse 4	☐	☐	☐	☐	☐	Incomprehensible Verse 4
Comprehensible Verse 5	☐	☐	☐	☐	☐	Incomprehensible Verse 5
Comprehensible Verse 6	☐	☐	☐	☐	☐	Incomprehensible Verse 6
Comprehensible Verse 7	☐	☐	☐	☐	☐	Incomprehensible Verse 7
Comprehensible Verse 8	☐	☐	☐	☐	☐	Incomprehensible Verse 8
Comprehensible Verse 9	☐	☐	☐	☐	☐	Incomprehensible Verse 9
Comprehensible Verse 10	☐	☐	☐	☐	☐	Incomprehensible Verse 10
Comprehensible Verse 11	☐	☐	☐	☐	☐	Incomprehensible Verse 11
Comprehensible Verse 12	☐	☐	☐	☐	☐	Incomprehensible Verse 12
Comprehensible Verse 13	☐	☐	☐	☐	☐	Incomprehensible Verse 13
Comprehensible Verse 14	☐	☐	☐	☐	☐	Incomprehensible Verse 14
Comprehensible Verse 15	☐	☐	☐	☐	☐	Incomprehensible Verse 15
Helpful footnotes for text comprehension	☐	☐	☐	☐	☐	Unhelpful footnotes for text comprehension

Q2. Overall, how would you rate the page layout of Chapter 1 on the following criteria?

	1	2	3	4	5	
Recognizable title	☐	☐	☐	☐	☐	Unrecognizable title
Short version	☐	☐	☐	☐	☐	Long version
Simple text layout	☐	☐	☐	☐	☐	Complex text layout
Familiar paragraph division	☐	☐	☐	☐	☐	Strange paragraph division
Clear verse numbers	☐	☐	☐	☐	☐	Unclear verse numbers
Clear text organization	☐	☐	☐	☐	☐	Unclear text organization

Q3. Overall, how would you rate the word usages of Chapter 1 on the following criteria?

	1	2	3	4	5	
Familiar words	☐	☐	☐	☐	☐	Strange words
Modern words	☐	☐	☐	☐	☐	Archaic words
Frequent English words	☐	☐	☐	☐	☐	Infrequent English words
Clear words	☐	☐	☐	☐	☐	Confusing words
Short words	☐	☐	☐	☐	☐	Long words

Q4. Overall, how would you rate the grammatical usages of Chapter 1 on the following criteria?

	1	2	3	4	5	
Short sentences	☐	☐	☐	☐	☐	Long sentences
Familiar English grammar	☐	☐	☐	☐	☐	Strange English grammar
Simple sentences	☐	☐	☐	☐	☐	Complex sentences
Clear sentences	☐	☐	☐	☐	☐	Confusing sentences

Q5. Overall, how would you rate the cohesion usages (links between elements) of Chapter 1 on the following criteria?

	1	2	3	4	5	
Coherent words	☐	☐	☐	☐	☐	Incoherent words
Coherent sentences	☐	☐	☐	☐	☐	Incoherent sentences
Familiar conjunctions	☐	☐	☐	☐	☐	Strange conjunctions

Sample (Chapter 86)

Read the following translation of a Qur'anic chapter completely and then move on to the next section to evaluate it. The numbers on the left are verse numbers.

Intro. to S. 86.

INTRODUCTION AND SUMMARY: SŪRAT *Aṭ-Ṭāriq*, 86.

This Sūra also belongs to the early Makkan period, perhaps not far removed from the last Sūra.

Its subject-matter is the protection afforded to every soul. The physical nature of man may be insignificant, but the soul given to him by Allah must win a glorious Future in the end.

- 1933 -

Aṭ-Ṭāriq, or The Night-Visitant.

In the name of Allah, Most Gracious, Most Merciful.

سُورَةُ الطَّارِقِ

بِسْمِ اللَّهِ الرَّحْمَٰنِ الرَّحِيمِ

1. By the Sky[6067]
And the Night-Visitant[6068]
(Therein);

وَالسَّمَاءِ وَالطَّارِقِ ﴿١﴾

2. And what will explain to thee
What the Night-Visitant is?–

وَمَا أَدْرَاكَ مَا الطَّارِقُ ﴿٢﴾

3. (It is) the Star
Of piercing brightness;–

النَّجْمُ الثَّاقِبُ ﴿٣﴾

4. There is no soul but has
A protector over it.[6069]

إِن كُلُّ نَفْسٍ لَّمَّا عَلَيْهَا حَافِظٌ ﴿٤﴾

5. Now let man but think
From what he is created!

فَلْيَنظُرِ الْإِنسَانُ مِمَّ خُلِقَ ﴿٥﴾

6. He is created from
A drop emitted–[6070]

خُلِقَ مِن مَّاءٍ دَافِقٍ ﴿٦﴾

6067. The appeal here is to single Sign, *viz.*: the Sky with its Night. Visitant; and the substantive proposition is in verse 4: "There is no soul but has a protector over it". In the last Sura we considered the persecution of Allah's votaries, and how Allah protects them. Here the same theme is presented in another aspect. In the darkest sky shines out most brilliantly the light of the most brilliant star. So in the night of spiritual darkness–whether through ignorance or distress–shines the glorious star of Allah's revelation. By the same token the man of Faith and Truth has nothing to fear. Allah will protect His own.

