

UNIVERSITY OF CAPE COAST

READABILITY AND ACADEMIC COMMUNICATION: THE CASE OF
THE HUMANITIES RESEARCH ARTICLES IN UNIVERSITY OF CAPE
COAST

WILLIAM KODOM GYASI

2017

UNIVERSITY OF CAPE COAST

READABILITY AND ACADEMIC COMMUNICATION: THE CASE OF
THE HUMANITIES RESEARCH ARTICLES IN UNIVERSITY OF CAPE
COAST

BY

WILLIAM KODOM GYASI

Thesis submitted to the Department of English, the Faculty of Arts, College of
Humanities and Legal Studies, University of Cape Coast, in partial fulfilment
of the requirements for the award of Doctor of Philosophy Degree in English
Language

APRIL 2017

©University of Cape Coast
William Kodom Gyasi

DECLARATION

Candidate's Declaration

I hereby declare that this thesis is the result of my own original research and that no part of it has been presented for another degree in this university or elsewhere.

Candidate's Signature:..... Date.....

Name: William Kodom Gyasi

Supervisors' Declaration

We hereby declare that the preparation and presentation of the thesis were supervised in accordance with the guidelines on supervision of thesis laid down by the University of Cape Coast.

Principal Supervisor's Signature: Date.....

Name: Professor Lawrence K. Owusu-Ansah.

Co-Supervisor's Signature: Date.....

Name: Professor Joseph Benjamin Archibald Afful

ABSTRACT

Research articles (RA) constitute the most important means of communicating research findings in the academic community. Despite the crucial role played by RAs within the academic community, numerous studies across the world have shown that RAs are difficult to read and understand. In Ghana, there has not been any earlier works to evaluate the readability of RAs. The objective of this study was to explore the level of readability of RAs produced by lecturers in the Faculty of Arts in the University of Cape Coast. The descriptive research design was used and the stratified random sampling technique was employed to collect 100 RAs across the eight (8) departments of the Faculty of Arts. Readability scores were computed using Flesch reading ease (FRE) and Flesch Kincaid – grade level (FKGL) indexes. With the help of SPSS (version 23.0) measures of central tendencies and dispersions, frequencies and percentage distributions, Wilcoxon signed ranked test, Kruskal – Wallis H test, Mann-Whitney U test and Spearman rho and phi correlation coefficient were used to analyse the data. The results showed that a majority (63%) of the research articles were graded as ‘*difficult*’ to read; that is, above the ‘*standard*’ readability level of 60 when measured on the FRE scale. In addition, Mann-Whitney U test showed that males and female researchers write RAs with equal readability level. The study further revealed a positive relationship between readability and text comprehension. To improve the readability of RAs in the Faculty Arts, the study recommended among others, that, researchers in the faculty should cut down on the excessive use of polysyllabic words and complex grammatical structures especially in the Departments of Music and Dance, and English.

ACKNOWLEDGEMENTS

In conducting this study, I received indispensable support from some people I wish to show appreciation to. First, my profound heartfelt appreciation goes to my supervisors, Professor Lawrence Kwadwo Owusu-Ansah and Professor Joseph Benjamin Archibald Afful, for their inestimable support in the course of this study.

I am also very thankful to Dr. Martin Bosompem, Department of Agricultural Economics and Extension and Mr. Simon Kyei Gyimah, Department of Fisheries and Aquatic Sciences, for helping me grasp basic concepts in statistics so that I could conduct the study. I also thank them for checking each stage of the work to ensure that I applied these concepts appropriately.

To Mr. Edward Tsatsu Kongo and Comsos Rai Amenorvi, I say a big 'thank you' for proofreading my work. I also have to mention my brother Mr. Osei Gyasi, and my bosom friend, Mr. Marcelinus Dery, and Professor Edoh Torgah for their emotional support.

I remember the immense support of Professor Dora Francisca Edu-Buandoh, who worked as my advisor at the proposal stage of this work; I am very grateful to her.

I cannot complete this section without mentioning Mrs Ama Amenakoma Gyasi, my beloved wife for giving me all the support I needed to complete the work.

DEDICATION

I dedicate this work to my mother, Abena Akoroma, and my father, Yaw
Manu.

TABLE OF CONTENTS

	Page
DECLARATION	ii
ABSTRACT	iii
ACKNOWLEDGEMENTS	iv
DEDICATION	v
TABLE OF CONTENTS	vi
LIST OF TABLES	xi
LIST OF FIGURES	xiii
LIST OF ACRONYMS	xiv
CHAPTER ONE: INTRODUCTION	1
Background to the Study	1
Statement of the Problem	6
Research Objectives	8
Specific Objectives	8
Research Questions	9
Hypotheses	10
Significance of the Study	11
Scope of the Study	13
Organization of the Study	13
Chapter Summary	14
CHAPTER TWO: THEORETICAL AND CONCEPTUAL FRAMEWORKS OF THE STUDY	15
Introduction	15
Theoretical Frameworks	15

<i>Discourse community</i>	15
Differences between discourse community and speech community	18
Criticisms of Discourse Community	21
The dominance theory in gender and language	27
Sex difference in language development	28
Mean length of utterance (MLU)	31
Syntactic complexity	31
Conceptual Frameworks	32
The Concept of Communication	32
Academic writing	35
Definition and functions	35
Features of academic writing	38
Tentative writing	38
Theses and Dissertations	39
Research articles	43
Conference paper	47
Critical review	48
Readability and readability formulas	49
Classical readability indexes	52
Readability formulas and cognitive theory	58
Formulas based on findings in statistical language modelling	61
Comparison of some popular readability indexes	63
Readability and text formality	65
CHAPTER THREE: REVIEW OF EXTANT LITERATURE ON READABILITY STUDIES AND ACADEMIC WRITING	69

Introduction	69
Readability of RAs	69
Readability of pedagogical materials	89
Readability of health materials	90
Readability in the financial sector	97
Readability and insurance policies	103
Readability in information technology	107
Readability – the Ghanaian context	111
Readability and text comprehension for L2 speakers of English language	114
CHAPTER FOUR: RESEARCH METHODOLOGY	118
Introduction	118
Research Design	119
Data Collection Instruments (Data Sources)	121
Validity and Reliability of Comprehension Tests items	122
Data Collection Procedures	124
Data Analysis	126
Ethical Considerations	130
Chapter Summary	131
CHAPTER FIVE: READABILITY OF RESEARCH ARTICLES IN THE HUMANITIES	132
Introduction	132
Readability of RAs in the Humanities	132
Syntactic features of research articles in the Arts across disciplines	134
Differences in readability of research articles in the Arts and standard readability scores	141

CHAPTER SIX: RELATIONSHIP BETWEEN READABILITY OF RESEARCH ARTICLES, DISCIPLINE AND GENDER	144
Introduction	144
Differences in the level of Readability of Research Articles across Disciplines/ Departments in the Arts	144
Differences in Readability Scores between Male and Female Authored Articles in the Arts	152
Gender and readability of research articles in the Arts	152
Test of significant differences in the readability of RAs between male and female authors in the Arts	153
CHAPTER SEVEN: RELATIONSHIP BETWEEN READABILITY, LEXICAL DENSITY AND TEXT COMPREHENSION	156
Introduction	156
Relationship between Readability and Lexical Density of RAs in the Humanities	156
Lexical density across departments in the Arts	157
Patterns in lexical density across departments in the Arts	159
Test of relationship between readability and lexical density	161
Relationship between Readability and Text Comprehension	163
Gender differences in text comprehension	167
Test of significant differences in text comprehension between Gender	171
CHAPTER EIGHT: SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS	176
Introduction	176
Summary	176

How readable are the research articles in the Arts when measured in terms of FRE?	178
Differences in the Readability of Research Articles across Disciplines/ Departments in the Arts	179
Gender variation in readability of research articles	180
Relationship between readability and lexical density of research articles in the Arts	180
Relationship between readability and text comprehension	181
Conclusions	182
Recommendations	183
Suggestions for Further Research	183
REFERENCES	185
APPENDIX A: MULTIPLE CHOICE TEST 1	200
APPENDIX B: MULTIPLE CHOICE TEST 2	202
APPENDIX C: MULTIPLE CHOICE TEST 3	205
APPENDIX D: REQUEST FOR PERMISSION FOR DATA COLLECTION	208

LIST OF TABLES

Table	Page
1: Suggested structure of students' thesis/dissertation	41
2: Flesch Readability Index	54
3: How to Test Readability	104
4: Recalibrated Flesch Index	106
5: Split-Half Coefficient showing the Reliability of the Text Items of the three (3) Comprehension Exercises	123
6: Flesch Readability score and Corresponding Interpretation	127
7: Descriptive Statistics of the Syntactic features of Research Articles in the Arts across Disciplines	135
8: Descriptive Statistics of Flesch Reading Ease for Articles from the Faculty of Arts	137
9: Mean Rank FRE Scores across Departments	145
10: K-W H test showing differences in FRE scores across Departments	146
11: Dunn's Post hoc Multiple Comparison Test of FRE Scores across Departments in the Arts	147
12: Descriptive Statistics of Readability of male and Female Authored Research Articles in the Arts Measured in terms of Flesch Reading Ease	152
13: Results of Mann - Whitney U test of FRE scores between male and Female Authored Research Articles	153
14: Descriptive statistics of lexical densities of articles from the Faculty of Arts	157
15: Correlation between other Readability indices and Lexical Density	162

16: Gender difference in Text Comprehension for Test 1	168
17: Gender difference in Text Comprehension for Test 2	169
18: Gender difference in Text Comprehension for Test 3	170
19: Chi-square Analysis of Text Comprehension between Dender	171
20: The odds of scoring above average in the MCCE between males and Females	174

LIST OF FIGURES

Figure	Pages
1: Discourse Community and conventions (Menezes, 2014)	17
2: Conceptual framework of readability and text comprehension of RAs.	51
3: Frequency distribution of readability scores of research articles in the Arts	133
4: One-sample Wilcoxon signed rank test	142
5: Lexical Density Distribution across the Arts	160
6: Bar Graph showing the pass rate of participants in MCCE 1	164
7: Bar Graph showing the pass rate of participants in MCCE 2	165
8: Bar Graph showing the pass rate of participants in MCCE 3	167

LIST OF ACRONYMS

AAC&U	Association of American Colleges and Universities
AAO	American Association of Ophthalmic Oncologists and Pathologists
AAPOS	American Association for Paediatric Ophthalmology and Strabismus
AGS	American Gem Society
ARI	Automated Readability Index
ASCRS	American Society of Cataract and Refractive Surgery
ASOPRS	American Society of Ophthalmic Plastic and Reconstructive Surgery
ATOS	Advantage – TASA Open Standard
BR	Beginning Reader
BYU	Brigham Young University
CLI	Coleman-Liau Index
CSR	Corporate Social Responsibility
ETS	Education Texting Service
FKGL	Flesch Kincaid Grade Level
FL	Foreign Language
FRE	Flesch Reading Ease
L1	Mother Tongue
IMRAD	Introduction, Methods, Results and Discussion
L2	Second Language
LD	Lexical Density
LSA	Latent Semantic Analysis

MLU	Mean Length of Utterance
NANO	North American Neuro-Ophthalmology Society
OECD	Organisation for Economic Cooperation and Development
OLS	Ordinary Least Squares Regression
ORF	Oral Reading Frequency
SMOG	Simple Measure of Gobbledygook
TASA	Touchstone Applied Sciences and Associates

CHAPTER ONE

INTRODUCTION

This chapter sets the background for the study. It discusses research articles as potent source of information in academia and advances five research objectives to determine whether this genre is written in a way that disseminates information effectively in the research community.

Background to the Study

Soodmand, Moradi, & Hamzavi (2014: 71) considers academic communication “as the most crucial and permanent way of communication, which has several forms, one of which is academic writing. Academic writing involves interrelationships between the author and their audiences based on certain conventions, which take part in an academic discourse”. In academia, “publish or perish,” is a well-known saying which is used to express the idea that it is important for teachers and researchers in colleges and universities to publish findings of their research, and that if they fail to do so it will have a negative effect on their career (Kampourakis, 2016). Hence, publications are an important component of an academician’s curriculum vitae (Priestley, 2015) and job progression. The number and quality of publications one has may contribute significantly to one’s chances of promotion, getting accepted for a new position, obtaining research grants, being invited to give presentations, or being elected to boards or other positions important to one’s area of expertise. In essence, by publishing one’s work, one presents one’s expertise to the academic community, and personal gain in terms of promoting one’s career is a significant motivation (Priestley,

2015). Thus, the art of publishing research findings is a key component in academia.

It is the duty of every academician to share new findings with the academic community, in an effort to advance the existing knowledge pool (Joubert & Rogers, 2015). A number of ways are used to communicate research findings with the academic community. Each method of communicating research findings has its approach and focus. These include books, essays, ethnographies, monographs, research articles, conference papers, technical reports, dissertations and theses. Among these, it has been indicated that Research Articles (RAs) published in journals are the most important mode of communication in the academic communities (Priestley, 2015 & Swales, 1981, 1990).

Therefore, every lecturer in the academic community considers publishing RAs as key ingredient in his/her career development. Even so, one's chances of academic progress are further enhanced, depending on the quality of the research publications to one's credit (Altman, 2015; Villiers & Dumay, 2014). In assessing the quality of research articles, literature primarily uses one of two different approaches. These are as follows:

1. using checklists covering specific aspects of the article and
2. applying broad principles that guide assessment

(Jaroongkhongdach, Watson, Keyuravong & Hall, 2012).

An example of the former is Bell's (1993) list of 22 items ranging from "Does the title indicate the nature of the study?" to "Is the referencing well done?" (p. 162). An example of the latter comes from the American Educational Research Association (2006), which identifies two overarching principles

governing RA quality: the sufficiency of the warrants, and the transparency of the report. It then suggests a list of criteria based on these principles which are sufficiently flexible to be applied to different kinds of research. The main difference between these two types is that the second assumes that there are certain generic elements that indicate the quality of RA.

The second approach has been adopted often because it uses broadly applicable and operationalizable criteria which can be applied across different research traditions (Jaroongkhongdach et al., 2012). Even so, there has not been an established list of criteria. Recently though, some authors have made considerable efforts to categorise the various themes or criteria for easy application (Jaroongkhongdach et al., 2012). In all, five criteria have been indicated as the yardstick for determining the quality of RAs: justification, clarity, coherence, appropriacy, and awareness. Justification refers to reasoning provided for decisions made in research. Since the goal of any RA is generally to convince its readers of its merits, it is important to ensure that a research is justifiable. This is particularly the case where there are competing paradigms and perspectives (Edge & Richards, 1998). Clarity refers to the sufficiency of descriptions or explanations of a term/concept/procedure, and the style of writing that makes the term/concept/procedure easy for an intelligent general reader to understand (Cottrell, 2005). Coherence refers to the logical relationships within a section or across sections in terms of contents or ideas.

Comment [P1]: Awareness of what?

The fact that coherence in articles is a publication problem mentioned by several editors and reviewers (Flowerdew, 2001) suggests that this aspect is crucial in research quality assessment. Appropriacy broadly refers to the

match/compatibility between two or more potentially related components such as the match between nature of data and the research question, or the consistency among research epistemology, methodology, and method (Carter & Little, 2007). Awareness means the thoughtful concern of alternative views or of possible impacts of research decisions. It is encouraged, though not explicitly stated, in literature, especially in terms of considering alternative views in reviewing literature, providing “limitations and cautions about the data collection procedures” (Smagorinsky, 2008, p. 395), or presenting “possible explanations for the results” (Mackey & Gass, 2005, p. 300).

Although each of these five criteria is important in the assessment of the quality of an RA, its clarity seems to me, to be the most important. This is because even if an RA has a persuasive justification, coherence, appropriacy, and awareness, all these will not mean much to a reader if the article is not clear and easy for an intelligent general reader to understand (Cottrell, 2005). Therefore, a major component of a quality RA is how readable the paper is. The readability of RA, is a measure of the RA’s clarity. Readability refers to the ease of understanding or comprehension of a text due to the style of writing (Dubay, 2004). Several factors affect the readability of a written text. These include the style of writing, the percentage of difficult words contained in the text, and the length of the sentences. In addition, the target audience informs the readability level a written material should target (Cutts, 2013).

The readability of a research paper is an important criterion for measuring the quality of the paper because incomprehensible texts hamper the transfer of understanding between readers and authors. For this reason, readability should be every researcher’s primary goal. No matter the

credibility, prestige, or popularity of a journal, book, magazine, etc. in which one's RAs are published, the basic function to transmit information will be hampered if the article is not readable.

The challenge to write readable text is especially pronounced for L2 users of the English language (Rummel, 2005; Sattari, 2012; Wennerstrom, 2003). Non-native English-speaking academics face enormous difficulties if they want to be successful in the discourse community through writing RAs (Paltridge, 1993), using the English language. The challenge in writing quality RA in English is compounded by issues of genre (traditionally defined as 'types of literary productions') and style (Rakedzon & Baram-tsabari, 2016). Indeed, a number of authors have indicated that foreign language learners tend to compose words and sentences in their L1 and then translate them into the L2 (Hussein & Mohammad, 2012; Khuwaileh, 1995b). However, learning to write specifically academic genres is imperative for beginning and middle level academics who want to be fully integrated into their academic community. Hence, at present, the acquisition and proficiency in foreign languages occupy an important position worldwide (Brookshire, 2012; Rummel, 2005; Shin & Bruno, 2003).

Conrad and Biber (2001) mention that English as a language enjoys a unique position worldwide and is preferred by academicians, such as lecturers, in the composition of their RAs. Underscoring an important reason for this preference, Mydan (2007, p. 9) indicates that riding the crest of globalization and technology, English dominates the world as no language ever has, and some linguists are now saying it may never be dethroned as the king of languages. Warschauer (cited in Myanmar Times, 2007, p. 5) further states; it

is the common language in almost every endeavour, from science to air traffic control, the emblem of a globalized workplace and has become the second language of everybody.

As the lingua franca for inter-ethnic communication and the main medium of instruction in educational institutions in Ghana, English provides a veritable platform for academic development, especially in the universities. The relevance of English for academic communication in the universities can also be attributed to its usage as the language for socio-economic mobility by educated Ghanaians, who grew into an elite group whose influence rested on their command of English (Afful, 2006; Albakry & Ofori, 2011; Sackey, 1997; Sey, 1973). Considering the enormous benefits that a researcher as well as members of the academic community derives from research articles, it is important that Ghanaian researchers in institutions of higher learning produce research articles that have high readability. To this end, the present study examines the readability of research articles in the humanities published by lecturers in the University of Cape Coast.

Comment [P2]: Fix it in the references

Statement of the Problem

In their studies, Swales (1990), Murray, Rowena & Moore (2006) and Brian (2010) indicate that RAs are a potent means of disseminating updated and relevant knowledge to the academic community and a useful platform for ensuring academic development. Priestley (2015) has also indicated that RAs are the most important means of communicating research results in the academic community. Consequently, it can be concluded that RAs play a major role in dispensing knowledge in every academic institution and that

members of the research community rely heavily on RAs for knowledge acquisition and dissemination.

Despite the critical role played by RAs within the academic community, it has been established in numerous studies that RAs are difficult to read and understand. For example, Dolnicar and Chappel (2014) found that all fourteen RAs evaluated for their readability proved to be difficult to read. Their work corroborated earlier findings of other researchers (e.g. Garcí'a-Merino & Santos-A'Lvarez, 2009; Gazni, 2011; Gizir & Simsek, 2005; Hall, 2005). Several other authors have reported similar findings (e.g. Hartley, Pennebaker, & Fox, 2003; Hayden, 2008; Lee & French, 2011; Shelley & Schuh, 2001). In addition, the readability of some RAs has become poorer over time (Bauerly, Johnson, & Singh, 2006). Dolnicar and Chappel, 2014; Gazni (2011) also add that the poor nature of RAs is exhibited in most RAs regardless of the geographic backgrounds of authors. If this was the case, then a similar situation is likely to be the case for RAs authored by Ghanaian scholars in all disciplines including RAs in the Arts considered to be good at writing and communicating in the academic community.

In his book, *Ghanaian English: An Exploratory Survey*, Sey (1973) articulated that the Educated Ghanaian (EG) does not only use learned and archaic forms but is also guilty of flamboyance of prose style and frequent cases of hyper-correctness. To Sey, it is not surprising that writers who are known for exhibiting this rhetorical disposition are viewed as being linguistically competent. Recent studies by some Ghanaian scholars have articulated a similar concern about the nature of the language of educated Ghanaians. For example, Mahama (2012) asserts that there are differences in

the use of English between highly educated Ghanaians, the averagely educated Ghanaian and those who have no formal education. Thus, the possibility that RAs written by researchers in Ghanaian universities are written in a way that will pose readability problems exist. Also, earlier works by Ghanaian authors (Gyasi, 2011 and 2013; Owu-Ewie, 2014) conclude that textbooks authored by Ghanaian scholars are very difficult to read. This gives further credence to the view that RAs by Ghanaian authors are likely to be difficult to read.

Yet, no research, as of now, has been published on the readability of RAs authored by Ghanaian researchers to the best of my knowledge, although this is the most important mode of communicating ideas in the academic community. This study, therefore, sought to fill the information gap with respect to the readability of RAs authored by Ghanaian university lecturers in the Faculty of Arts and to determine the readability levels of these RAs.

Research Objectives

General objective:

The general objective of the present research was to explore the level readability of RAs written by lecturers in some Humanities Disciplines in University of Cape Coast, Ghana.

Specific Objectives

The following specific objectives were set in order to achieve the general objective of the study:

1. To determine the level of readability of RAs written by Faculty in the Arts in the University Cape Coast.

2. To determine the differences in the level of readability of RAs written by faculty across disciplines/ Departments in the the University of Cape Coast.
3. To determine the differences in the level of readability of female and male authored RAs in some selected departments in the Humanities.
4. To explore the relationship between readability and lexical density of research articles in the Humanities.
5. To explore the relationship between readability and text Comprehension among faculties.

Research Questions

The following research questions were derived from the specific objectives to guide the study:

1. How readable are RAs written by faculty in the University of Cape Coast when compared to standard readability scores?
2. What are the differences in the level of readability of RAs written by faculty among disciplines/Departments across the Arts?
3. Are there differences between female-authored and male-authored RAs in the Arts at Univesity of Cape Coast?
4. What is the relationship between readability and lexical density of RAs in the Arts?
5. What is the relationship between readability and text comprehension?

Hypotheses

The following four (4) hypotheses were formulated to guide the study:

Hypothesis 1:

H₀: There are no significant differences in the level of readability of RAs in the Arts and recommended readability level (FRE = 60)

H₁: There are significant differences in the level of readability of research articles in the Arts and recommended readability level (FRE = 60)

Hypothesis 2:

H₀: There are no significant differences in the level of readability of research articles across the disciplines/Departments in the Arts

H₁: There are significant differences between the levels of readability of research articles across the disciplines/Departments in the Arts

Hypothesis 3:

H₀: There is no significant difference between the level of readability of female and male authored research articles in the Arts.

H₁: There is a significant difference between the readability of female and male authored research articles in the Arts.

Hypothesis 4:

H₀: There is no significant relationship between the level of readability and lexical density.

H₁: There is significant relationship between the level of readability and lexical density.

Significance of the Study

The relevance of this study is rooted in the fact that English Language is a national language and that it is used for education in Ghana. Researchers in Ghanaian universities therefore use solely English language in the production of their RAs. Therefore, the way the language is used in composing their research articles is very important since the RAs are the major means of communicating research findings in the academic. Thus, the study sets a benchmark for the production of RAs for the academic community to be done with readability in view. In other words, the present work serves the purpose of quality control in the writing of RAs. Thus, the findings of this study will constitute a reference point as far as ensuring clarity and readability, which are key features of academic communication, are concerned.

The Department of Communication Studies of the University of Cape Coast established *The Writing Unit* to offer services such as editing and proofreading to lecturers and students in the university community and others outside the university community. Thus, both students and lecturers who send their write-ups to *The Writing Unit* for editing and proofreading can also ask for readability check to be conducted to ascertain the level of readability of their documents. General consultancy services on readability issues can also be obtained from *The Writing Unit*. This will no doubt improve academic communication in the University of Cape Coast and other universities in the country can also follow suit in that respect.

This study also opens a large door of activities in the area of readability and academic communication which has not gained much attention in Ghana. Studies in readability abound in the US and Europe in general.

Research in the UK and USA mainly focused on readability of educational and examination materials, health manuals and insurance policy documents. However, in Ghana, it appears that educational materials, health manuals, and insurance policy documents among others are written without targeting a particular readership. Considering target readers means pegging the readability statistics of the texts on the educational level of the targeted readers.

RAs are genres written by both experts and novice researchers. Thus, it is paramount for researchers to bear in mind the specific target for whom the articles are written. Targetting the readership of the RA will help researchers to tone down or tone up the complexity of the RA. Apart from Fosu (2014), Gyasi (2011, 2013a, 2013b and 2013c), and Owu-Ewie (2014) who did some work on readability in Ghana, some researchers may lack knowledge about readability statistics. Therefore, the current study seeks to bring to light the various levels of complexity of most RAs.

To bring to the fore the complexities embedded in RAs in terms of readability, the study could serve as the yardstick for publishing companies as well as journals to measure the readability statistics of texts before they are published. Since readability study is not common in Ghana, it is likely that publishing companies in Ghana do not have this scientific method of measuring the complexity of the documents they publish and might result in the comprehension difficulties for readers. A readability analysis of RAs by Ghanaian academics is therefore a step in the right direction to expose the incommodious experiences of readers. The study also has relevance to editors of journals. They can increase the readability of their journals thereby

Comment [P3]: Rephrase

Comment [P4]: Not clear

increasing the readership of their journals. By ensuring that their journals have high readability, they may succeed in increasing or widening of their journals.

Scope of the Study

The University of Cape Coast is chosen for the study because it is not only a wellknown university in the country but it also has international acclaim. Its strategic location in Cape Coast, the first capital city of the Republic of Ghana, makes it conspicuous to the world. The University of Cape Coast has a high reputation for being the only tertiary institution in Ghana that produces graduate teachers not only for the country but also for the outside world. The researcher chose the Arts because as a member of that research community he wanted to start from “home” in order to ascertain what is happening in his own discourse community in terms of readability. The study is restricted to published RA’s by lecturers in the Arts of the University of Cape Coast. Both empirical and non-empirical RAs were used since both types are chief scholarly sources of information within the academic community.

Comment [P5]: source

Comment [P6]: which RAs (empirical and non empirical) were your focus?

Organization of the Study

The study is organised in eight chapters. Chapter one focuses on the background to the study, statement of the problem, research questions, and significance of the study. This chapter also considers the significance of the study, scope of the study as well as the organization of the study.

Chapter Two reviews the theoretical framework that is related to the present study. In chapter three, empirical studies are discussed. The fourth

chapter discusses the methodology of the research. Here, the researcher identifies the population as well as sampling technique that are used to select the sample size for the study. The chapter also considers other relevant issues like instrumentation, research site, data collection and data analysis procedure. The subsequent three chapters concern data analysis. While chapter five considers the readability of RAs in the Humanities, chapter six compares the readability of RAs across the eight Departments in the Faculty of Arts. It also considers differences in the readability of RAs written by males and females. In chapter seven, the relationships between readability of RAs and lexical density, and text comprehension, are considered. Summary of the research findings, the implications of the study, and recommendations for further study are presented in Chapter eight.

Chapter Summary

In this chapter, I discussed the fact that RAs are potent sources of information in the academic community and raised the possibility of readability problems of RAs in the Humanities. I further advanced five research objectives as the basis for investigating the research problem and discussed the significance of the study. Finally, the limitation and scope of the study were set. The next chapter examines the theoretical and conceptual underpinnings of the study.

Comment [P7]: Shd be provided.

CHAPTER TWO

THEORETICAL AND CONCEPTUAL FRAMEWORKS OF THE STUDY

Introduction

This chapter presents the theoretical and conceptual underpinnings of the study. Here, I discuss discourse community by Swales (1990) and Dominance Theory by Tannen (1990) as the theories for the study. Further, I explore academic writing and readability as the major concepts of the study.

Theoretical Frameworks

Discourse community

Swales (1990) described Discourse Community as a group that has goals or purposes and uses communication to achieve these goals. Central to his analysis is the notion of genre, the organizational patterns of written communication which according to Borg (2003), Swales sees genre as belonging to Discourse Communities and conversely helping to define those communities (1990:9).

Mooney (2011) indicated that genres are textual tools used by members of a DC to work toward their desired ends. Genres and conventions are special languages and words the community uses for communication, and according to Swale and Feak (2004), these change over time as the community discovers several efficient adaptations, as group membership changes, or as the group's desired goals change. Illustrating the importance of genre and conventions, Mooney (2011) cited the example of a computer support specialist who was not familiar with the conventions of his DC. It

made him an outcast because he was unaware of the genres of the group. A lecturer is required to present an RA in line with the lexis, genre and convention guiding such a publication.

Illuminating Swales' definition of DC from the perspective of pedagogies associated with writing across the curriculum and academic English, Teubert (2010) posits that DC is now viewed as a 'nest of ideas'. Language use in a group is a structure of social behaviour while discourse is a means through which the group's knowledge is maintained and extended, underscoring the fact that discourse is epistemic or constitutive of the group's knowledge. Hence, as Routledge (2005) puts it, DC could comprise academics or the readers of articles in magazines. It can refer to several overlapping groups of people. It could also refer to individuals as a written text is aimed at (such as the targeted audience of a lecturer's research article). It could also refer to people who participate in a set of discourse practices both by reading and writing (such as the lecturer who writes and the academics or students who read).

Summarizing the notion of DC, Menezes (2014) stated that the term has been appropriated by the social 'perspectivists' for their variously applied purposes in writing RAs. It operates within conventions defined by communities, be they academic disciplines or social groups and that the idea of DCs is not a settled notion, rather the center of a set of ideas. As depicted in Figure 1, each of the DCs will describe or provide an in-depth exposition on Campus Crime utilizing the lexis, genres and conventions in that lexis. However, discussing the same topic, each of the communities would express what campus crime entails according to their group's special language and

terms. For example, “offenders group” may refer to it as “bullshit” “teaching faculty group”, will refer to it as “misdemeanor” and “law enforcement agents group” will refer to it as “offence.”

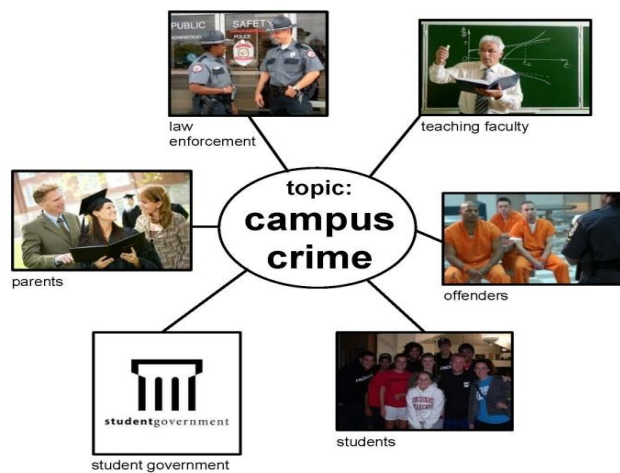


Figure 1: Discourse Community and conventions (Menezes, 2014)

A DC is a group of communicators with a common goal or interest that adopts certain preferred ways of participating in public discussion Swales (1990). These preferred ways of discussion are called discursive practices. Discursive practices in academic DC involve various genres (academic papers, books, lectures, debates, TV and radio programming, etc.) and require the mastery of certain special terminology or jargon. Generally, "membership" in a DC requires a certain level of expertise in the common goal; the more "expert" one is considered to be, the more influence one has over the preferred discursive practices. The boundaries of DCs are often hazy, and frequently overlap, and many broad DCs have smaller, more specialized sub-

communities. Most people participate regularly in several different discourse communities.

Initially, Discourse Communities were viewed in the perspective of spoken communication and studied along that line. However, scholars saw the urgent need for underscoring DC as comprising written text rather than the spoken text and studied for writers and readers (Menezes, 2014) hence the change of the term Discourse Community to 'Discourse Communities'. This feat was attained via Swales' exposition in his book, *Genre Analysis: English in Academic and Research Settings*.

Irrespective of the clarifications and explanations given about DC, scholars including Swales have argued that there is no need to clarify the concept of Discourse Community. Their reason is that any further explanations to the concept of DC will yield nothing more than a translation of the long-established concept of a Speech Community" (SC) (Swales, 1990 p. 23). However, I would like to establish the fact that DC and SC are quite distinct in a number of respects.

Differences between discourse community and speech community

A Speech Community is considered as "any human aggregate characterized by regular and frequent interaction by means of a shared body of verbal signs and set off from similar aggregates by significant differences in language usage" (Gumpers, 1982). In this sense, a person is born into a speech community, can constantly relocate to a place at one point and leave the next time. On the other hand, the membership of a person into a Discourse Community requires some qualification which is spelt out by members or

made known to people before they join often by the means of education, apprenticeship or some other form of socialisation.

Furthermore, DC and SC differ in membership and participation of members in the community. The concept on which a discourse community is formed depends on how someone can become a part of it. The member of DC is two folds. For one, the prospective members must have a certain qualification before they are registered into the community and for the other, prospective members, once they have the qualification become automatic members and begin to contribute towards a common goal. Swales believes anyone can acquire discourse and become a member of the community once they have met the requirements or cluster of ideas for that particular discourse community.

However, the concept on which speech community is formed is based on sharing knowledge of rules of conduct and interpretation of speech. Such sharing comprises knowledge of at least one form of speech, and knowledge also of its patterns of use". In this sense the sharing of knowledge in the community is open to both natives and non-natives of the community. The community accepts members on the basis that the individual being admitted into the community will comply by rules of conduct.

Swales illustrates that a Discourse Community requires a network of communication and common goals among members because there may be a considerable distance between them both ethnically and geographically, while a speech community requires physical proximity (Martin-Martin, 2005, Nordquist, 2015). In this sense, one's initial socialization into a speech

community may occur within a culture with communicative values that differ from other cultures and communities that one encounters later in life.

Again, Swales mentions that a Discourse Community is a socio-rhetorical unit where language is used in diverse contexts to achieve specific goals. However, a Speech Community is a socio-linguistic unit that shares linguistic forms, regulative rules and cultural concepts (Borg, 2003). This suggests that language use in SC represents, embodies, constructs and constitutes meaningful participation in society and culture. In addition, SC develops through prolonged interaction among those who operate within these shared and recognized beliefs and value systems regarding forms and styles of communication.

Borg (2003) claims that in DCs, the communicative needs of the goals tend to predominate in the development and maintenance of its discursal characteristics. Simplifying Borg's expression, Menezes (2014) states that members of a Discourse Community actively share goals and communicate with other members to pursue those goals which often focus on the use and analysis of written communication.

Martin-Martin (2005) says that discourse communities are centrifugal in nature; that is, they tend to separate people into occupational or specialty-interest groups while speech communities are centripetal; that is, they tend to absorb people into the general fabric of the society. According to Routledge (2005), a speech community typically inherits its membership by birth, accident or adoption while a Discourse Community recruits its members by persuasion, training or relevant qualification in order to form what we could refer to as a specific interest group.

Every discipline within a university constitutes a discourse community since members of each discipline share a common goal, namely teaching, research and community service. Using language, scholars produce and circulate RAs in their DCs to contribute to knowledge. For this goal of disseminating information through the production and publication of RAs to be achieved, such articles must be written in concise and readable for. The academic discipline as a DC determines how language should be used in the production of research articles. It is against this background that the researcher deems it appropriate to use Swales' DC as the theoretical underpinning of the present study (Swales, 1990; Swales & Feak, 2004).

Criticisms of Discourse Community

In spite of the laudable and convincing information provided by Swales on the concept of Discourse Community and its definition, critiques such as Bizzell (1992) contends that Swales' (1988) criterion for defining DC was misleading. When Swales (1988) offered the criteria for defining DCs, he illustrated the application of the criteria by considering some hypothetical café owners and hobby groups. He argued that cafe owners would not form a DC because they would not have common channels of interaction or a common project.

Bizzell (1992) adds that she had intended to make world-view and social practices, not simply language-using conventions, central. She suggested that the hypothetical cafe owners were likely to be a DC because of social and class-based or ethnic discursive practices of the people likely to

become cafe owners and because they engaged in similar discourses (ordering supplies, talking to customers and employees).

In fact, Bizzell expressed skepticism about Swales' stamp collectors, wondering if this hobby would lead to a shared worldview. She concluded that perhaps really dedicated stamp collectors might, through intense and long-term engagement in this hobby, develop "habits of mind" that would shape many areas of their lives. Or perhaps, as in the cafe owners' problem, selection issues could be relevant.

Again, Swales (1988) characterises Discourse Community as Goals, Intercommunication, Participation, Genres Lexis, and Expertise. These parameters are contended by Bizzell that the concept of DC is fuzzy, overly homogeneous, and overly consensual utopias. Her reason was that in the face of the emerging situated research on academic and disciplinary writing complex spaces sprang up with multiple discourses, practices and identities. Particularly, the existence of multiple discourses and complex spaces make the concept of DC quite cumbersome and uninteresting. This is because the concept encompasses a lot of aspects that go beyond the parameters of the theory.

Prior (2003) adds that Swales had noticed the shortcomings of the explanation to the concept of Discourse Community based on the issues his critiques raised against the theory. Initially, Swales was skeptical about the concerns raised by his critiques. Critics such as Bizzell (1992) and Prior (2003) state that DC is not necessarily defined by mutual engagement but consists of individuals who co-participate in discursive practices with some purposeful focus even when they are separated by time, language, geography, and so on.

As a result, Swales in 1998 concluded that his skepticism had been premature; therefore, and revised the framework.

Harris (1989) argues that discussions about DCs had been sweeping and vague. The reason is that the communities described were fuzzily marked by consensus and homogeneity. Similarly, Cooper (1989) notes the tendency for discourse communities to be conceived of as stable entities separate from and governing individuals and events. Drawing on Geertz (1973) argued for seeing discourse communities instead as the products of continual hermeneutic work, as social phenomena where varied values and practices intersected, as ways of being in the world, not narrow intellectual commitments.

Seeking to clarify the notion of DCs, Cooper and Harris both suggested that the concept needed to stay closer to concrete, local groupings, avoiding expansive abstractions like “the academic discourse community,” and to acknowledge disagreement and conflict. Harris saw conflicts particularly as marks of commitments to multiple communities. He submits “...one is always simultaneously a part of several discourses, several communities, is always already committed to a number of conflicting beliefs and practices” (p. 19). Situated research on the academic and disciplinary writing and the enculturation of undergraduate and graduate students (e.g., Casanave, 2002; Chin, 1994; and Chiseri-Strater, 1991) has complicated notions of discourse. Despite the fact that critiques have pinpointed some shortcomings of the theory, the theory still can serve some purposes that are of great importance to all students and faculty members because discourse communities use rigid conventions for language use, in choice of words, genre, and style. According to Kutz (1998) the theory has enlightened language users to understand that

the use of language is mainly based on context and specialized vocabulary; therefore, when writing, language users should be conversant with such conventions and employ them.

Moreover, White and Lowenthal (2011) espoused that DC use a variety of discursive styles used by different categories of people. The university, for instance, has colleges, departments and other areas of study where each maintains and perpetuates its own unique discursive style(s). As a result, linguists and educational researchers who are familiar with DCs have acknowledged that the university does represent a definitive example of a discourse community complete with specific rules for participation therein (Bizzell, 1986, 1992; Gravett and Petersen, 2007; Williams, 2005). The researcher of the present study also believes that the university is a collection of DCs and is interested in identifying the readability levels of the writing styles of faculty members.