6068. This is explained in verse 3 below. The "Star of piercing brightness" is understood by some to be the Morning Star, by others to be the planet Saturn, by others again to be Sirius, or the Pleiads or shooting stars. I think it is best to take the "Star" in the collective or generic sense, for stars shine on every night in the year, and their piercing brightness is most noticeable on the darkest night.

6069. If man has a true spiritual understanding, he has nothing to be afraid of. He is protected by Allah in many ways that he does not even know. He may be an insignificant creature as a mere animal, but his soul raises him to a dignity above other creation. And all sorts of divine forces guard and protect him.

6070. See n. 5832. to lxxvi. 2. See also last note.

7. Proceeding from between The backbone and the ribs:[6071]	يَخْرُجُ مِنۢ بَيْنِ ٱلصُّلْبِ وَٱلتَّرَآئِبِ ﴿٧﴾
8. Surely (Allah) is able To bring him back (To life)![6072]	إِنَّهُۥ عَلَىٰ رَجْعِهِۦ لَقَادِرٌ ﴿٨﴾
9. The Day that (All) things secret Will be tested,	يَوْمَ تُبْلَى ٱلسَّرَآئِرُ ﴿٩﴾
10. (Man) will have No power, And no helper.[6073]	فَمَا لَهُۥ مِن قُوَّةٍ وَلَا نَاصِرٍ ﴿١٠﴾
11. By the Firmament[6074] Which giveth the returning rain,	وَٱلسَّمَآءِ ذَاتِ ٱلرَّجْعِ ﴿١١﴾
12. And by the Earth Which opens out[6075] (For the gushing of springs Or the sprouting of vegetation),–	وَٱلْأَرْضِ ذَاتِ ٱلصَّدْعِ ﴿١٢﴾

6071. A man's seed is the quintessence of his body. It proceeds from his loins, *i.e.*, from his back between the hip-bones and his ribs. His back-bone is the source and symbol of his strength and personality. In the spinal cord and in the brain is the directive energy of the central nervous system, and this directs all action, organic and psychic. The spinal cord is continuous with the Medulla Oblongata in the brain.

6072. The Creator who can mingle the forces of psychic and physical muscular action in the creation of man, as explained in the last note, can surely give a new life after physical death here, and restore man's personality in the new world that will open out in the Hereafter.

6073. In that new world, all our actions, motives, thoughts, and imaginings of this life, however secret, will be brought into the open, and tested by the standards of absolute Truth, and not by false standards of custom, prejudice, or partiality. In that severe test, any adventitious advantages of this life will have no strength or force whatever, and cannot help in any way.

6074. The Firmament above is always the same, and yet it performs its diurnal round, smoothly and punctually. So does Allah's Revelation show forth the Truth, which like a circle is ever true to its centre,–which is ever the same, though it revolves through the changing circumstances of our present life.

6075. The earth seems hard, but springs can gush forth and vegetables sprout through it and make it green and soft. So is Truth: hard perhaps to mortals, but through the fertilising agency of Revelation, it allows our inner personality to sprout and blossom forth.

13. Behold this is the Word
That distinguishes (Good
From Evil):[6076]

14. It is not a thing
For amusement.

15. As for them,[6077] they
Are but plotting a scheme,

16. And I am planning
A scheme.[6078]

17. Therefore grant a delay
To the unbelievers:
Give respite to them
Gently (for awhile).[6079]

إِنَّهُۥ لَقَوْلٌ فَصْلٌ ﴿١٣﴾

وَمَا هُوَ بِٱلْهَزْلِ ﴿١٤﴾

إِنَّهُمْ يَكِيدُونَ كَيْدًا ﴿١٥﴾

وَأَكِيدُ كَيْدًا ﴿١٦﴾

فَمَهِّلِ ٱلْكَٰفِرِينَ أَمْهِلْهُمْ رُوَيْدًا ﴿١٧﴾

6076. See the last two notes. Revelation–Allah's Truth–can pierce through the hardest crusts, and ever lead us back to the centre and goal of our life: for it separates Good from Evil definitely. It is not mere play or amusement, any more than the Sky or the Earth is. It helps us in the highest issues of our life.