The university culture is a unique community based around discourse (Bizzell, 1982; Gravett & Petersen, 2007) and comprises its own discourse style. Correspondingly, full participation within this academic DC requires that one learn and adopt the unique discourse structure of the university. Bizzell's work reveals that traditional four-year colleges and universities are themselves a unique culture in which participants are required both to employ certain kinds of discourse and to adapt themselves to a specific and corresponding set of values and identities unique to that setting. It is against this background that it is very essential that everyone become familiar with the discursive styles of DC he/she belong to. Prior (1998) adds the need for members of a particular discourse community to intellectually, linguistically,

and socially employ the conventions of the DC to which one belongs. The conventions govern members' spoken and written interaction because the ways of thinking and communicating of one culture may differ significantly from that of another discourse community in terms of ideology and linguistics. White and Lowenthal (2011) further support DC and indicate that for any member of a discourse community to be able to shift into any kind of discourse community, they must first know and understand the characteristics that define this discursive style. That said, DC seeks to bring to the fore various conventions that are required by various discourse communities. The work of researchers in academic literacy who have explored many of the characteristics that define the academic DC (Chiseri-Strater, 1991; Elbow, 1998; Macken-Horarik, 1996; Spellmeyer, 1998), has largely remained in the realm of research.

Elbow (1998), Gravett & Petersen (2007), Hindman (1997), Tannen, (2002) and Turner (2003) also shared in the same view with earlier researchers in favour of the theory that Discourse Communities tend to focus on a number of essential components: verbal assertiveness and voluntary participation, formality and explicitness, binary agonism, objectivity, specialized jargon, elements of display, and selectivity. Each of these components of academic discourse is unique and almost all of them are based on white, western linguistic norms (Elbow, 1998; Scheurich, 1993; Sleeter, 1993; Turner, 2003). Another opinion that deserves mentioning in connection with views in favour of DC is the concept of Cognition and Metacognition intimated by Bruner (1986 & 1990); and Gee (2002 & 2003). Cognition and Metacognition develop largely with language. Language serves as the primary scaffold for cognition;

without well- developed language skills, humans are largely incapable of developing high-order thinking (Vygotsky, 1986). Similarly, because language is requisite for cognition, it affects identity (Schwartz, Montgomery & Briones, 2006).

People come to understand themselves and their respective place in the world largely through uses of language (Vygotsky, 1986). It is not surprising then that changes in language often bring with them concurrent changes in identity (Gee, 2005); through repeated and extended interaction with communities of practice and their associated uses of language come changes in how individuals perceive themselves and their respective roles both within a discourse community and in the culture(s) outside of that community (Gee, 2000). In short, people begin to identify themselves through the various communities of discourse and practice of which they are a part. Based on this assertion, this research is meant to study the discursive practices of university lecturers by analysing their research articles in terms of how readable their articles may be.

In view of the aforementioned, it could be seen that DC is well suited for the academic context. It is very relevant in the dissemination and gathering of useful information and feedback, which in turn, can provide an upward surge for pedagogy, especially in the universities. However, Bazerman and Prior (2005) contend that DC has been shown to be imprecise and inaccurate, by emphasizing the uniformity, symmetrical relations and cooperation within text circulation networks.

Explicating the 'bone of contention', Bazerman and Prior indicated that social collectivities in communication are often contentious, by design or

accident. Individuals within these social collectivities are cast into or they adopt different roles with different discursive power, rights, obligations, and expectations. Written texts are circulated somewhat in heterogeneous groupings, as lecturers write to administrators, colleagues, students, and reputable organizations. This means, circulation of texts may form groupings that might not otherwise have any regular communicative relations prior to being brought together by the circulation of documents. Due to such occurrences and other social complexities, it is appropriate that a more subtle and varied sociological vocabulary is needed to describe the set of relations within text circulation networks.

The dominance theory in gender and language

Comment [P8]:

According to the difference theory propounded by Tannen (1990), men and women, even those within the same group (e.g. age, culture etc.), live in different or separate cultural worlds and, as a result, they promote different ways of speaking (Uchida, 1992). This theory is sometimes called “two-culture theory” (Nemati & Bayer, 2007). In simple terms, although men and women live in the same environment they establish different relations with society as if each belonged to a different environment and culture, the result of which is consequently reflected in the language of both genders as in other aspects of their lives (Nemati & Bayer, 2007). So, in this theory, cross-gender communication is to be taken as cross-cultural or bi-cultural communication. Tannen (1990) believes the difference starts in childhood, where parents use more words about feelings to girls and use more verbs about action to boys. Males and females belong to different sub-cultures and therefore speak

differently. Her book, *You Just Don't Understand* (1990), claims that there are six main differences between the ways males and females use language:

1. Status vs. support – men see language as a means of asserting dominance; women see it as a way of confirming/supporting ideas.
2. Independence vs. intimacy – men “go it alone”; women seek support.
3. Advice vs. understanding – men see language as problem solving; women see it as a means of empathy.
4. Information vs. feelings – males are concerned with the facts; women with emotions.
5. Orders vs. proposals – men use imperatives; females use hidden directives.
6. Conflict vs. compromise – men will argue; women will try to find a middle ground.

Sex difference in language development

From Western Europe, numerous studies have reported that females have advantage in language skills development over males (Joseph 1993, 1996, 2000). Several of these studies reviewed by Joseph (1993, 1996, 2000) have shown that females tend to display language, articulation, word knowledge, syntactic, and linguistic superiority over males (Hampson & Kimura, 1992; Hyde & Linn, 1988; Kimura, 1993; Levy & Heller, 1992; Lezac, 1983; McGlone, 1980). Hence, across many domains of language, female language skills are more highly developed and often more complex than the language skills of their male counterparts (Cornett, 2014). For instance, in a vast study of over 13,000 children in ten different language communities, Eriksson et al. (2012) found girls to be more advanced than boys

Comment [P9]: Refer to comment in the draft.

in language abilities in each language community. Specifically, results showed girls to be ahead of boys in early communicative gestures, in productive vocabulary, and in combining words. Although there existed great variation between the children's language abilities from community to community, the female advantage persisted throughout. In a similar study, Tse, Kwong, Chan, and Li (2002) set out to determine sex differences in language ability among Cantonese-speaking children.

In particular, Tse et al. (2002) focused their efforts on the syntactic domain of language. They analysed utterances spoken by children of ages 3 to 5 during spontaneous play. They found significant sex differences between girls and boys in syntactic development. Girls outperformed boys in mean length of utterance (MLU), some sentence types and structures, and syntactic complexity (Tse et al., 2002). Essentially, sex differences in language development appear to persist across various languages and cultures as well as across the different domains of language.

Similarly, Shaywitz, Shaywitz, and Pugh (1995) demonstrated marked differences in bilateral cerebral activation between males and females engaged in language tasks. Females tend to speak more rapidly than males, vocalise more as infants, speak their first words earlier, and acquire vocabulary at an earlier age. Their speech as children is easier to understand; they acquire articulation and grammatical skills at a faster rate, and the length and complexity of their sentences is greater than those of males (Hyde & Linn, 1988; Levy & Heller, 1992; Lezac, 1983; McGlone, 1980). Such differences seem to persist, with males tending to suffer more from such language-related disturbances as stuttering (Corballis & Beal, 1983; Lewis & Hoover, 1983).

Males tend to lose language-related capabilities more than females do as they reach old age, are more likely to become aphasic following stroke, and tend to recover lost language capability less quickly and fully (Joseph, 1993, 1996, 2000). In school, girls tend to learn how to read more quickly and more proficiently than boys and they possess superior reading comprehension, writing, and spelling skills (Lewis & Hoover, 1983; Warwick, 1992). At the same time, they are far less likely to suffer from such reading disorders as dyslexia (Corballis & Beale, 1983). The US Department of Education reports that the writing skills of 9-year-old girls (4th grade) are equal to those of 13-year-old boys (8th grade), and that at all age levels females outscore males in reading and writing proficiency (see Joseph, 2000).

Opinion differs, however, over whether the above differences are (a) primarily maturational, (b) culturally conditioned, or (c) a combination of the two. In this respect, Preston (1962) claimed that boys surpassed girls in reading achievement in Germany, but were consistently behind girls in the United States because, at that time, male teachers predominated in German primary schools whereas the reverse was true in American schools. Johnson (1974) investigated sex differences in reading at grades 2, 4, and 6 in four countries and found support for Preston's assertion. Some 20 years later, Warwick (1992) found support for the hypothesis that boys' performance relative to that of girls is reduced as a result of the preponderance of female teachers in primary schools. Warwick also proposed that boys are too language immature to receive formal reading instruction at age 5. This possibility prompted the writers to investigate whether this hypothesis is also

true for young children in Hong Kong, where female teachers exclusively staff preschools and nurseries.

Mean length of utterance (MLU)

Brown (1973) was one of the first to use MLU to assess and track syntactic development in young children. He used the total number of morphemes (units of meaning) rather than words to indicate the length of utterances produced in natural contexts. In all, length of sentence has been found to increase during preschool years with the age of the child (Chen, 1995; Jin, 1994). Since a large-scale study by Yang and Zhang (1974) reported significant sex difference in MLU in Mandarin, with girls outperforming boys, the writers were interested in seeing whether there are similar sex differences in MLU in early childhood in Cantonese-speaking children. In addition, it has been found that men use longer words than females (e.g., Gleser, Gottschalk, & John, 1959; Mehl & Pennebaker, 2003; Mulac & Lundell, 1986).

Syntactic complexity

A major indicator of syntactic development, syntactic complexity, has been extensively explored in studies of different languages. For example, Kwong (1990) and Zhu (1979) found that sentences are gradually refined as children develop, with increases in complex modifications such as modifiers, verbs in serial expression, and subject/predicate as object or subject. Kwong found that, between the ages of 3 to 4 years, the proportion of declarative sentences with simple modifiers increased significantly and achieved 60% by the age of 5 years. This figure is much lower than the 85.2% figure for the age of 5 years reported in Zhu's study. This difference in proportion may be a

consequence of the small and possibly unrepresentative samples in Zhu's and Kwong's studies. The small sample size in Kwong's study (60 children aged from 3 to 5 years) limits the generalisability of her findings. It is interesting to note that Kwong (1990) noted a spurt at the age of 4 years, with syntactic complexity increasing significantly. No such significant differences were found between the ages of 4 and 5 years.

Conceptual Frameworks

This section discusses the concept of communication. Academic communication, which is a form of communication, is also given attention to, and arguments are put forth to show that it is the conceptual basis of this study. The various features of academic communication are reviewed as well.

The Concept of Communication

Keyton (2011) defines communication as the process of transmitting information and common understanding from one person to another. Similarly, Cheney (2011) adds that communication is a process whereby information is enclosed in a package and is channelled and imparted by a sender to a receiver via some medium. The definitions of both Keyton (2011) and Cheney (2011) underscore the fact that the receiver then decodes the message and gives the sender a feedback. Communication requires a sender, a message, and an intended recipient, channel, and feedback; however, the receiver need not be present or be aware of the sender's intent to communicate at the time of communication in order for the act of communication to occur. For this reason, senders could send written messages to their recipients who are closer or at far distances (Antos, 2011).

The sender is the initiator of the communication in that he/she is the one who has a need or desire to convey an idea or concept to others. Ludlow and Panton (1992) claim that for a sender to communicate effectively, he/she must use effective verbal as well as nonverbal techniques. Flesch (1996) adds that the act of speaking or writing clearly, organizing one's points to make them easy to follow and understand, maintaining eye contact, using proper grammar and giving accurate information are all essential in the effectiveness of the message. On the contrary, the sender can lose his/her audience cooperation if the audience becomes aware of obvious flaws on the sender's part (Lunenburg: 2010). It is for this reason that it important that the sender has some understanding of who the receiver is in order to modify the message to make it more relevant.

The receiver is the individual or individuals to whom the message is directed. According to Lunenburg (2010) the extent to which the receiver comprehends the message depends on a number of factors, which include the following: how much the individual or individuals know about the topic, their receptivity to the message, and the relationship and trust that exists between sender and receiver. All interpretations by the receiver are influenced by their experiences, attitudes, knowledge, skills, perceptions, and culture. It is similar to the sender's relationship with encoding.

The message is the outcome of the encoding, which takes the form of verbal, nonverbal, or written language. The message is sent through a medium or channel, which is the carrier of the communication (Gibson & Hodgetts, 1990). The message may be the most crucial element of effective communication because it can come in many different forms, such as an oral

presentation, a written document, an advertisement or just a comment. In the basic communication model, the way from one point to another represents the sender's message traveling to the receiver. The message is not necessarily what the sender intends it to be. Rather, the message is what the receiver perceives the message to be. As a result, the sender must not only compose the message carefully, but also evaluate the ways in which the message can be interpreted.

Wright and Noe (1995) indicate that the route through which the message passes is what is referred to as the channel. When an appropriate is one chosen, the message enters the decoding stage of the communication process. Decoding is conducted by the receiver. Once the message is received and examined, the stimulus is sent to the brain for interpreting, in order to assign some type of meaning to it. It is this processing stage that constitutes decoding. The receiver begins to interpret the symbols sent by the sender, translating the message to their own set of experiences in order to make the symbols meaningful. Successful communication takes place when the receiver correctly interprets the sender's message.

Feedback is the final link in the chain of the communication process. After receiving a message, the receiver responds in some way and signals that response to the sender. The signal may take the form of a spoken comment, a long sigh, a written message, a smile, or some other action. "Even a lack of response, is in a sense, a form of response" (Bovee & Thill, 1992). Without feedback, the sender cannot confirm that the receiver has interpreted the message correctly. Feedback is a key component in the communication process because it allows the sender to evaluate the effectiveness of the

message. Feedback ultimately provides an opportunity for the sender to take corrective action to clarify a misunderstood message. Feedback plays an important role by indicating significant communication barriers: differences in background, different interpretations of words, and differing emotional reactions (Bovee & Thill, 1992).

Communication can be categorised into various forms. According to Mikoluk (2013) communication has been grouped into verbal, non-verbal and written kinds. The verbal form of communication is associated with auditory means, such as speech, song, and tone of voice. The non-verbal form has to do with body language, sign language, paralanguage, touch, eye contact, through media, i.e., pictures, graphics and sound. The written form concerns the use of elements such as memos, letters, dissertations, research articles among others. Since this research used research articles for analysis, it could be mentioned that the written form of communication was the form chosen for this research.

Academic writing

In this section, I give attention to academic writing as a discipline by defining and discussing the major features of academic writing. Some types of academic writing are also examined. I then narrow down to a detailed discussion of RAs as a major genre in the research community.

Definition and functions

The primary purpose of universities and other research-oriented institutions is to transfer and disseminate knowledge. It is the duty of every academic to share any new findings with the academic community, in an effort to advance the existing knowledge pool. In this way, the writing produced in

the academic setting aims to educate the next generation of scholars (Joubert & Rogers, 2015). This effort by which academics share new findings and educate others in an academic community through writings such as RAs is what is termed academic writing, in its most simplistic definition. Indeed, as Salski (2013) points out, a definition of academic writing is not easy to find or formulate. Different schools of thoughts dictate the focus of each definition. For example, Nordquist (2013) approaches the task from a textual perspective and sees academic writing as “the forms of expository and argumentative prose used by university students and researchers to convey a body of information about a particular subject.” Generally, academic writing is expected to be precise, semi-formal, impersonal, and objective.

This perspective characterizes academic writing through the prism of genres typically employed by the academic community. On the other hand, the word *writing* can also be read as referring to human activity, and in this sense the best academic writing has one underlying feature: it is deeply engaged in some way with other people’s views. Too often, however, academic writing is taught as a process of saying ‘true’ or ‘smart’ things in a vacuum, as if it were possible to argue effectively without being in conversation with someone else. (Graff and Birkenstein 2007:3).

Academic communication includes articles, books, essays, ethnographies, monographs, research papers, conference papers, explications, technical reports, dissertations and theses which are the outcomes of the different forms of inquiry that academics undertake (Badley, 2013). They are mostly written within an academic context such as a university and written for other academics in their own particular discourse communities. For example,

engineers usually write their research papers for other engineers while historians usually write their monographs for other historians. Sometimes, however, academic writers write articles and books for multi-disciplinary or even more general audiences. By learning to write for a wider readership, academic texts which are not just original, rigorous and significant but which are also accessible or even enjoyable (Glassick et al., 1997) may sometimes be produced.

Academic writing serves a number of functions. Among other things, academic communication helps to educate, contribute to knowledge, and raise the reputation of the academic setting from which they are produced. For example, good universities are considered to be those whose research performance, as measured by the number of publications and citations per year, is excellent (Priestley, 2015). Although the various systems applied to rank the quality of universities employ various criteria, academic excellence and research performance are always among the main factors. Some ranking systems, for example, the Academic Ranking of World Universities (ARWU; www.shanghairanking.com), do not only rate the number of publications produced but specifically focus on the number of articles published by the prominent journals *Nature* and *Science*.

The ability to effectively write standard written English is particularly important in higher education, where proficiency in written communication is considered a critical student learning outcome (SLO). A survey conducted by the Association of American Colleges and Universities (AACU, 2011) found that 99% of the chief academic officers from 433 higher education institutions rated writing as one of the most important intellectual skills for their students.

More recently, the Educational Testing Service (ETS, 2013a) conducted interviews with provosts or vice presidents of academic affairs from more than 200 institutions regarding the most commonly measured general education skills, finding that written communication was the most frequently mentioned competency considered by respondents as critical for both academic and career success.

Comment [P10]: Recast sentence.

The focus on written communication is also apparent internationally. Notably, written communication is included as a *generic* skill expected of all students in the Assessment of Higher Education Learning Outcomes (AHELO) project, an effort to evaluate general learning outcomes of college students across nations, which is sponsored by the Organisation for Economic Co-operation and Development (OECD, 2012).

Features of academic writing

Although different disciplines and genres dictate different topics, interests, copora, and technical language, all forms of academic writings have some common features. These include the use of tentative statements, specific academic wording, writing objectively, and acknowledging one's sources.

Tentative writing

Research in academia is cumulative which means that research conducted today may depend to a large extent on the work of numerous other researchers over a long period. The results of experiments are rarely conclusive and findings may later be shown to be inaccurate or based on false assumptions. For these reasons, it is a good idea to be fairly tentative in one's writing in the results and discussion sections. For example, rather than writing

Comment [P11]: Refer to comment in the draft.

‘This result proves...’ it is preferable to use expressions such as: *tends to*, *appears to*, *suggests that*, *would seem to*, and so on and so forth. Academic writing is a cooperative enterprise where new concepts, theories and hypotheses are built on the foundations of existing ideas. The context of an author’s scientific report or essay is therefore the scientific work that has been done before. Therefore, an author must acknowledge the earlier works to which he/she referred. This enables the reader to trace the source of an idea or fact to its origin and to avoid plagiarism.

Academic communication takes several forms. Each has its approach and focus. Yet all share similar features, as has been described earlier, despite the differences in approach. These include books, essays, ethnographies, monographs, research papers, conference papers, explications, technical reports, dissertations and theses. Among these, it has been indicated that RAs are the most important mode of communication in the academic communities (Priestley, 2015). Several authors have discussed specific forms of academic communication.

Theses and Dissertations

For many students, the thesis (also referred to as dissertation) represents the first attempt at writing a formal scientific document. Although the internet supplies ample advice on how to write a thesis and universities tend to supply good templates, most students face a major challenge when writing their master’s thesis. They are aware of the importance of this document in that it represents the formal product of their studies and serves as the basis of which their performance and achievements can be assessed. Ideally, thesis and dissertations are written in a manner that renders them

Comment [P12]: Refer to comment in the draft.

suitable for subsequent publication in an appropriate scientific journal. For most academicians, the publication track record is of fundamental importance, and the sooner they can establish their articles in the scientific community, the better are their chances of advancement in academia.

However, not all students' projects are suitable for eventual publication, and this may not necessarily reflect the student's ability to address the academic community in question. Some projects simply do not deliver publishable results, or they just form a part of a larger study that will be published by other authors. Students whose projects involve collaboration with a pharmaceutical company may additionally be faced with confidentiality issues that prevent them from making their findings available to a broader audience. In any case, students' theses have to meet high standards in terms of content, format, and style, but there is no general consensus on how to present and structure the data.

The type of structure chosen depends primarily on the nature of the study, as well as on guidelines and example documents provided by the university or the institution at which the research was conducted. If the outcome of a student's project is suitable for publication, the structure of the manuscript is essentially guided by the specific author instructions of the chosen journal. Most journals follow the classic IMRAD structure (an acronym based on the first letters of Introduction, Materials & Methods, Results, and Discussion), or a modification of this. For example, Introduction may be replaced by Background, Methods by Procedures, and Results by Findings. Clearly, this simplifies the task of writing a student's thesis to some extent because author guidance tends to be concise, and novice writers can

consult examples of papers published by the journal in question. If you do not intend to publish the data generated within the study for whatever reason, the thesis should be written in the form of a book consisting of chapters.

Although the number and nature of chapters depends on the specific research project and extent of information accumulated, the structure suggested in Table 1 can be applied to most situations.

Table 1: Suggested structure of students' thesis/dissertation

Title page	Use the format suggested by your institution. The title page (one page only, not numbered!) should indicate that the document is a master's thesis and should include the title of the study, author details (i.e., your name, current degree such as B.Sc., and affiliation), as well as the date of release. Most universities also require supervisor details. A formal statement is commonly required, e.g., "Submitted in partial fulfilment of the requirements for the degree of Master of Science" or "Research project submitted in partial fulfilment of the requirements for the degree of Master of Science"
Abstract	Abstract Provide a brief (usually ≤ 250 words) abstract using IMRAD, in line with the main part of the thesis. Make sure the most important findings (including numerical data, e.g., percentages and p-values) appear in the abstract
Chapter 1: Background or Introduction	Describe the rationale behind your project. Why did you embark on this research project? Summarize what is known already by crediting the relevant authors (see List of References, below). Do not reveal any of your new findings in the Introduction
Chapter 2: Literature Review	This chapter should only be added if you have reviewed the literature supporting your research question. If the existing literature is limited, include a summary in Chapter 1
Chapter 3: Methods and Materials	This section should detail all methods used during the study, including statistical methods and data analysis. Some templates foresee a separate section for data analysis, but this may not make sense for your specific project. Provide sufficient and accurate information that would allow other scientists to apply your methods to their own work or repeat experiments that you have conducted
Chapter 4: Results	Summarize your results, using sensible subheadings that allow readers to quickly find the information they are looking for. Use tables for numerical findings and graphs for developments, trends, progress, etc.

Table 1: (Continued)

Chapter 5: Discussion and Conclusions	Discuss your findings in a qualitative sense, elaborating sensible explanations for them. Do not “over-hypothesize” if there is no plausible explanation. Place your findings in the context of what is already known, giving credit to those who reported any previous findings (see List of References, below)
Chapter 6: List of References or bibliography	List all references cited in your thesis, using an accepted format (as suggested by your institution or formatted automatically by a reference manager tool). The reference format should be used consistently throughout
Chapter 7: Appendix	Include any additional materials, such as raw data, detailed method descriptions, randomization codes, tables and figures not included in the main part, etc.

Source: Adapted from Priestley (2015)

In a section placed before or after the main text, you may wish to acknowledge any help you have received during the studies. This may include supervisory efforts, laboratory assistance, statistical help, or even editorial support.

It is noted that thesis/dissertations at undergraduate, master’s and doctoral levels, essentially follow the same pattern of writing. However, a more popular approach nowadays for doctoral thesis, is to structure it as a series of articles suitable for publication in scientific journals (Priestley, 2015). As pointed out earlier, scientific endeavours live from sharing information and contributing to the “knowledge pool.” Moreover, a scientist’s (academic’s) success hinges on his or her publication track record, and early visibility in the scientific community is clearly advantageous. Moreover, the thesis structure based on individual manuscripts spares students the effort of having to write the thesis and articles for publication separately. Nonetheless, not all universities support this approach, and one’s institution and/or supervisor will have to advise on this. The Internet provides ample advice on the preparation of a doctoral thesis, but the most appropriate inspiration usually comes from

good examples written by members of your group or department. There is no hard and fast rule on the composition, contents, and structure of a doctoral thesis; you are the author and the expert!

Research articles

RA is potent means of disseminating updated and relevant information to the academic community and a veritable platform for academic development. A number of avenues exist for publicising one's research findings. These include conferences, and posters. The most common approach, however, is to publish in an academic journal. An academic journal is a periodical publication in which scholarship relating to a particular academic discipline is published (Blake & Bly, 2000). Academic journals serve as forums for the introduction and presentation of new research, and the critique of existing research. To make academic journals scholarly, they are usually peer-reviewed or refereed. A basic requirement for any good journal is that the articles therein must be easy to read. In fact, readability is a prerequisite for any good writing. Incomprehensible texts hamper transfer of understanding between readers and authors. For this reason, readability should be every researcher's primary goal. Therefore, no matter the credibility, prestige, or popularity of a journal, the basic function to transmit information will be hampered if the journal is not readable.

An RA is a substantial piece of academic writing in which the author does independent research into a topic and writes a description of the findings of that research. An RA is more complex than a simple reporting of events. It is a document written by experts for experts in one's academic discipline

(Majumder, 2015). It is a construction intended to persuade the reader of the validity of the claims. An RA reports the results of original research, assesses its contribution to the body of knowledge in a given area, and is published in a peer-reviewed scholarly journal. A given academic field will likely have dozens of peer-reviewed journals. For university professors, publishing their research plays a key role in determining whether they are granted tenure (Hall, 2017). An RA serves not only to further the field in which it is written, but also to provide the student with an exceptional opportunity to increase her knowledge in that field (Purdue University, 2007).

Once, research articles had only a limited audience consisting mainly of other scholars and graduate students. Today, websites such as Google Scholar and the proliferation of electronic academic journals have broadened the potential audience for research articles.

Types of RAs

There are a number of RAs that can be written by an academic. Primarily, the type of RA is dependent on the type of research design. Basically, RAs are grouped into two main types:

1. Primary Research Articles (PRA) and
2. Review Research Articles (RRA)

Primary RA

A primary RA reports on an empirical research study conducted by the authors. These articles describe an original experiment (empirical research) or analysis that adds to current knowledge on a particular topic. Such articles will

include background information, the methods the scientist used, a description of the results, and an analysis of what the results mean in the context of current knowledge. Primary RAs are mostly discipline-specific. Generally, PRAs are termed empirical RAs, and convey the idea that the authors will have collected data to answer a research question. Empirical RAs contain observed and measured examples that inform or answer the research question. The data can be collected in a variety of ways such as interviews, surveys, questionnaires, observations, and other various quantitative and qualitative research methods (Dalal, 2016). Below are some types of primary RAs.

Clinical case study RAs: Clinical case studies present the details of real patient cases from medical or clinical practice. The cases presented are usually those that contribute significantly to the existing knowledge on the field. The study is expected to discuss the signs, symptoms, diagnosis, and treatment of a disease. These are considered as primary literature and usually have a word count similar to that of an original article. Clinical case studies require a lot of practical experience and may not be a suitable publication format for early career researchers.

Clinical trial RAs: Once again, specific to the field of medicine, clinical trials describe the methodology, implementation, and results of controlled studies, usually undertaken with large patient groups. Clinical trial articles are also long, usually of about the same length as an original research article. Clinical trials also require practical work experience, as well as, high standards of ethics and reliability. So this format is more useful for experienced researchers.

Review RAs

Review of RAs synthesizes current research on a specific topic. Often an article will summarize past research, identify important people in the field, outline recent advances, and point out gaps in a body of knowledge. Review articles are often located in the same journals PRAs, but do not report original research. Review articles are a great resource if you're looking for an overview of a small topic, with complete and current information. Review articles are well-cited, so they can provide a starting point for more extensive research (University of Toronto, 2017).

Book review: Book reviews are published in most academic journals. The aim of a book review is to provide insight and opinion on recently published scholarly books. Book reviews are also relatively short articles and less time-consuming. Book reviews are a good publication option for early-career researchers as it allows the researcher to stay abreast of new literature in the field, while at the same time, adding to his publication list.

Theoretical review article: Review articles give an overview of existing literature in a field, often identifying specific problems or issues and analysing information from available published work on the topic with a balanced perspective. These are considered as secondary literature and can be a particularly efficient way for early career researchers to begin publishing. Review articles can be of three types, broadly speaking: literature reviews, systematic reviews, and meta-analyses. Review articles are usually long, with the maximum word limit being 3000-5000 or even more, depending on the journal. However, some journals also publish short reviews.

Perspective, opinion, and commentary: Perspective pieces are scholarly reviews of fundamental concepts or prevalent ideas in a field. These are usually essays that present a personal point of view critiquing widespread notions pertaining to a field. A perspective piece can be a review of a single concept or a few related concepts. These are considered as secondary literature and are usually short articles, around 2000 words. Opinion articles present the author's viewpoint on the interpretation, analysis, or methods used in a particular study. It allows the author to comment on the strength and weakness of a theory or hypothesis. Opinion articles are usually based on constructive criticism and should be backed by evidence. Such articles promote discussion on current issues concerning science. These are also relatively short articles. Commentaries are short articles usually around 1000-1500 words long that draw attention to or present a criticism of a previously published article, book, or report, explaining why it interested them and how it might be illuminating for readers.

Conference paper

Academic conferences have various formats, but in English, the default mode is a series of panels in which 2–4 scholars deliver presentations and then respond to comments and queries. A conference paper is the text for one of these presentations. The presentation length for the BYU English Symposium is 15 minutes, or around 7 pages of text (Brigham Young University, 2013).

Presenting at a conference is a great opportunity for gaining valuable feedback from a community of scholars and for increasing your professional stature in your field. Unlike RAs which do not enable readers to seek clarification from

what they read by asking questions, a conference paper is written such that the author stands in front of an audience of experts as well as novices, to present his findings and thoughts (UNC, 2010).

Comment [P13]: ?

Critical review

Although the term critical review has been used quite extensively in academia, there seem to be no definite and straightforward definition. Different people have applied the term to mean different things (Edgley, Stickley, Timmons, & Meal, 2016). As has been noted by Edgley et al., (2016), a brief online search suggests that the term ‘critical review’ is used in a variety of different ways within the literature, which can be a source of confusion, particularly for students. Together, the words might be understood as a compound noun – a ‘critical review’ (similar to other research exercises such as a ‘systematic review’ or ‘literature review’). Alternatively, it can be used as an adverb indicating a judgement, for example ‘this is a critical review of the literature’.

Somehow, however, Grant, Booth, and Centre (2009) attempt to ‘normalise’ the situation by giving a description of what critical review encompasses. They note that a critical review aims to demonstrate that the writer has extensively researched the literature and critically evaluated its quality. It goes beyond mere description of identified articles and includes a degree of analysis and conceptual innovation. An effective critical review presents analyses and synthesizes material from diverse sources. Its product perhaps most easily identifies it—typically manifest in a hypothesis or a model, not an answer. The resultant model may constitute a synthesis of

existing models or schools of thought or it may be a completely new interpretation of the existing data.

Readability and readability formulas

Several authors have discussed readability since the 1930s (Bailin & Grafstein, 2016). Generally, readability refers to the ease with which a document can be read (Ahmed, Zeeshan Shaukat, & Islam, 2013). The study of readability is the study of those properties of written texts that aid or hinder the effective communication of ideas and information (Bailin & Grafstein, 2016). It has been used as a proxy for the comprehension of a written material (Richards & Van Staden, 2015). In the broadest sense, however, Dale and Chall (1949) define readability as follows:

The sum total (including 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 optimum speed, and find it interesting (p. 19).

From Dale and Chall's definition, it implies that there are a number of factors that dictate the ease (difficulty) of reading a text. While some of these factors are directly related to the text itself (such as the syntactic elements), others depend on the ability of the reader. The measure of readability in classical terms, focus on the former factor rather than the latter or both (Bailin & Grafstein, 2001).

In order to measure readability, a variety of techniques have been developed. The most common approach is the use of readability formulas,

which have been widely used since the 1950s. Richards and Van Staden (2015) and Bailin and Grafstein (2016) have discussed these approaches adequately, and their discussion has been adopted in this study. They described readability formulas as mathematical equations derived by regression analysis, in which a model or equation is constructed that best predicts the reading grade level of readers who comprehend a given text. Readability formulas are based on the assumption that how difficult a text is to read is related to whether or not the words in the text are understood, and whether or not these words are put together in an easy-to-follow manner (Bailin & Grafstein, 2001). Simply put, readability formulas are based on vocabulary difficulty and syntactic complexity. Vocabulary difficulty refers to the degree to which a text contains words that are unfamiliar and/or difficult to understand. Syntactic complexity refers to the degree to which the sentences in a text have complicated grammatical structures. That is, the longer a word is, the more difficult it is to comprehend. The interrelationships between the various syntactic components, readability, and text comprehension have been illustrated with Figure 2.

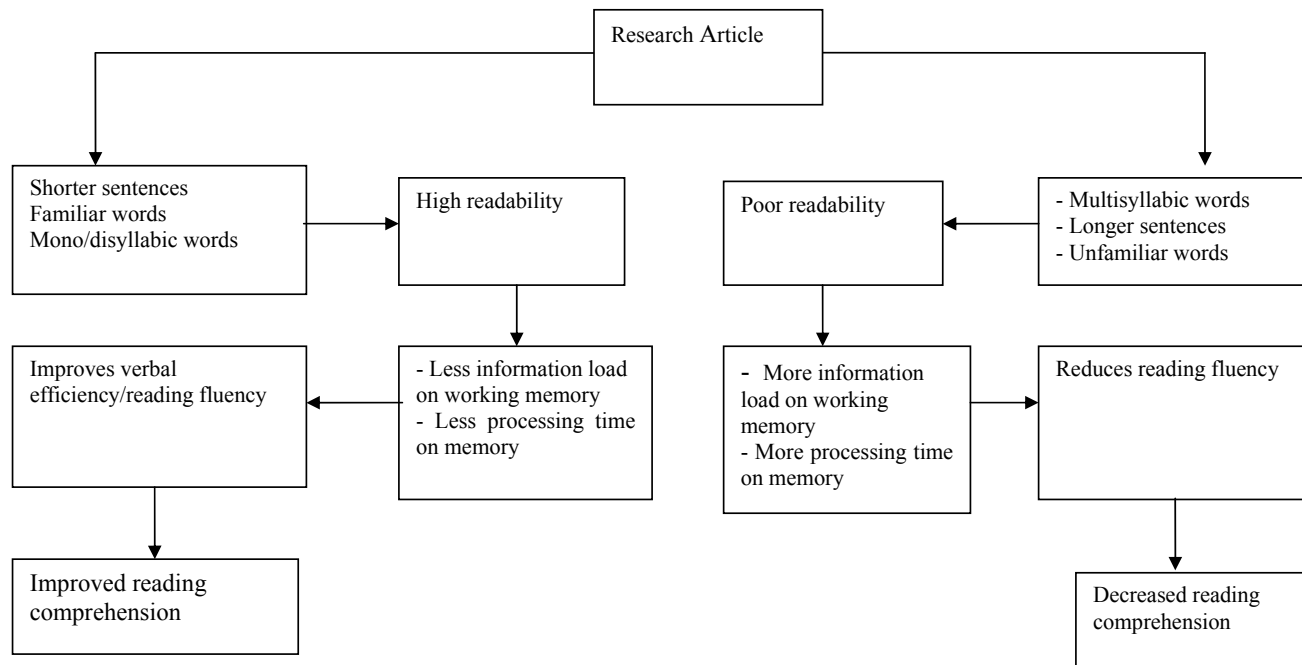


Figure 2: Conceptual framework of readability and text comprehension of RAs.

Source: Author's Construct (2016)

In the construction of readability formulas, comprehension is assessed by a specific pass score on test questions based on the content of the text (Mandic et al., 2012). Chavkin (1997) noted that the two strongest and most common elements used in analysing written material for ease of understanding are vocabulary difficulty and sentence length. Indeed, it is variations in these two elements that make up the majority of readability formulas. Stevens et al. (1992) supports this, noting that “readability formulas determine the readability level of a passage by examining word difficulty and sentence length” (Stevens et al., 1992, p.1).

Although there are over 200 readability formulas that have been developed since the inception of the concept of readability about a century ago, all could be classified into three categories (Benjamin, 2012) as follows:

1. Formulas based on traditional readability features
2. Methods inspired by advances in cognitive theory and
3. Formulas based on findings in statistical language modelling

This review thoroughly discusses these categories of readability indexes. In some instances, some popular individual formulas are described as well. It will be of little point to cover all individual indexes because earlier researchers have published enormously in this area.

Classical readability indexes

Readability formulas based on traditional readability features utilize traditional syntactic variables such as sentence length, percentage of familiar words, and word length. These formulas are termed classic readability formulas. Passages that contain shorter sentences, shorter words, and more

frequent words would be considered more readable or less difficult than passages with longer sentences, longer words, and rare words. A common shortfall of all classical readability indexes is that the formula might judge even a nonsense passage as quite readable if the text's jumbled words are frequent, short, and organized into brief sentences. This weakness has warranted several criticisms (Bailin & Grafstein, 2016). Yet, classical readability indexes remain very popular. It appears that the larger proportion of all readability indexes use the traditional syntactic variables. In fact, the total number of readability indexes based on these criteria cannot be reviewed in this thesis. Instead, few are described here based on their wide use, simplicity, popularity, and some modern creative adaptations of traditional features like word frequency.

In the academic setting, the Flesch indexes have been used more extensively than all other readability indexes. The Flesch Reading Ease Readability Formula is one of the oldest and it is considered to be the most accurate of all the formulas. It was developed in 1948 by Rudolph Flesch who is a writer and a reading consultant. It is a simple approach to evaluating the grade-level of readers. This formula is mostly used for academic text. It is largely used to assess the difficulty of a reading text written in English language. This formula uses a scale of 0 to 100; where 0 corresponds to a minimum of college education and 100 is equivalent to 4th grade (Primary 4). This simply means that the higher the score the easier the passage to be read and the lower the score the more difficult the passage. In addition, a zero score means text has more than 37 words on the average in each sentence and the average word is more than 2 syllables. Further, a score between 60 – 75

indicates Plain English Score. To assist users, Flesch provided an interpretation table (Table 2) to aid users convert the scale to estimated reading grade and estimated school grade completed.

Table 2: Flesch Readability Index

Reading Ease Score	Description	Reading Grade
0-29	Very difficult	Post graduate grade
30-49	Difficult	College grade
50-59	Fairly difficult	10 th – 12 th grade
60 - 69	Standard	8 th – 9 th grade
70-79	Fairly easy	7 th grade
80 – 89	Easy	5 th – 6 th grade
90 – 100	Very Easy	4 th – 5 th grade

Source: Adapted from Wyatt and Schnelbach, 2008

Flesch-Kincaid Grade Level Test is related to the Flesch reading ease test, and translates the Flesch Reading Ease Test scores to grade level. The formula was propounded by Peter J. Kincaid and his team in 1975. It is mostly used in pedagogy. This formula is used to determine the readability level of a variety of educational materials especially books. This formula makes it easier for parents, teachers, and librarians to select suitable reading texts for their children/learners (Owu-Ewie, 2014).

The Gunning Fog Readability index is mostly referred to as FOG Index. It was propounded by an American textbook publisher by name Robert Gunning in 1952. The formula is used generally to confirm that a text can be read easily by the intended audience. The salient objective of the Gunning Fog Index formula is that short sentences in plain English achieve better

readability scores than long sentences written in complex language. The best score for readability with the fog index is 7 or 8 and anything beyond 12 is too difficult and hard for most people to read.