6077. Though Allah in His Mercy has provided a piercing light to penetrate our spiritual darkness, and made our beings responsive to the growth of spiritual understanding, just as the hard earth is responsive to the sprouting of a seed or the gushing of a stream, yet there are evil, unregenerate men who plot and scheme against the beneficent purpose of Allah. But their plots will be of no avail, and Allah's Purpose will prevail. It happened so with the Quraish who wanted to thwart the growth of Islam. It will be so in all ages.

6078. *Cf.* iii. 54.

6079. Gentle forbearance with Evil shows our trust in Allah and Allah's Plan: for it can never be frustrated. This does not mean that we should assist or compromise with evil, or fail to put it down where we have the power. It means patience and humility where we have no visible power to prevent Evil.

Q1. Overall, how would you rate the comprehension ease of Chapter 2 on the following criteria?

	1	2	3	4	5	
Comprehensible title	☐	☐	☐	☐	☐	Incomprehensible title
Helpful introduction for text comprehension	☐	☐	☐	☐	☐	Unhelpful introduction for text comprehension
Comprehensible Verse 1	☐	☐	☐	☐	☐	Incomprehensible Verse 1
Comprehensible Verse 2	☐	☐	☐	☐	☐	Incomprehensible Verse 2
Comprehensible Verse 3	☐	☐	☐	☐	☐	Incomprehensible Verse 3
Comprehensible Verse 4	☐	☐	☐	☐	☐	Incomprehensible Verse 4
Comprehensible Verse 5	☐	☐	☐	☐	☐	Incomprehensible Verse 5
Comprehensible Verse 6	☐	☐	☐	☐	☐	Incomprehensible Verse 6
Comprehensible Verse 7	☐	☐	☐	☐	☐	Incomprehensible Verse 7
Comprehensible Verse 8	☐	☐	☐	☐	☐	Incomprehensible Verse 8
Comprehensible Verse 9	☐	☐	☐	☐	☐	Incomprehensible Verse 9
Comprehensible Verse 10	☐	☐	☐	☐	☐	Incomprehensible Verse 10
Comprehensible Verse 11	☐	☐	☐	☐	☐	Incomprehensible Verse 11
Comprehensible Verse 12	☐	☐	☐	☐	☐	Incomprehensible Verse 12
Comprehensible Verse 13	☐	☐	☐	☐	☐	Incomprehensible Verse 13
Comprehensible Verse 14	☐	☐	☐	☐	☐	Incomprehensible Verse 14
Comprehensible Verse 15	☐	☐	☐	☐	☐	Incomprehensible Verse 15
Comprehensible Verse 16	☐	☐	☐	☐	☐	Incomprehensible Verse 16
Comprehensible Verse 17	☐	☐	☐	☐	☐	Incomprehensible Verse 17
Helpful footnotes for text comprehension	☐	☐	☐	☐	☐	Unhelpful footnotes for text comprehension

Q2. Overall, how would you rate the page layout of Chapter 2 on the following criteria?

	1	2	3	4	5	
Recognizable title	☐	☐	☐	☐	☐	Unrecognizable title
Short version	☐	☐	☐	☐	☐	Long version
Simple text layout	☐	☐	☐	☐	☐	Complex text layout
Familiar paragraph division	☐	☐	☐	☐	☐	Strange paragraph division
Clear verse numbers	☐	☐	☐	☐	☐	Unclear verse numbers
Clear text organization	☐	☐	☐	☐	☐	Unclear text organization

Q3. Overall, how would you rate the word usages of Chapter 2 on the following criteria?

	1	2	3	4	5	
Familiar words	☐	☐	☐	☐	☐	Strange words
Modern words	☐	☐	☐	☐	☐	Archaic words
Frequent English words	☐	☐	☐	☐	☐	Infrequent English words
Clear words	☐	☐	☐	☐	☐	Confusing words
Short words	☐	☐	☐	☐	☐	Long words

Q4. Overall, how would you rate the grammatical usages of Chapter 2 on the following criteria?

	1	2	3	4	5	
Short sentences	☐	☐	☐	☐	☐	Long sentences
Familiar English grammar	☐	☐	☐	☐	☐	Strange English grammar
Simple sentences	☐	☐	☐	☐	☐	Complex sentences
Clear sentences	☐	☐	☐	☐	☐	Confusing sentences

Q5. Overall, how would you rate the cohesion usages (links between elements) of Chapter 2 on the following criteria?

	1	2	3	4	5	
Coherent words	☐	☐	☐	☐	☐	Incoherent words
Coherent sentences	☐	☐	☐	☐	☐	Incoherent sentences
Familiar conjunctions	☐	☐	☐	☐	☐	Strange conjunctions