SMOG (Simple Measure of Gobbledygook) is reading readability formula which predicts the years of formal education needed to understand or comprehend a particular text. This readability formula was created by G. Harry McLaughlin in 1969. The SMOG readability formula was propounded to deal with the lapses in other readability formulas like the Gunning Fog. This formula was developed specifically for checking health messages (Hedman, 2008) but has been applied to language learning texts. Although the SMOG readability formula is seen as being too simplistic, it is preferred in evaluating the difficulty of the language of customer health related materials (Fitzsimmons, Hulley, Micheal, & Scott, 2010). According to McLaughlin (1969) the researcher should select 10 successive sentences at the beginning of the text, 10 from the middle, and 10 from the end, and then count every polysyllabic word and take the square root of the total. The result obtained after this process, represents the reading grade that a person should attain to understand the text.

The (new) Dale-Chall index has been suggested to be the most accurate of all classical readability indexes (DuBay, 2004). Ironically however, it seems to be the least used index in evaluating the readability of texts in all published papers (Benjamin, 2012). The apparent accuracy with the Dale-Chall index stems from the fact that this index focuses on measures of assessing text difficulty that were based on theories in cognitive science. This slight change in paradigm from the use of typical syntactic elements in classic

readability indexes to include other measures of readability propounded in theories in cognitive sciences seem to make up, at least partially, for the pitfalls of classic readability indexes. Following this formula was the digression of readability formulas based on traditional variables to those that fully utilise theories in cognitive sciences.

Several other indexes similar to the ones discussed above have emerged over time, with improvements of sorts, over earlier ones. An example is (but not limited to) the Advantage-TASA open standard for readability index, which is based on the same traditional variables—word length, sentence length, and grade level of words. As an improvement, over typical readability indexes (such as has been described above), the Advantage-TASA open standard for readability index for books also takes book length into account, a factor that was found to significantly influence book difficulty. Another index, which is quite new (developed sometime in 2007) is the Read-X index. This index uses some traditional readability variables—number of sentences, number of words, number of “long words,” and number of letters in the text—to analyse the readability of texts on the Web in real time so that a person can perform a web search and filter results by reading level. The uniqueness of this program (index) lies specifically in its ability to categorize search results by theme (e.g., science, music, history, etc.). In future versions, Read-X is expected to be able to take a user’s existing topical knowledge into account and customize reading level filtering based on a reader’s level of content knowledge about a particular topic (e.g., long words in a particular domain are not necessarily difficult for a reader with a lot of background knowledge in that domain).

Aside these popular examples of classical readability indexes discussed above, there are other forms which are relatively new. These forms of classical readability indexes differ somewhat from earlier forms due to their computational complexities. One such index is the Lexile framework. The Lexile framework (scale) was developed in 1989 by MetaMetrics co-founders A.J. Stenner and Malbert Smith III. The Lexile scale runs from below 0L (Lexile) to above 2000L, though there is not an explicit bottom or top to the scale. Scores 0L and below are reported as Beginning Reader (BR). These books or students may be coded as Lexile: BR. In some cases, a student will receive a BR code followed by a number (e.g. Lexile: BR150L). A measure of BR150L indicates that the Lexile measure is 150 units below 0L. Creators of the Lexile scale attempted to design a scale that matched their definition of reading comprehension, thereby creating a measure with a high degree of construct validity in addition to the predictive validity sought by many traditional scale developers. Determining that reading comprehension depends upon the familiarity of the semantic units and the complexity of the syntactic structures used in constructing the message (Smith et al. 1989), the creators devised a scale that included measures of word frequency (the semantic variable) and sentence length (a proxy for syntactic complexity). Although the Lexile scale readability measure is similar to almost all the other classical readability indexes, the actual measures taken and the application of the Lexile scale differ from prior methods and is complex such that it cannot easily be computed manually as the other readability indexes can (Benjamin, 2012). This is because The Lexile word frequency measure is calculated as the mean log word frequency from a 5-million-word corpus (Carroll et al. 1971), and

their sentence length measure is the log of the mean sentence length in the text. The use of logarithmic functions to determine word frequency and sentence length makes this index quite complex and relatively new in its approach compared to earlier forms of classical readability indexes. Yet, the Lexile framework of high appeal, especially in schools. The appeal of the scale for wide use seems to be, according to Benjamin (2012), largely based on its application: a person receives a Lexile score based on his or her ability to answer comprehension questions correctly; a text also receives a Lexile score. If the person and the text are matched, then the person has a 75% chance of answering a comprehension item correctly for that text. Teachers, then, can look at the Lexile score for a text and determine whether or not that text would be appropriate for a student based on the student's Lexile score.

Readability formulas and cognitive theory

Methods inspired by advances in cognitive theory basically hangs on the idea that text difficulty and readability are more related to coherence and the relationships between elements in a text rather than simply the sum or averages of individual surface features (as used in classic formulas) (e.g. Britton, 1991; Kintsch, 1988; McNamara & Kintsch, 1996).

Just as classical readability indexes are based on specific syntactic features, readability formulas based on theories of cognitive science have two variables, which are used in its assessments. The first variable is termed as propositions and inferences. Simply put, sentences can be broken down into propositions, or brief meaningful units that do not take into account information like tense, voice, or aspect (Graesser et al. 1997). Propositions are

units comprised of a predicate and at least one argument. Propositions can also include other propositions. An argument serves a functional purpose within a proposition, indicating the relationships between meaningful words in the sentence. The example given by Benjamin (2012), is reproduced here to explain this concept. In the sentence “*The nurse placed the scalpel on the table and grabbed the sponge*”, the propositional breakdown is as follows:

1. Place (AGENT = nurse; OBJECT = scalpel; LOCATION = on table)
2. Grab (AGENT = nurse; OBJECT = sponge)
3. And (PROP 1; PROP 2)

Benjamin (2012) explains that in order to carry on coherent discourse or write a coherent and readable text, there must be some propositional or at least argument overlap among successive sentences. Likewise, at the macro level, there must be some propositional connections across the larger text or throughout a conversation if the text or conversation is supposed to address a particular topic. When there are few or no gaps in overlap across sentences, then a text is seamlessly moving from one point of information to another while giving the reader all the help he or she needs to build new knowledge. This type of text is a highly cohesive text: inferences are explicit and the reader does not have to fill many gaps using his or her own knowledge about the topic. However, in a text where less propositional overlap exists, the reader will be required to fill gaps of information with his or her own knowledge. This type of text has low cohesion: inferences are implicit and require more work on the part of a novice reader. Kintsch and van Dijk (1978) describe these latter texts as more difficult texts because a novice reader may not have the schema in place to make the necessary inferences to comprehend the text.

Hence, in cognitively oriented text difficulty analysis, propositions and inferences play an important role, and the analysis and manipulation of these variables can yield significant differences in reader comprehension (Britton and Gülgöz, 1991; Britton et al. 1993).

The second variable in cognitively oriented text difficulty analysis is Latent Semantic Analysis (LSA). Again, Benjamin (2012) provides sound discussion of this variable. He notes that Latent semantic analysis is an automated tool that represents text content (e.g., an individual word and all the contexts in which it appears, for example) as a vector in semantic space. LSA analyses the semantic relatedness either between texts or among segments of text in a more expanded way than simple measures of word overlap. According to him, researchers can train the system on a large corpus of topical text so that it begins to develop a “knowledge” of which words tend to appear in particular contexts. For example, the word *music* probably appears frequently in the same contexts as *guitar*. Thus, music and guitar would have a strong semantic relationship and music would be considered to be an important word in texts containing the word guitar.

In addition, LSA examines the indirect relationships among words within the contexts in which they are used. For example, in a text about *trumpets* where the word *music* exists but *guitar* does not exist, the word *guitar* will still have a positive semantic relationship to the word *trumpet* even though they may have never appeared in the same context (Benjamin, 2012). This happens because through *guitar's* relationship to *music* and other words, the system understands that *guitar* is indirectly semantically related to *trumpet*.

Formulas based on findings in statistical language modelling

As the world moved towards the use of computers and with the development of the world-wide web, the use of information from web pages became popular. Hence, the traditional approach to the readability of texts was employed to evaluate the readability of web pages. Unfortunately, however, traditional formulas have often performed poorly when analysing Web documents. This has been attributed to the significant amount of what Benjamin (2012) describes as “noise” found in web documents (i.e. punctuation errors, sidebar menus, photograph captions) as well as the large number of web pages containing fewer than 100 words. This necessitated a new approach to readability. The resulting formulas developed have been based on findings in statistical language modelling (e.g. Collins-Thompson & Callan, 2005; Si & Callan, 2001). Language modelling provides a powerful means for expanding simple word frequency indices by analyzing the probability of a particular model (e.g., a model for fifth grade texts) generating a word or combination of words. It must be indicated that in contemporary times, much of what readers consume are on webpages. Hence, this approach to readability is laudable, and certainly shows promise.

Critique of readability indexes as a tool

The validity and utility of quantitative formulas began to be questioned during the 1980s (Cain & Miller, 1981; Gordon, 1980; Rygiel, 1982; Templeton; Wheeler & Sherman, 1983).

Recently, Bailin and Grafstein (2016) made some valid arguments against the use of readability indexes to measure text comprehension. First, these authors argued that the assumption that all instances of the property

being counted contribute equally to difficulty or, at the very least, on average contribute an equal amount of difficulty is flawed. They point out that no readability formula using vocabulary difficulty provides evidence that this assumption is actually the case. In fact, it is not clear what evidence could possibly be used to support this assumption. Using simple pairs of monosyllabic and polysyllabic words to contrast and make their case, they question how someone could know how much more difficult it is to use the word *idiot* than it is to use the word *dog*— accepting, for the moment, the assumption that polysyllabic words are necessarily more difficult to understand than monosyllabic words. They further question why someone should assume that the difference in difficulty between those two words (*idiot* and *dog*) is the same as the difference in difficulty between *assessment* and *read*. The authors proffered similar critique about the way in which syntactic complexity is measured. Average sentence length is often used as a measure. The number of prepositions or prepositional phrases is also a metric we have seen in formulas. Still another measure is the number of simple sentences. Whatever the measure, all suffer from the same problem as the vocabulary measures. Counts of formal properties do not translate into units of reading difficulty. If one text has an average sentence length of ten words and another of 15 words, this does not correlate to a difference of some function of five units' difference of difficulty. The same point can be made about all counts of formal syntactic properties. Since all readability indexes use aspects of syntactic complexity as the variable that underpins the index, Bailin and Grafstein (2016) ask: If the individual variables that constitute the readability

formulas cannot be considered meaningful factors in readability, then how can we consider the formulas themselves to be meaningful measures?

In addition, although the different readability indexes have been indicated to have strong correlation with one another, different indexes have produced different scores, even for the same texts. The differences in readability scores across indexes for the same text have been shown to be quite large in a number of cases. For example, Oakland and Lane (2009) have documented that two popular readability formulas, Fry and Dale-Chall, were used to analyse passages from Plato's (2003) *Parmenides*. The readability level of this difficult piece was found to be between the 6th and 10th grades (i.e., ages 12 and 16, respectively) using the Fry formula and, even lower, between the fourth and eighth grades (i.e., ages 10 to 14, respectively) using the Dale-Chall formula (Gordon, 1980). For this reason, data from readability formulas may lack credibility (Oakland & Lane, 2009).

Nevertheless, Chall and Dale (1995) consider the formulas good enough for practical purposes to estimate the reading level of any text. For this reason, readability indexes were deemed adequate to be used as the proxy to measure readability and reading comprehension.

Comparison of some popular readability indexes

There are over 200 readability indexes proposed by researchers (DuBay, 2004). Almost without exception, readability in these formulas is based on syntactic and semantic complexity. The number of words per sentence determines syntactic complexity. The basis for measuring semantic complexity is either by word familiarity as defined by inclusion on a word list or the number of syllables per word. According to Fry (2002), readability of

classroom materials or public documents, usually refers to a numerical or grade-level score that is obtained by applying a mathematical formula to a sample of text.

The classic readability formulas predict comprehension (Pikulski, 2002). Most do so by providing a numerical score representing the educational level necessary to read a document with ranges of 50% to 75% comprehension. However, Holland (2012) indicated that the SMOG formula predicts 100% comprehension. In connection with the popular readability indexes, Burke and Greenberg (2010) intimated that the mathematical formula used to determine the difficulty of a given text typically takes into consideration issues such as sentence complexity (measured by sentence length) and vocabulary difficulty (measured by either the number of letters or syllables in words or by a comparison to lists of easy or difficult words).

In effect, readability formulas measure the relationship between the difficulty experienced when reading a text and the linguistic features, specifically word meaning and sentence structure, of that text (Holland, 2012). Of the readability formulas, Fry Readability Formula is very easy to use because it is based on the average sentence length and average syllable count of 300-word samples from different places in selection. The averages of the scores are plotted on a graph to ascertain the approximate grade level. The Fry formula helps in quickly judging the approximate grade level of a selected text when no other information is readily available.

However, in his study, Bravos (2010) posits that the Flesh Kincaid Readability Formula is the most reliable when used with upper elementary and secondary materials. Gunning Fog, according to Bravos is widely suitable in

the health care and general insurance industries for general business publications. Coleman-Liau is very useful in the area of study involving fourth-grade to college level readers. Due to its simplicity and relative ease of use, Fry Graph according to Bravos is very useful for elementary through secondary materials. The Simple Measure of Gobbledygook (SMOG) is widely used for measuring readability in the healthcare sector.

The Spache readability formula is very useful for third grade level students. Hull Formula and Automated Readability Index (ARI) could be put to good use in the area of technical or specialized writings. Bravos indicated that Bormuth index is very useful in ascertaining readability in documents written for academic purposes. Linsear Write could be used in attesting readability for Air Force technical manuals. McAlpine EFLAW Readability formula is also indicated to be very useful in measuring readability for documents written in connection with individuals learning English as a second language. Fernandex-Huerta, RIX, Hayashi, Douma and Kandel & Moles' Readability Formulas could be used for languages other than English (Chall & Dale, 1995).

Readability and text formality

Formality has been defined from different perspectives. Atkinson (1982) adopted everyday conversations as a standard of informality and defined formal as nonconversational, based on noticeable features that differed from everyday conversations in unfamiliar settings (e.g., court hearing), persons involved (e.g., attorney), or status of readers (e.g., professionals). The scope of this definition applies primarily to spoken registers, including an attempt to account for silence, repair, and turn allocation. Andren, Sanne, and

Linell (2010) adopted organizational or institutional conventions as a standard of formality. They defined formality to be achieved when officially standardized and recognized institutional conventions or prescriptions are employed. Four classes of informal linguistic characteristics were adopted by Andren et al (2010) to explain formality in conversations. These are informal lexical embedding (e.g., “hi there”; Li et al., 2015), colloquial style or jargon (e.g., “what do you say to that?” instead of “how do you plead, guilty or not?”; Linell et al., 1993), omissions of formally required parts (e.g., abbreviations), and addition of nontask talks (e.g., phatic talk). These definitions provided features that distinguish formality, albeit not adequate for general acceptance by all experts of the field.

According to Heylighen and Dewaele (2002) a text is recognized as formal by the use of actual sequence of linguistic symbols rather than by the cluster of implicit, poorly delimited, and subjective factors that constitute a context. The context itself of a text is considered as a feature of formality by Heylighen and Dewaele (2002). They defined context to include all those features of a discourse (oral or written) which helps shed light on the ideas been expressed. According to them, low context increases formality, whereas high context decreases formality. In low-context situations, the characteristics of explicitness are borrowed from logic and the mathematics of formal language (Bussmann, 1996; Chomsky & Miller, 1963) to describe formality. For example, the formal language of legal scholarship is intentionally explicit and unambiguous; no matter who reads it and where it is read, it is designed to be understood with the same precision. In contrast, the informal language at a party would be implicit and ambiguous, which would require shared

knowledge and common ground for accurate understanding. Graesser et al. (2014) claim that formal discourse, either in print or preplanned oratory, occurs when there is a need to be precise, coherent, articulate, and convincing to an educated audience. Its opposite end of the continuum is informal discourse in oral conversation, personal letters, and narrative, which are replete with pronouns, deictic references (e.g., here, there, this, that), verbs, and reliance on common background. In the current study, I adopted this theoretical definition and measure of formality but further evaluated it to see whether it accounted for human judgments of formality.

Graesser et al (2014) specifies formality as being on a continuum rather than discrete. Graesser et al. (2014) consider many levels of language and discourse in their theoretical framework and formulation of the composite formality score. Thus, this measure is broader than alternative measures of formality. The theoretical framework also considers important aspects of context (such as purpose, discourse planning, audience, and common background), even though the formality measures are based on the features of the text rather than context. I developed the broad definition of formality in the directions for human judges when they scaled texts on formality.

Formal language can theoretically be distinguished by linguistic features, such as choices of words, syntactic structures, semantic and pragmatic meanings, and multitextual levels of discourse (Biber, 1988, 1995; Carroll, 1966; Conrad & Biber, 2001, 2009; Crystal, 1987; Graesser et al., 2014; Graesser, McNamara, & Kulikowich, 2011; Sardinha & Pinto, 2014). In other words, formal texts are characterized by polysyllabic words and complex grammatical structures, which are ingredients for low readability. Therefore, a

formal text has low readability. Conversely, informal texts have high readability. RAs are written in formal language, and so reflect all the characteristics of formal texts which include the use of polysyllabic words and complex grammatical structures. Because academic writings are types of formal writing, it is important not to use the informal short forms (contractions) that are used in writings meant for friends. For example, one should use: do not instead of don't; cannot instead of can't and will not instead of won't.

CHAPTER THREE
REVIEW OF EXTANT LITERATURE ON READABILITY STUDIES
AND ACADEMIC WRITING

Introduction

This chapter explores and reveals existing literature on readability and academic writing across disciplines. Here, the researcher positions his work as a maiden study in the area of readability and academic writing in Ghana. There is also attention given to readability studies in some other domains such as insurance, health, information technology, and others. This is to establish that readability analysis has been very useful in other domains other than pedagogy and academic writing.

Readability of RAs

Over the years, several authors have evaluated the readability of RAs. Different authors have focused on RAs in different journals. Fields of studies that have been evaluated include the humanities (e.g. Dolnicar & Chapple, 2014; Lee & French, 2011), science and medicine (e.g. Hayden, 2008; Severance & Cohen, 2015), business (e.g. Lee & French, 2011), and education. This section of this study presents a review from extant studies that have evaluated the readability of research articles from different disciplines. The various methodological approaches that were used, the readability indexes, as well as the major findings are discussed. Objective critique of the approaches of each author has been included.

In the humanities, Dolnicar and Chapple (2014) evaluated the readability of research articles in tourism journals. Abstracts from full-length

Comment [P14]: Refer to comment in the draft.

original research articles published in *Annals of Tourism Research*, *Tourism Management*, and the *Journal of Travel Research* were included. Thinking that the readability of research articles should improve over time due to improved review process, articles published at three points in time were chosen: 20 years before the research (1993), ten years before the research (2003) and at the time of undertaking the study (2013). The authors selected RAs from these journals because these were the tourism journals with the highest impact factors. As a consequence, they received a large number of submissions annually. The authors held the view that since the journals they selected had the highest impact factor (indicating high citation) the research articles from these journals were also subjected to the most rigorous reviewing process which, according to them, was expected to improve manuscript readability. The sample sizes used in their study were: Ninety-five articles published in 1993, 143 in 2003 and 257 in 2013. One hundred and nine appeared in *Annals of Tourism Research*, 131 in *Tourism Management* and 58 in the *Journal of Travel Research*. Using Flesch's readability index, the authors found that the average Flesch scores for the three journals were similar, ranging from 17 to 19, indicating that the three tourism journals are very difficult to read. Only three articles (1% of the articles) were found to be fairly easy to read, while 75 articles (15%) were difficult and 372 articles (75%) were very difficult to read. Although the sample size used was quite substantial, using abstracts may not be fair representation of the readability of the entire research paper. Being apparently aware of the limitation of using abstracts in assessing the readability of full – length research papers, the authors analysed the correlation between Flesch reading ease score of the

abstracts for the abstract, the introduction and the conclusions section for articles published in 2003. The correlation was found to be significant at the 99% significance level for abstract and introduction (0.373) and abstract and conclusions (0.421). Hence, the authors concluded that abstract readability reflected readability of other sections.

It must be noted however, that the correlation coefficients (r) were only 'moderate' in strength. In addition, the statistical significance attained could have been to the relatively large sample size used (495), as has been indicated elsewhere (Field, 2011; Julie Pallant, 2013; B G Tabachnick and Fidell, 2007). A strong correlation (r at least .7) between the readability of the abstracts and the other sections of the research articles used in this litmus test, would have given greater confidence in their conclusions. This view is attested to by the findings of Hartley, Pennebaker, and Fox (2003) who analysed separately the abstracts, introductions, and discussions of 80 research articles from the *Journal of Educational Psychology* that were published between 1997-2001. The results showed that the abstracts scored worst on most of these measures of readability, the Introductions came next, and the Discussions did best of all. Therefore, using text from just one section of a research article to judge the article's readability is likely to lead to bias conclusions.

James, Sotto and Fox (2004) investigated the readability of research articles across three disciplines namely: the Sciences, Social Sciences, and Arts and Humanities. In doing this, they considered features such as sentence length and percentage use of passive tenses. Focusing on readability of research articles from the Arts and Humanities (the area related to this current work), James, Sotto & Fox (2004) found that the Arts and Humanities

performed worse on the use of long sentences, implying that the Arts and Humanities, on the average, used long sentences their publications. Hence, RAs from the Arts and Humanities were concluded to be the least readable of all research articles. It seems that the writing of long sentences is the pattern in the Arts and Humanities. This is attested to by an earlier work by Bazerman (1981), which was replicated and enhanced in a series of works by James, Sotto, and Fox (2004). They did this by employing a computer-based style program (Microsoft's Office '97) to assess the mean sentence lengths (in words), the percentage of passive sentences, and a Flesch Reading Ease score for each extract. It was found out that Hartman's 1978 humanities article, had the longest sentences. It is noted that the sample size used in Bazerman's study was small (since he selected just one article each from the three disciplines) which could bias the findings and corresponding generalisation of their work. Noting this limitation, James, Sotto and Fox (2004) replicated the work of Bazerman, and enlarged the sample size by selecting 30 articles each from the three disciplines. Interestingly, the findings showed that the Arts and Humanities scored best in all three areas under consideration: sentence length ($M = 32.3$; $SD = 13.4$), passiveness ($M = 8.3$ $SD = 11.8$), and readability ($M = 25.4$; $SD = 14.2$). Although this finding seemed to have contradicted earlier works by Bazerman (1981), James, Sotto and Fox, (2004) provided a sound explanation to the possible cause of this seemingly contradictory results. They noted that the authors in the *American Historical Review* used the numbered footnote method of referencing, whereas authors in the *American Journal of Science* and in the *American Journal of Sociology* use the Harvard (name-date) system of referencing. This meant that any article that used the Harvard

system to present long lists of references in the text (as many did) would be bound to score worse on a measure of sentence length than would similar articles using footnotes. However, because names and dates in the text often consist of only two or three short words, the overall Flesch measure is not greatly affected. These reasons therefore accounted for the seemingly improved readability of RAs in the Arts and Humanities.

In the same study by James, Sotto and Fox, (2004), chapters written in three similar textbooks for colleagues in the sciences, social sciences, and arts and humanities were examined. Using 20 passages from the sciences and Arts and Humanities (and 17 for the Social Sciences) it was found out that the arts and humanities texts consisted of the lengthiest sentences. The mean scores on the Flesch measure, showed that the texts from the Arts and Humanities seemed to be the lowest of the three disciplines. The difference in readability was not statistically significant, thereby giving only partial support to the findings of earlier works by both Bazerman (1981) and James, Sotto and Fox, (2004).

In two other findings from the series of studies by James, Sotto, & Fox, (2004) drew similar conclusions was drawn to what has been reviewed above. This gives additional support that the Arts and Humanities writes very difficult to read research articles.

Other researchers have also used abstracts of research papers to evaluate the readability of research articles. But some took into consideration the apparent biases that are introduced in the analysis and conclusions by using only the abstract for the assessment. Hence, they limited the interpretation of the results to the abstracts alone and did not extend the

implications to the entire RAs. For example, Gazni (2011) examined the abstracts of articles of the five most cited institutions in the world (Harvard University, Johns Hopkins University, Stanford University and Washington University as well as the Max Planck Institution) to determine their text reading level. Around 260,000 articles published between 2000-2009 were analysed and the Flesch reading ease (RE) formula was applied to calculate the difficulty level of the abstracts according to the readability scores. The results revealed that the mean RE scores across the institutions ranged between 15 and 19.

Hence, the texts of the abstracts were all classified as *very difficult* to read. Paralleling the study of Gazni (2011), Severance and Cohen (2015) measured how the readability of abstracts of research articles of medical research journals changed from 1960 to 2010. Employing the Coleman-Liau Index (CLI) readability score, the mean CLI score across all the years was between 16 and 17. This suggested that all the abstracts were very difficult to read. It is noted that these authors used a very large sample size (close to 7000000 abstracts) which gave confidence to their findings.

Similar to the use of abstracts to evaluate the readability of research articles, some authors have analysed other sections other than the abstracts. For example, Bauerly, Johnson, and Singh (2006) evaluated the readability of *Journal of Marketing* using Flesch reading ease formula. Sample of text from the introductions of the first five articles in the first issue of the journal in each five-year period, starting from 1936 and ending in 2001 were used to assess the readability of the research articles. In all, 70 separate texts were used for the analysis. It was found that the samples scored FRE ranging from 12.6 to

41.2, with an overall mean of 27.1 ('very difficult'). Their justification of using only texts from the introduction sections of each selected article was that the introduction section avoids the use of the most technical language.

However, the authors did not define what they meant by 'technical language'. If the term 'technical language' refers to the use of mathematical formulae, then their justification will be true. On the other hand, if the term 'technical language' referred to terminologies peculiar to a field of study other than mathematical algorithms, then the reason for using only texts from the introductory section is somewhat questionable. This is because the use of technical language directly influences the readability of the research paper. Yet, texts from just the introductory section, seems to me, to not likely be a fair representation of the entire research article. Nevertheless, their findings parallels Dolnicar and Chapple's (2014) work, and further strengthens the conclusion that research articles are difficult to read. Using a similar approach, Lee and French (2011) investigated the readability of academic papers in the *Journal of Property Investment & Finance (JPIF)*. Texts from the introductions of all articles in all the issues of *JPIF* from 1997 to 2009 (Volume 17 to Volume 27), which resulted in a total of 297 valid sample size, were used for the study. Unlike Bauerly, Johnson, & Singh (2006), Lee and French (2011) employed five different readability indexes for their analysis. Their findings showed that Flesch reading ease scores range from -2 to 70, with an overall mean of 30.4 ("very difficult") and with three quarters (75%) of the papers having FRE scores in the "very difficult to difficult" to read range, i.e. scores of 50 and less. It is noted that Lee and French (2011) took extra care to ensure that their conclusions were accurate. This they did by first

eliminating sample research articles which had introductory text of less than 100 words. This was necessary because texts which consist of less than 100 words produced negative FRE scores. This stems from the regression equation of the FRE index. Nevertheless, some negative scores were still recorded which was quite surprising. In addition, despite the effort to ensure accuracy in the conclusions they drew, Lee and French (2011) still made generalisation of the readability of the introduction sections they used to the entire research articles. In all, they conceded that it was possible that the sample that generated the findings in their study could have been unrepresentative in some systematic way of the entire research articles.

Some authors have noted the shortfall of using just a single section of a research paper to determine the readability of that entire paper. Some have therefore, used a combination of texts from different sections of the same research article for their assessment. For example, in the field of Medicine, Hall (2005) evaluated the readability of research articles in medical journals. Flesch scores for papers published in the *Archives of Surgery*, the *British Journal of Surgery*, and the *ANZ Journal of Surgery* were calculated. The first 30 original articles published in each journal in 2005 were selected for study. Hall found that the overall median score was 15.1 (0.0 – 29.1). The median scores for the journals were 12.4 (*Archives of Surgery*), 14.4 (*British Journal of Surgery*), and 18.6 (*ANZ Journal of Surgery*). Hence, the conclusion is that the research articles were very difficult to read. Using texts from different sections of the same research paper to evaluate the paper's readability is likely to produce a fair and balanced assessment of the entire article.

In fact, using texts from all sections of an RA to evaluate the article's readability is the recommended approach to measuring readability which can be generalised to the entire RA (Fry, 1977). Thus, as an improvement to the approach of Hall (2005), Hayden (2008) considered the readability of *British Journal of Surgery*. Manuscripts of 189 articles published in the *British Journal of Surgery* from March 2006 to April 2007. Flesch reading ease scores were calculated for each article. For each article, the words from the introduction, methods, results and discussion were used in the assessment. The mean FRE score was found to be 20.6, implying that the research articles were generally very difficult to read. In addition, it was found that the mean readability score was lower when English was the first language of the principal author ($P = 0.016$) but there was no significant difference in mean Flesch score between the submitted and accepted (peer reviewed) manuscripts. But a significant increase occurred after editing ($P < 0.001$). Similarly, in the field of Marketing, Sawyer, Laran and Xu (2008) studied readability of articles in four marketing journals: *Journal of Marketing (JM)*, *Journal of Marketing Research (JMR)*, *Journal of International Marketing (JIM)*, and *Journal of Public Policy & Marketing (JPP&M)*. Using texts from the entire research articles of 162 papers, they found that the average readability score across all articles was 35.3, which is considered difficult for people to read.

In certain studies, that evaluated the readability of research articles, the authors did not clearly indicate which sections of the research article they sampled the texts they used for the analysis from. For example, in a comparative study of nine popular journals in the area of information systems (IS) Otto et al., (2010) evaluated the readability of twenty articles that were

randomly selected from each of the nine journals for the year 2010. Using the Flesch – Kincaid grade level index, the readability score of the nine journals ranged between 11 and 17, which corresponds to difficult and very difficult to read text. The authors, however, failed to indicate which section of the research articles they sampled the text they used for the analysis. Hence, no fair critique could be offered.

Although the vast majority of all extant research articles on readability of research articles have largely reported the average readability scores to help determine the readability of the research articles evaluated, some studies have only been comparative in nature without indicating the level of readability of the texts. For example, Shelley and Schuh (2001) examined the relationship between writing quality, readability, and selectivity of seventeen education journals. But, they failed to provide the average readability of any of the texts they used.

Considering that almost all research articles that have been evaluated to determine their readability have all been indicated to be difficult to read, some academicians have theorised that scientists gain prestige by writing in a manner that is difficult to read. This theory is said to be a motivating force that may have driven the difficult manner in which almost all research articles have been written. This theory has been termed the bafflegab theory (Armstrong, 1988). To be sure, Armstrong (1988) reviewed studies that have been made of relationships between a journal's readability and its prestige. Using ten experimental studies which they

- i. manipulated the textual difficulty of extracts from journal articles and examined their prestige ratings, or

Comment [P15]: ?

Comment [P16]: ?

ii. calculated correlation coefficients between readability measures of journal articles and the journals' prestige rankings;

they concluded that, there is little evidence to support the notion held by some that prestige is inversely related to reading difficulty.

A limitation of Armstrong's (1988) study is that the sample size used (10) was small. In addition, the sampling approach was not indicated. Therefore, it will be inappropriate to generalise his findings to all research articles and conclude that there is there is little evidence to support the notion held by some that prestige is inversely related to reading difficulty.

In more recent times, some authors have reported empirical findings of text difficulty by using analytical procedures based on cognitive theories rather than classical readability theories. For example, using principles developed based on Kintsch's model of text comprehension (Kintsch and van Dijk 1978), Britton and Gulgoz (1991) developed principles for revising texts at a local level to improve comprehension. They took a text that was used to train Air Force recruits and used Kintsch's computer program (Miller and Kintsch 1980) to find places in the text where inferences were lacking. Having devised modification principles based on Kintsch's theory, they modified the text by linking each sentence to the previous sentence via overlapping propositions and arguments using only one term for each concept that appeared in the text, arranging sentences so that old information precedes new information, and making important implicit inferences explicit for the reader. The authors found that participants performed better on free recall tasks and multiple-choice inference questions when given the revised version rather than the original version of the text even though traditional readability statistics

(e.g., Flesch–Kincaid, Coleman–Liau, and Automated Readability Index) between the passages were the same. Additionally, when the novices' ratings of relationships among terms were compared to experts' ratings (experts had read only the original text), ratings of novices who had read the revised text correlated much higher with the experts than ratings of novices who had read the original text.

Britton et al. (1993) later found similar results when they conducted a review of studies in which textbooks, had been revised according to similar principles, providing a promising contrast to studies in which revisions made according to readability formulas had little effect (e.g., Coleman 1962; Klare 1963). These studies by Britton and colleagues demonstrate that even if readability formulas are not able to discern differences between texts, analyses of explicit inferences within a text can show that one text is more comprehensible—at least for novices—than another.

The findings from these two studies shed light on the inadequacy of classical readability formulas to measure text difficulty. However, these findings do not suggest that classical readability indexes are inappropriate tool to be used to measure texts difficulty. It only indicates that some extra factors that influence texts readability, are not captured by the traditional readability indexes.

Similarly, Crosier (2004) considers the readability of 475 RAs from 14 English-language journals in the field of marketing, by using texts from all sections of the article (except the abstract). As an earth science-turned-marketing professional, Crosier was motivated by the potential effectiveness of typical marketing journals in transferring useful learning from academics to

practitioners. The articles used were published in 2013, and were from the following journals:

Admap (35 articles), *Australasian Marketing Journal* (16 articles), *European Journal of Marketing* (54 articles), *Harvard Business Review* (93 articles), *Industrial Marketing Management* (66 articles), *International Marketing Review* (29 articles), *Journal of Brand Management* (23 articles), *Journal of Marketing* (32), *Journal of Marketing Management* (44), *Journal of Product & Brand Management* (29 articles), *Marketing Bulletin* (3 articles), *Marketing Management* (19 articles), *Singapore Management Review* (10 articles), *Sloan Management Review* (28 articles).

Fleisch Reading Ease (FRE) was employed to analyse the readability of the RAs. The finding from this work was that 472 out of 475 articles scored a mean FRE of 36, denoting that the RAs were overwhelmingly difficult to read. ANOVA showed significant differences in the average FRE scores ($F_{1/4} = 53.76$; $df = 13$; $p = 0.001$) across journals. The *Journal of Product & Brand Management* contained the easiest to read RAs ($M = 57.0$) while *International Marketing Review* RAs were the most difficult to read ($M = 10.7$). Although the other took extra care to get quite 'substantial' sample size which was representative of the field of management, her analysis had a few draw backs. First, in comparing the readability of the RAs across journals, Bonferroni's post hoc test was employed, even though the sample sizes of RAs across journals were different. Under such circumstance (when the sample size differs), either the Gabriel's procedure or the Hochberg's GT2 test gives the best comparison because these two exhibit greater power. Using Bonferroni's test for different sample sizes raises the type error. Secondly, the author failed

to indicate whether or not a test of homogeneity of variance was performed prior to the ANOVA. The test of homogeneity of variance is critical to ANOVA. Homogeneity of variance is the assumption that the spread of scores is roughly equal in different groups of cases, or more generally that the spread of scores is roughly equal at different points on the predictor variable. A violation of this assumption could bias the findings of this author's work, and therefore invalidates it (Field, 2011).

In 2016, Lei and Yan analysed the readability of abstracts and full texts of the articles published in four journals of information science from 2003 to 2012. The authors followed the earlier works of other researchers who considered readability of RAs in other disciplines, and sought to fill the gap for the case of Information Science, since no earlier research(es) had been reported in this field. Articles from four journals (*Scientometrics*, *Journal of Informetrics*, *Research Policy*, and *Research Evaluation*) were used for the analysis. These journals were selected because they are considered to be important journals in Information Science. The time span of articles selected for this work was informed by two reasons. First, the authors considered that a time span of 10 years was enough (and current) to serve as a time window for them to discern the possible change of the difficulty level of the articles.

In addition, no articles published after 2012 was included because the span from 2012 to 2015 would not leave enough time for the publications to accumulate citations. Using FRE and SMOG formulas, the authors found that the abstracts were very difficult to read in terms of readability indices such as FRE and SMOG. The results also showed that some of the readabilities of the abstracts and full texts changed in the examined decade, though the effect

sizes were minuscule. Meanwhile, the readability scores were not significantly correlated with the number of citations. It is noted in this study that the span of years considered for the study was long enough to observe practical changes in the readability of the RAs in these four journals over time. However, two shortfalls to the approach could be highlighted. First, the sample size used for the work was rather too small. Although a total sample size of 200 may seem 'large', this figure is deceptive. For each year, journals could publish over 20 articles per issue. By this assumption, a total of at least 100 articles per journal per year, is published. Therefore, selecting just 5 articles for each journal per year is woefully inadequate. This point is crucial because the authors generalized their findings to the entire set of journals for the 10-year period they investigated. When sample sizes are that small, representativeness is not achieved, hence, the results may not be accurate. Secondly, the authors failed to indicate their criteria for selecting each of the five set of articles for each journal. Thus, a replication of their work will be extremely difficult.

Libraries are hosts of knowledge in the form of books, journals, and several other academic writings. Hence, Librarians are very much interested in the quality of academic writing, including the readability of the texts in the books to be hosted in any library. Therefore, some librarians have been interested in evaluating the readability of RAs in journals published by certain library associations. For example, Metoyer-Duran (1991) evaluated the readability of published, accepted, and rejected papers appearing in *College & Research Libraries*. She was convinced that during this period where there is proliferation in the sources of information RAs written poorly (that is RAs that are not readable) may not be published in the long run. Therefore, her

objective was to evaluate the readability of RAs published in the *College & Research Libraries*, and to determine whether readability has changed over time, and in what direction. She hypothesized that the texts of published RAs were likely to be more readable than those accepted or rejected. Texts used for this study was obtained from the editor of *College & Research Libraries* and included copies of all manuscripts accepted and rejected during 1990 and 1991, excluding the names of the authors and associated editorial correspondence. In all, 82 refereed papers appeared in print, 70 papers were accepted but not yet published, and 119 were rejected. This author demonstrated strong statistical ability by employing all the principles of proper sample and sampling techniques, and thus inspires confidence in her findings. For example, given the hypotheses and the large size of a sample necessary to achieve a precision of ± 5 , with 95 percent confidence, the investigator examined all 271 papers and did not draw a sample. In addition, the author noted the need for representativeness by randomly sampling paragraphs, including the first and final paragraph of each paper. The investigator numbered the unique paragraphs in each paper which aided in the selection of which paragraph to be included in the sample. Her findings were that both the text and abstracts are "difficult" to "very difficult" to read. In addition, rejected RAs were found to be more readable than the published or accepted RAs. Yet, for abstracts of each of these three groups, there were no statistically significant difference in the readability. In all, this author demonstrated (among all RAs reviewed in this work) profound knowledge of research design and statistics and was apt in her applications of the laws governing research design, statistics, and generalization of findings.

Earlier, Whatley, Ray and Reidy (1973) applied the approach of using texts from the entire RA to evaluate the readability of 10 selected management journals. The journals were *Academy of Management*, *Journal of Administrative Science Quarterly*, *Advanced Management Journal*, *Business Horizons*, *California Management Review*, *Industrial Relations*, *Personnel*, *Supervisory Management*, and *Systems & Procedures Journal*. Whatley et al. employed FRE index to evaluate the readability of these 10 journals. Their findings were that the reading ease scores for eight of the ten journals decreased from 1967 to 1971, indicating that those journals became harder to read. Also, reading ease score averages for the 1971 versions of the journals combined was 3.67 points lower than the 1967 combined journal average. Three of the ten journals showed significant changes in readability from 1967 to 1971. In all cases, the journals were found to be difficult to read. It must be noted that these authors failed to indicate how the texts used for the analysis were sampled. In fact, the paper is silent on how the RAs were obtained. In addition, the entire paper cites only two authors, which is woefully unsatisfactory for an academic writing of this nature. The failure to indicate the sample and sampling procedure, coupled with the limited references cited, may raise concerns on the accuracy of the findings in this paper, and makes replication of their study virtually impossible.

It has been noted in this review that all research works that have been reviewed are comparative in nature. The issue of readability has been treated, in all cases, as a standalone. That is, authors evaluate readability of RAs and compare to earlier works in order to put their own findings in context. One author, though, deviated from this approach, and considered readability of

RAs as one determinant (among several others) of the visibility of an RA. Inspired by the logic that the extent to which the field pays attention to what an academic publishes (through the number of citations received for a publication) determines the rest of his or her academic career, Stremersch, Verniers and Verhoef (2007) evaluated the factors that influence whether a given paper, therefore a given scholar, is cited. The FRE was used to compute the readability of the RAs used in this study. It was found that readability was negatively correlated to citation, and this correlation was statistically significant. In other words, readability of RAs was found to be a significant determinant of citation, but the effect on citation is negative such that more readable RAs were cited less as compared to RAs which were less readable. This work was very extensive with respect to the number of determinants used. The comprehensibility of the work however, seemed to have dwarfed the thoroughness with which the readability aspect should have obtained. For example, although the authors documented the use of FRE in their analyses, the paper is silent on how texts were sampled. Nevertheless, the paper seems satisfactory (when the readability component alone is considered).

In a rare evaluation of readability of RAs in the field of textile and clothing, Oliver, Dallas and Eckman (1998) considered Flesch scores to measure the readability of passages taken from published research. Their work was inspired by the lack of earlier research on readability in the field of textile and clothing, although several other aspects of published research have been considered by authors in this field. Three objectives were set to evaluate the readability of RAs by these authors:

- ❖ To investigate, through objective and subjective testing, the readability of research utilized by and representative of subject interests of ITAA members.
- ❖ To investigate, through subjective testing, the perceptions of quality of research and journals utilized by and representative of subject interests of ITAA members.
- ❖ To interpret how readability factors influence perceptions of research and journal quality.

Included in this study were journals identified as publication sites for articles related to textiles and apparel subject areas (Oliver & Kuruvilla, 1992). Among these 23 journals were ten journals cited most often in *Clothing and Textiles Research Journal* (Oliver & Mahoney, 1991). The remainder included the *Clothing and Textiles Research Journal* and 12 journals in which ITAA members published most often, according to publication listings in the ITAA Newsletter (formerly ACPTC Newsletter). All 23 were academic journals rather than popular or trade publications. The researchers collected all issues from the previous five years of each of the 23 journals and cataloged articles specific to the eight textiles and apparel subject areas as listed in various ITAA membership documents. Operational definitions of subject matter areas were developed by the researchers to classify articles. The purpose and theoretical perspectives of each article were examined by all authors who agreed on classification of subject matter. In each journal, articles were counted to determine distribution by subject area. Based on the distribution, a stratified random sample of five articles was taken from each journal for a total of 115 articles. These 115 articles served as the population

of textiles and apparel research for instrument development. Eight 100-plus word passages were selected from each article as suggested by Dobler and Gilroy (1987). Passages from the first paragraph of each section of the article (introduction, methods, results, and conclusions) and four additional passages approximately halfway through each section comprised the writing sample. Passages were scanned, filed by article on a computer disk, and reexamined to correct any errors from the scanning process. Flesch scores for selected passages ranged from very difficult (*Journal of Marketing Research*, 18.39) to a standard readability level (*Canadian Home Economics Journal*, 68.04). This range of readability is consistent across publications in most academic fields (Armstrong, 1980; Clark & Geisler, 1986; Loveland, et al., 1973). Respondents' subjective evaluation of readability clustered around 3.5, indicating moderate readability. No passages were perceived as very difficult to read, although some were perceived as somewhat easy to read. A non-parametric test could not be identified to compare objective and subjective ratings of readability.

However, within each subject area, rank order of Flesch scores did not correspond with rank order of respondents' perceptions of readability. Armstrong (1980) had similar results. Although Flesch scores are valid quantitative measures, they cannot capture the holistic evaluation of readability. For example, an individual's interest in a specific subject area may influence familiarity with research and, therefore, perception of readability. Correlation coefficients were calculated to analyze the relationship between perceived readability and perceived quality of research. Readability ratings were combined across all subject matter areas and compared with ratings of

research quality. A significant ($p < .01$) relationship between the two concepts was found for all three passages. The more readable the research, the higher the perceived quality. It is noted that the findings of this research contrasted the earlier findings of almost all published works which reported negative correlations between quality of RAs and readability.

Finally, Bottle, Rennie, Russ, and Sardar (2015) scrutinized the readability of research articles in multiple disciplines. However, their article appears so disjointed such that it cannot be determined whether the findings reported in the paper were from their own research or it was the report of some other authors (as if they were reviewing literature). Hence, it is difficult to review this paper thoroughly.

In all, one pattern has been demonstrated in all the papers reviewed so far: that RAs are difficult to read when considered in the light of readability indexes.

Readability of pedagogical materials

Comment [P17]:

Berndt and Wayland, (2014) have listed nine different studies on the readability of textbooks in the field of business studies. In all, the general conclusion was that the textbooks were difficult to read. Owu-Ewie (2014) considered the readability of comprehension passages in Junior High School (JHS) English textbooks in Ghana. Using the Gunning FOG Readability test, Flesch Reading Ease Formula, Flesch-Kincaid Grade Level, SMOG Index, Coleman-Liau and Automated Readability Index, it was found that most of the passages were above the age of learners and were therefore difficult for them to read and comprehend. Similar methods and conclusions have been drawn

by others in different fields including the sciences, mathematics, and economics from different countries (see Barret, 2014; Barrett, Mtana, Osaki, & Rubagumya, 2014; EdQual, 2010; Li, 2011; MacAllister & Duckworth, 1981; O'tkeeffe & O'donoghue, 2014; Percorari, Shaw, Malmstrom, & Irvine, 2011; Pertekeme & Agbor, 2012; Tinkler & Woods, 2013; Toriki, 2012). The conclusion from these numerous research is that most textbooks are much difficult to read than the level of the intended readers. The findings of Owu-Ewie (2014) throws a spotlight on the situation in Ghana, suggesting the need for further research in the readability of other textbooks and manuals of other schooling levels. This study attempts to bridge the gap in information on readability of RAs from the Ghanaian context.

Readability of health materials

Several literatures have been published in the field of readability of healthcare materials. Several of such studies have focused on specific disease conditions. For example, some have considered the readability of HIV/AIDS. Considering that, Buccini et al., (2010) measured the readability of Australian based informed consent documents employing the SMOG and Fog indexes. Using a total of 200 consent documents, the HIV/AIDS informed consent documents were, on average, written at a grade 13 reading level. This was above recommended reading age of 8. Earlier, Wells (1994) evaluated the readability in 136 HIV/AIDS educational items using the SMOG Index. The result was similar to the findings of Buccini et al., (2010).

Other authors have considered the readability of health materials on other diseases. For example, Barbosa and Martins (2007) evaluated the

Comment [SKG18]: Review more of this

Comment [P19]:

readability of 49 internet information about floaters and light flashes found by two search engines, MetaCrawler and MSN. Using FKGL and quality component scoring system, the Mean FKGL was found to be 9.9 which suggested that the internet information and floaters were fairly difficult to read. Although the FKGL has been used as a popular index for analysing readability written texts both in academia and other fields, it has been suggested that the SMOG index works best for health literature. Hence, Barbosa and Martins work did not use the best readability index for their evaluation. On the other hand, Brown, et al. (2004) used the SMOG index to consider the readability of Cataract information leaflets from 12 ophthalmology departments in England. The mean SMOG score was found to be 10, which suggested that those materials were fairly difficult to read.

In 1997, Ebrahimzadeh, et al considered the readability of 22 health educational brochures from the AAO using the Flesch reading ease and grade level, as well as the gunning for index. It was revealed in that study that 32 % of all those educational materials were at or below an 8th-grade reading level, 55 % between 8th- and 10th-grade levels, and 15 % were a 10th-grade reading level or higher. Edmunds, et al. (2013) also used a combination of a number of readability indexes as was the case with Ebrahimzadeh et al (1997). Using a combination of the FRE, FGKL, Gunning fog index, and the SMOG index to evaluate the readability of the top 10 patient-oriented websites for 16 different ophthalmic diagnoses, these authors found the mean FKGL = 11.3. They concluded that regardless of the index used, readability scores were poorer than those recommended. Just a year after this study, Edmunds, et al. (2014) evaluated the readability of 50 online literature specifically for Graves' disease

and thyroid-associated ophthalmopathy through Google search. Using FKGL, FRES, SMOG, and GFI indexes, the mean FKGL was found to be 11 which denotes fairly difficult to read materials. In the same year, Hansberry, et al conducted a study that appeared to be a replication of Ebrahimzadeh et al (1997) work. These authors used information from the same Patient education material on AAO website, as was the case of Ebrahimzadeh et al (1997). The difference with the work of the earlier authors was the unspecified number of educational materials used and the use of many readability indexes (FKGL, FRES, SMOG, GFI, Coleman-Liau Index, the New Fog Count Formula, the New Dale-Chall Readability Formula, FORCAST formula, Raygor Readability Estimate, and the Fry Graph). The mean reading age recorded was 11.7, indicating fairly difficult to read texts. Although not indicating any shortfalls of the works of either Edmunds et al (2014) or Ebrahimzadeh et al. (1997), Huang et al. (2015) recently replicated and expanded the works of these two earlier groups of researchers. By considering the readability of 339 Websites from 7 ophthalmologic organizations (AAO, American Association of Ophthalmic Oncologists and Pathologists, AAPOS, AGS, ASCRS, ASOPRS, American Society of Retina Specialists, American Uveitis Society, Cornea Society, and NANOS) using numerous readability indexes, (FKGL, FRES, SMOG, GFI, Coleman-Liau Index, New Fog Count, New Dale-Chall Readability Formula score, FORCAST score, Raygor Readability Estimate Graph score, and Fry Readability Graph score), Huang and his colleagues found the mean reading age to be between 10.4 to 12.6. This suggested readability level classified as fairly difficult to read and corroborates the

findings of those earlier authors whose work Huang and his colleagues mimicked.

Still on the readability of ophthalmology educational materials, John, et al. (2015) considered the readability of the first 10 PEMs to appear in search on Google search for 10 paediatric ophthalmology conditions. Using FKGL, FRES, SMOG, GFI, Coleman-Liau Index, New Dale-Chall, FORCAST Formula, Fry Graph, Raygor Reading, the mean reading age was found to be 11.75. Hence these materials were largely fairly difficult to read. Only 12 % of articles were written below a 9th-grade level and only 3 % met recommended criteria. Similar findings have been recorded by Khurana et al (2003), Martins and Morse (2005), Muir and Lee (2010) and Zaidi and Jones (2009).

In other healthcare materials Corcoran and Ahmad, (2016) investigated the readability and suitability of sexual health promotion leaflets using the SMOG and Fry indexes. These authors found that the leaflets were at an average reading level of grade 9, which was above recommended reading level of 6-8. Paralleling this finding, Hadden et al., (2016) evaluated the readability of patient education materials in hand surgery using the multiple readability tools. Although they found improvement in readability over earlier studies, the readability was still above recommended reading grade levels. Similarly, Suleiman et al (2016) appraised the readability of educational materials that support parent sexual communication with their children. Fifty brochures, pamphlets, and booklets were analyzed using the Flesch-Kincaid, Gunning Fog, and Simple Measure of Gobbledygook (SMOG) index methods. Mean readability grade-level scores were found to be 8.3 (range = 4.5–12.8), 9.7 (range = 5.5–14.9), and 10.1 (range = 6.7–13.9), respectively. These analyses

indicated that the majority of educational materials available online to support parents' communication with their children about sex and sexuality do not meet the needs of many or most parents. Efforts to improve the accessibility of these materials are warranted.

Again, in sexual health, Carol and Hosei (2008) examined the readability of printed health information materials collected from multiple sources. Using 21 materials as the sample size and employing the SMOG readability index, it was found out that 53% to 86% of the printed materials had a reading level at or higher than 9th grade.

The field of rhinology has also seen literature on readability of educational materials. Kasabwala et al (2013) evaluated the readability of online patient education information available from the American Rhinologic Society (ARS) website using 9 different assessment tools that analyze the materials for reading ease and grade level of the target audience. These indexes were Flesch Reading Ease, Flesch- Kincaid Grade Level, Simple Measure of Gobbledygook (SMOG) Grading, Coleman-Liau Index, Gunning-Fog Index, FORCAST formula, Raygor Readability Estimate, the Fry Graph, and the New Dale-Chall Readability Formula. The findings from this study was that all the healthcare education materials assessed were written between a 9th grade and graduate reading level and were considered "difficult" to read by the assessment scales.

Some authors have used systematic review approach in analysing the readability of health education materials. For example, Morony, Flynn, Mccaffery, Jansen, and Webster (2015) reviewed the readability of written materials for chronic kidney disease (CKD) patients. Analysing 80 materials

using the Lexile Analyzer and Flesch-Kincaid tools, the results showed that most materials required a minimum of grade 9 (age 14-15 years) schooling to read them. Only 5% of materials were pitched at the recommended level (grade 5).

Extensive study has been conducted in assessing the readability of malaria medicines information leaflets. For example, Auta, Shalkur, Banwat and Dayom (2011) collected data based on forty-five leaflets obtained from community pharmacies in Jos, Nigeria. The researchers used only leaflets in respect of antimalarial medicines of artemisinin-based preparation and those registered by Nigeria National Agency for Food, Drug Administration and Control (NAFDAC). In all, 68 malaria medicine information leaflets were collected and 45 selected via lottery method. SMOG readability formula was applied in checking the readability of the selected malaria medicine information leaflets. The findings revealed that malaria information leaflets scored SMOG value from 9 to 16. In addition, several were having a reading grade level of 14. Hence, Auta et. al. (2011) concluded that malaria medicine information leaflets in Nigerian health sector are not readable to an average reader in Nigeria.

Jabbari and Saghari (2011) conducted a comparative study to determine the readability of English medical texts and their Persian translations using 50 translated booklets and their corresponding texts in English. Gunning Fog Index and SMOG Readability Index were used to measure the level of readability. Original texts were drawn from the British Medical Association (BMA), while the corresponding Persian Translations were translated by different translators and publications kept in National

Library of Iran. Findings of the study revealed that the number of words and multi-syllable words in English medical texts and their corresponding Persian translations were not normally distributed. Gunning Fog and SMOG scores indicated that the number of sentences in English medical texts and their corresponding Persian translations were distributed normally. Results emanating from a t-test based on the readability scores, revealed that the difference between the number of sentences in English medical texts and their Persian translations turned out not to be significant ($P = 0.835$). It was also discovered that the Persian translation had a better readability score than its English counterparts due to the significant difference in the number of multi-syllable words between the two texts.

As technology keeps growing, more and more individuals have gotten used to fishing for healthcare information from the internet. Hence, a number of researchers have evaluated the readability of online healthcare materials. For example, Adam et al., (2015) evaluated whether the American Orthopaedic Society for Sports Medicine (AOSSM) website's patient education materials meet recommended readability guidelines for medical information. Using Flesch-Kincaid readability index, these authors found that the average readability of all 65 articles was of grade level of 10, which was above recommended readability of sixth grade level. Paralleling this, Cortez, Milbrandt, Kaphingst, James, and Colditz, (2015) assessed the overall readability of breast cancer risk assessment tools and accompanying information on the internet. The average SMOG reading grade level was grade was found to be 12.1 (SD 1.6, range 9–15), suggesting that the materials were generally difficult to read.

Further, Falconer et al (2011) analysed fourteen educational brochures on the APTA website using the following assessments: Flesch-Kincaid Grade Level, Flesch Reading Ease, Fry Readability Formula, Simple Measure of Gobbledygook (SMOG), Checklist for Patient Education Materials, and Consumer Health Web Site Evaluation Checklist. Findings of the study showed that Flesch-Kincaid and Flesch Reading Ease score for over 90% of the brochures were written at greater than a sixth-grade level. The mean reading level was grade 10.2 (range= 3.1 to 12), with a Reading Ease score ranging from 31.5 to 79.9. SMOG formula revealed that the brochures scored a mean reading level of grade 11.5 (range= 9 to 13). The Fry Readability formula showed that 85% of the brochures were written higher than a sixth-grade level, with a mean reading level of grade 9.5 (range = 6 to 14).

Similar findings using Flesch indexes have been reported, all indicating corpora of text that are above the readability of the target population (see AlKhalili et al., 2015; Cheng & Dunn, 2015; Evans et al., 2016; McInnes & Haglund, 2011).

Numerous other findings in the health field have been documented including paediatric health, mental health, orthopaedics, and oral health.

Readability in the financial sector

Comment [P20]:

In several countries, it is required by law and ordinances that public companies provide an annual report to their investors. Several efforts, via regulations, have been made by the Security and Exchange Commission (SEC) to make these disclosures less complex and free of incomprehensible language. However, Li (2008) contended that despite these regulations,

companies could use vague language and format in disclosure to hide adverse information, and average investors may be unable to understand these disclosures leading to capital market inefficiency.

In order to ascertain whether firms use vague and incomprehensible language for disclosures, Kumar (2014) conducted extensive investigation to address the following issues: 1) the effect of domestic culture, i.e secrecy; 2) the effect of agency theory, i.e ownership dispersion, and 3) the effect of profitability on the readability of annual reports of U.S. – listed Asian companies. Data gathering samples were drawn from all U. S. – listed Asian companies in the year 2010. However, there are 85 companies (registered and reporting with the SEC) from nine Asian countries which were selected; only 68 forms the focus of the study because they are listed on the NYSE (New York Stock Exchange) / NASDAQ (National Association of Securities Dealers Automated Quotation).

Three hypotheses were tested to guide the study as stated below:

H₁: U. S. – listed Asian Companies with low/ (high) secrecy are likely to provide more/(less) readable annual reports in the U.S.

H₂: U. S. – listed Asian Companies with higher / (lower) ownership dispersion will provide more/(less) readable annual reports.

H₃: U. S. – listed Asian Companies with higher / (lower) profitability are likely to provide more/(less) readable annual reports in the U.S.

The researcher applied Flesch Reading Ease (FRE) as calculated in MS word, based on readability of the Management Discussion and Analysis section of the sample companies' Form 20F annual report. OLS regression was used by the researcher to examine the effect of secrecy, ownership

dispersion, and profitability. Findings of the study revealed that companies whose domestic culture is more secretive are providing less readable annual reports. Results also show that companies with higher ownership dispersion are providing readable annual reports. The researcher recorded no significant effect of profitability on the readability of annual reports. Findings of the study failed to reject the hypothesis related to the effect of profitability, though it upheld the rest two hypotheses on the effect of secrecy and ownership dispersion.

The implication of the study, according to Kumar (2014), is very significant in the arena of international accounting literature. This becomes serious because many of these companies are using IFRS (International Financial Reporting Standards) or U.S GAAP (Generally Accepted Accounting Principles) to prepare their financial statements and in spite of using these global standards the domestic culture has an effect on the readability of their annual report. It, therefore, behooves investors of these companies to painstakingly peruse the annual report of these companies.

Additionally, the global standard-setting bodies should consider the effects of culture on financial statements while issuing new standards. Undoubtedly, the impact of poor readability of important document could also be felt in accounting for dire consequences to one's hard-earned money. Likewise, when research article is written with low readability or higher than the grade level of the intended audience, all efforts directed toward its production is wasted and the useful information it contains, becomes useless.

In a bid to explain factors involved in communication gap or breakdown, several researchers, including Chall (1958), Flesch (1949) and Smith and

Smith (1971) tend to view the terms readable and understandable as being very closely related, in fact, some researchers, such as Adelberg (1983) sees the two expressions (readable and understandable) as similar. However, these expressions are far from similar, because according to Krakoff (2011), readable refers to formatting, bullet points, etc., while understandable means taking time to re-read your report as it were, with the eyes of your targeted readers and weed out ambiguity.

In order to ascertain the possibility of a difference in the two expressions, Brennan, Pierce, and Guillamon-Saorin (2009), conducted in-depth study on the literature review of 40 readability studies on annual reports for a period of 59 years (1948 – 2007). Findings emanating from their study revealed unerringly that the reading level of the reports was either difficult or very difficult to understand or comprehend. In other words, most of the executives or directors were affected by communication breakdown; the desired message for their intended user was missed by significant number of readers.

According to Brennan et al. (2009), 25 studies reflected the assumption that readability means understandability. The studies (25) employed Flesch's method of testing for readability with a score of 19. Based on the score, the 25 studies concluded that corporate or institution reports were written with the cold natured academic or scientific style of which an average reader would find it difficult to understand, or very difficult to comprehend. Overall, Brennan et al.'s study revealed executives/directors should compose their report with their targeted audience in mind so that it is clear and understandable.

Soper and Dolphin (1964), Dolphin and Wagley (1977) also conducted an extensive study on annual report in connection with American institution/corporations. The study dwelled on the reports provided by the same establishment in 1961 and 1977. Findings of the study revealed a sharp decline in the firm's corporate report readability. Indicating that much attention was not given to rendering the report understandable to the intended readers, hence it was difficult for an average reader. Another interesting study with ample discoveries was carried out by Lewis, Parker, Pound and Sutcliffe (1986). Among other things, it was discovered in their study, spanning four years that instead of accentuating readability, difficulty or low readability gained ascendancy as indicated by Flesch's readability index. This means most report studied were not fully understood by the targeted readers.

Jones (1988) also conducted extensive study of a particular company in the United Kingdom over a period of 34 years (1952 – 1985). He focused his study on Chairman's narratives and examined the relationship between readability and some company's variables. His in-depth study revealed that readability had declined significantly over the years under review. Moreover, he also discovered that as turnover of the corporation increased over the period under review, readability also declined.

Loughran and McDonald (2010) conducted extensive inquiry into deciphering the readability of company disclosures by applying three different measures to a large sample of 10-Ks within 1994 – 2007. The researchers based their findings on the assumption that better written 10-Ks are more informative to the market. Loughran and McDonald used a readability measure derived from SEC documentation surrounding the plain English

initiative, because when the Fog Index readability formula was used for the text, substantial measurement issues were identified due to business context within which the text is located. The readability measure developed by Loughran and McDonald (2010) captured text informativeness when compared to traditional readability formulas.

Findings of the investigative study by Loughran and McDonald revealed that readable 10K documents produce a greater price impact on their filing date, implying that well written and readable disclosures appear to be more informative, thus improving the financial status of investors and firms. The study also revealed that investors' improved readability has profound impact on investors. It greatly influences investors' behavior, because it produces greater participation by "average" investors. It was also observed that improved readability signals a higher probability of managers issuing seasoned equity. Loughran and McDonald also reported that companies whose management is shareholder-friendly also create 10Ks that are more readable. The textual analysis of 10-Ks document for a period of 13 years (1994 – 2007), clearly demonstrate the power of words, because it can either make or mar investors' responses to a particular firm. As a result, it does much good when relevant company documents are rendered in plain English and with improved readability.

In a study conducted in Malaysia, Curtis and Hassan (2002) examined the reading ease between the English and Chinese versions of 65 corporate annual reports in Hong Kong coupled with the English and Malay versions of 53 annual reports in Malaysia using Flesch and Yang Formulas for Hong Kong and Flesch and Yunus Formulas for Malaysia. Though the results were

not conclusive, it thus provides a tentative impression that the indigenous language version is easier to read than the English-written versions. Findings also revealed that the English passages in Malaysian annual reports are easier to read than the English passages in Hong Kong's annual report.

Abubakar and Ameer (2011) also conducted an impressive study on readability in Malaysia. In their study, they examined the readability of the annual reports on Corporate Social Responsibility (CSR) for a sample of listed companies. Abubakar and Ameer employed readability formula and discovered that the extent of syntactic complexity influences the comprehension of CSR report. In essence, the higher the level of syntactic complexity in a given report of a company under review, the less comprehensible the report would be. They also examined the relationship between readability and companies' performances and findings revealed that the executives or directors of corporations with poor performance often render their CSR report with difficult language which often results in obfuscation for readers.

Readability and insurance policies

In his study entitled *Recalibrating the Flesch Readability Index for the Twenty-first Century*, Stewart (2003) focused on underscoring the fact that the same reading materials Flesch studied in the 1940s, in relation to insurance policies, as indicated in Table 3, do not yield the same results today when analyzed in line with his formula. A modified interpretation of Flesch's scores compatible with the twenty-first century texts was suggested by Stewart. In order to achieve the objectives of the study, the researcher analyzed texts of over 100,000 words, including passages from insurance publications.

Comment [P21]:

Table 3: How to Test Readability

Score	School Level	Example
90 – 100	5 th grade	Comics (92)
80 – 90	6 th grade	Consumer Ads (82)
70 – 80	7 th grade	Movie screen (75)
60 – 70	8 th and 9 th grade	Seventeen (67)
		Reader's Digest (65)
		Sports illustrated (63)
		N. Y. Daily News (60)
50 – 60	10 th to 12 th grade	Atlantic Monthly (57)
		Time (52)
		Newsweek (50)
30 – 50	College	Wall Street Journal (43)
		Harvard Business Review (43)
		New York Times (39)
		NY Review of Books (35)
		Harvard Law Review (32)
0 – 30	College graduate	Auto Insurance Policy (10)

Source: adapted from Harper and Row 1951

The analysis was carried out by means of computer, using software developed by Micro Power and Light Co embedded with a Microsoft Windows application. The computer software analyzed 100-word samples for readability, selected across the publications. In the case where a prose paragraph contained more than 100 words, the sample was extended to the end of the paragraphs, rather than limiting it to 100 words.

Findings revealed that the readability of a *Standard Auto Insurance Policy* scored 59.76 could be considered as an artifact of the works of Flesch. In arriving at the score, Flesch's interpretation (0-30 for insurance policies) was based on his reliance on school grade levels (from fifth grade to college graduate). However, a closer look, according to Stewart reveals a number of inconsistencies that must be resolved in order to achieve an interpretation suitable now – twenty first century. In recent times, insurance policies are being written at a specified level of difficulty, taking into account the reading abilities of their targeted audiences as stipulated by the law.

For instance, according to Stewart (2003), "plain English laws" governing consumer contracts stipulates that insurance policies, that is, consumer contracts, be written in language that is easy to understand rendered in a clear and coherent manner using words with common and everyday meanings. Some of the laws defines "plain language" in terms of the average number of words per sentence, words per paragraph, and syllables per word, and gave detailed instruction on how such specification could be obtained.

As a result, the stipulated range (0-30) indicated by Flesch is no longer necessary for an individual to understand insurance policies. The above stipulated laws have prescribed the level of difficulty and are constantly being scrutinized so that insurance policies are accessible to the public. This discrepancy in Flesch's interpretation as indicated in fig. 1 begs for a more practical method of interpreting the readability scores based on the concept of audience appeal rather than school grade level as observed by Stewart.

Stewart also indicated that several studies have indicated that a correlation exists between reading ability and occupation. Individuals who are

better readers often enter occupations that involve intellectual skills, while poor readers are prone towards doing work that is more physical. In view of the aforementioned, Stewart recalibrated Flesch’s original interpretation in order of their present-day difficulty. The recalibration was based on some different parameters. It was done to correspond to occupation rather than school grade level as depicted in Table 4.

Table 4: Recalibrated Flesch Index

20 th century Level	Flesch Index	21 st Century level	Examples
5 th – 6 th grade (80 – 100)	90 – 100	Elementary school	Peanuts (95) Alice Otter (91) Stuart Little (86)
	80 – 90	Junior High Sch.	Lord of the Rings (84) Comic Books (81)
	70 – 80	High School	Seventeen (77) How to write Plain Eng. (74)
Junior High School (60 – 80)	60 – 70	Blue – collar	Reader’s Digest (70) Sports Illustrated (67) New York Daily News (65) Consumer Ads (62) Standard Insurance Policy (60)
High School	40 – 60	White – collar	Newsweek (58) Time (57) Wall Street Journal (53) Atlantic Monthly (52) Internal Revenue Code (52)
College (50 – 60)			Harvard Business Review (50) New York Review of Books (50) New York Times (48)
College Graduate	20 – 40	Professional (Attorney – at – law, Physicain)	Medical Journals (34) Harvard Law Review (34)
	0 – 20	Statesman	Preambles, Solutions

As shown in Table 4, the new range for insurance policies was interpreted not in terms of what people should read rather in terms of what

they actually read. Hence, there was a centripetal direction in the difficulty of insurance policies and tax laws because they have direct relevance to consumers and their suppliers. A score of 60 is now obtained for the insurance policy which is comparable to an average reader in Junior High School on Flesch's interpretation.

Stewart's study clearly demonstrated that Flesch's Reading Ease scores should be brought into par with current realities.

Readability in information technology

Comment [P22]:

Grandisar, Humar and Turk (2006) conducted an extensive study on factors affecting the readability of coloured text in computer displays. The focus of the study was to examine the readability of different color combinations of text and background, presented in Cathode Ray Tube (CRT) display by measurement of speed of reading. Colours for the study were chosen from non-dithering web-safe colour palette which consists of 216 different colours. These colours comprised elementary colours: 1) white 2) yellow 3) red 4) magenta, 5) blue 6) cyan 7) green and 8) black. All the colours were combined with every other different colour to make the 56 texts/background used for the study.

Participants for the study were selected among the students of the University of Ljubljana. The students were in Level 100 studying Informatics. Though, 300 students were initially selected with 19 (ranging from 18 – 21) as the mean age only 270 respondents were used for the study after a thorough eye examination for visual acuity via Ishihara text for colour blindness and other defects of the eyes.

In order to ascertain or achieve the objectives of the study, the tasks were presented on 21 inches Dell CRT display. The screen resolution was 1280 x 1024 pixels without interpolation and refresh rate was 85 Hz (non-interlaced). After appropriate settings for the study has been achieved, the speed of reading was measured. The subjects were divided into three groups. Each group of students performed experiments by reading 10 color combinations for reference and the Chapman-Cook Speed of Reading tests (slightly adjusted for measurement of readability performance from a computer screen).

Findings of the study revealed that reading is a complex process influenced by many interrelated factors which, according to Grandisar et. al. (2006) cannot be easily identified and controlled. The results also indicated that there are no statistically significant differences in readability between 30 colour combinations. However, the STRESS values for each group of experiments and dimensions setup revealed that there exist at least five factors having influence on the speed of reading. In other words, readability, according to Grandisar et. al. (2006), depends on a mixture of at least five factors influencing readability. Of the five factors, visibility/legibility is the most important common factor influencing readability and it is independent of aesthetic, content, and context of text.

Another impressive, informative and educative study was conducted recently by Kim, Park and Seo (2014). The driving force for the investigative study was to validate the notion that visualizing readability gives intuitive impact on how difficult the texts will be before examining the texts further.

Text visualization, according to Kim et al. focuses on providing structural characteristics of text contents in an efficient way.

Adequately using massive text data, such as books or articles, the study proposed readability measurement factors and formulas for the suggested methods that visualize texts by extracting a key factor 'length' for readability. The study also authenticates the effectiveness of visualization through the text of the case studies. The findings of the case studied illustrates that readers can have readability information not from independent texts, but from the comparison of previous texts, which make it possible to accommodate books with texts that are difficult to read.

Based on the concept of readability formula involving three measurement factors, average number of words, average number of sentences, and average sentence length, Kim et al. (2014) designed readability visualization. In order to accomplish this task, they validate each of the three measurement factors by choosing different texts that have distinctive features and difficulty levels and analyzed the average sentence and paragraph length of each text through graphs. These graphs were based on four books, namely: *Alice's Adventure in Wonderland*, *My Sister's Keeper*, *Ethics*, and *the Critique of Pure Reason*

Finding of the study based on readability visualization algorithm suggested by the researchers, as shown in fig. 3, indicates that when sentences and paragraphs of a book are simple and short, the overall image appears light, and the level of lightness fluctuates with various patterns. This means readability visualization, according to Kim et al., for stories comprising simple

conversations and stories for children have light visualization images, while classics and philosophical or scientific stories have darker images.

In order to ascertain the authenticity of their proposed readability visualization algorithm, Kim et al. (2014), conducted an investigative study into the readability of nine (9) books. The participants of the study comprised 20 graduate students. The researchers investigated the relationship between readability measured from skimming through the text, and the readability predicted from their proposed readability visualization images. The scores were based on a 5-point scale (1: very easy, 5: very difficult).

Findings of the study revealed that 180 pairs of responses indicated a meaningful relationship between readability predicted from reading the actual text, and readability predicted from the readability visualization image. The study also showed that when participants predict or measure the text readability by reading it, it took an extended amount of time, although there were individual variations among the subjects. On the other hand, the rest of the participants measure readability almost intuitively based on the visualized images, the result obtained were as correct as the participants who actually read the text.

It could be seen from the aforementioned study conducted by Kim et al (2014) that since text readability visualization shows visualized images of the whole text, readers could have readability information on the whole text not just on parts of the text. It is possible for readers to compare the readability of a new book to the readability of books they enjoyed reading. It is therefore recommended that these visualized images by Kim et al. (2014) could be instrumental in providing ample information on text readability and different

text characteristics, just like book covers do. It could be observed in all the studies discussed that readability formulas play prominent roles in ascertaining the readability of a written report or research article.

Readability – the Ghanaian context

Comment [P23]:

Considering the Ghanaian context, research works on readability that have been published cover pedagogy (Gyasi, 2013a, 2013c; Owu-Ewie, 2014), newspapers (Fosu, 2014), and medical information manuals (Gyasi, 2013b). All these studies drew similar conclusions, like those reviewed above, that those documents were all difficult to read. For example Owu-Ewie (2014) considered the readability of comprehension passages of Junior High School (JHS) English textbooks in Ghana. Using the Gunning FOG Readability test, Flesch Reading Ease Formula, Flesch-Kincaid Grade Level, SMOG Index, Coleman-Liau and Automated Readability Index, Owu-Ewie (2014) was found that most of the passages were above the age of learners and were therefore difficult for them to read and comprehend.

Comment [P24]: Check the structure

Similarly, Gyasi (2013a) investigated the readability of science textbooks for senior high schools in Ghana employing the survey research design. Using stratified sampling technique, a sample size of 300 SHS 1, 2 and 3 students were drawn from five senior high schools in the Cape Coast Metropolis, Ghana. Gunning Fog and Cloze Test readability indexes were used to examine the difficulty level of the Physics, Chemistry, Biology and Integrated Science textbooks for senior high schools, written by the Ghana Association of Science Teachers. Findings of the study revealed that, on the average, the books are difficult to read. The study further showed that the

Integrated Science textbook is the most difficult among the textbooks, followed by the Physics textbook.

Along the same line of readability of textbooks, Gyasi (2013b) investigated the readability of handbooks of three reputable universities in Ghana, namely: University of Ghana, Kwame Nkrumah University of Science and Technology and University of Cape Coast. These readability indexes include: Flesch-Kincaid Grade Level, Flesch Reading Ease, Gunning Fog Index, Coleman-Liau Index, SMOG Index, Automated Readability Index and Lixear Write Formula. The readability consensus for the 7 readability indexes showed that these handbooks were very difficult to comprehend when measured in terms of readability indexes and that they were generally written to be understood by university graduates and in some cases even above the reading level of university graduates. The study also established that there were no statistically significant differences across the mean scores of the readability of the three handbooks.

Shifting focus from the readability of texts written by experts (e.g. textbooks published by learned lecturers), Gyasi (2011) considered the readability of essays written by undergraduate students in one Ghanaian university (University of Cape Coast). Two hundred and eighty students were selected across four (4) programmes of study namely Business, Arts, Science and Education. The selected students were made to write an essay on the topic, *“Ways of Conserving Electric Energy in the Various Halls of residence of the University of Cape Coast”*. Electronic forms of the essays were collected and fed to a computer. Using Microsoft word (2007) version, a Flesch- Kincaid Reading Ease analyses of the essays were run. The study revealed that essays

of first year level hundred students of Ghanaian universities were largely difficult to comprehend. Also, gender and programmes of study did not significantly influence the readability of essays of first year Ghanaian university students.

In the field of health studies, Gyasi, (2013c) conducted in-depth analysis of the readability of commonly used malaria medicine information leaflets in Cape Coast, Ghana. Seven leaflets of malaria medicines that are very popular in Cape Coast, Ghana were analyzed using the Flesch-Kincaid Reading Ease and Gunning Fog readability indexes. These leaflets were also examined based on the legibility of print, type of paper and bilingual information. Data collected from the two indexes revealed that all the leaflets are very difficult to read. Values ranged from 10.4 to 38 for the Flesch-Kincaid Reading Ease, indicating that the leaflets are very difficult to comprehend and that one needs university education to comprehend them. 14.2 to 18.8 were recorded for the Gunning Fog index, indicating that the leaflets are very difficult and that in some cases are even not readable for people with university education. The mean value for the readability consensus for the two readability indexes was 21.04. The implication is that in terms of the Gunning Fog index, the leaflets are very difficult to comprehend and that they were written far above the reading comprehension level for university graduates. Also, in the case of the Flesch reading ease regarding the readability consensus of the two indexes, the leaflets are very difficult to comprehend in terms of level of reading difficulty but match the appropriate reading level of the university graduate.

Fosu (2014) considered the readability of newspapers in Ghana. The study employed a research design that triangulated approaches in corpus linguistics, readability and survey studies. A computer-aided Linguistic analysis was carried out on the front-page stories of four influential national newspapers of the country to assess the extent to which the language is complex. The research established that the language used to communicate socio-political news to readers is complex and difficult for a significant proportion of readers across the educational categories of the country.

Conclusively, then, all the studies from the Ghanaian context have shown that texts authored by Ghanaians are at least 'difficult' to read when measured in terms of readability indexes. Yet, no study has been carried out to explicitly determine the readability of research articles published by lecturers of Ghanaian universities. Thus, this study contributes to the growing interest in readability of literature in the Ghanaian context, with the aim of helping to improve these literatures' readability.

Comment [P25]: Establish the gap well.

Comment [P26]: Place under readability ... the Ghanaian context.

Readability and text comprehension for L2 speakers of English language

Comment [P27]:

Although readability of written text is important in all contexts (that is, regardless of the language background of the reader), it is noted that for L2 language learners (e.g. Ghanaians using English language), the effect of readability on text comprehension is much profound. There is strong evidence to suggest that Word reading efficiency, or *Verbal Efficiency*, has significant impact on reading comprehension. Word reading efficiency refers to the ability of readers to identify printed words rapidly and accurately. Word reading efficiency has the power to predict reading success, which in turn is

used to predict reading comprehension. In short, when texts are made readable by using words that are easy to identify and pronounced accurately, it affects reading efficiency. As a result, reading comprehension is improved. Since it is more likely for L2 learners to have difficulty in accurately identifying and pronouncing words compared to their L1 counterparts, it should be the point for authors who write for L2 readers to especially present their thoughts by using simple, easy-to-read words and sentences. By this approach, readability and reading comprehension will be improved.

In recent years, a number of authors have shown the relationship between readability and reading comprehension, especially for L2 speakers of English. For example, in Israel, Saiegh-Haddad (2003) studied The relevance of oral reading fluency (ORF) to reading comprehension in the native language (L1) and in English – a foreign language (L2). Using fifty university students, twenty-two of them being Arabic and twenty-eight Hebrew native speakers, these respondents were made to read both L1 and English texts aloud and reported their comprehension on-line. Their Results showed that ORF was significantly correlated with reading comprehension in English (L2) but not the L1 languages (Arabic and Hebrew).

Fuchs et al. (2001) similarly reported high correlations between ORF and various kinds of reading comprehension measures such as high stakes state-mandated comprehension tests, as well as a variety of comprehension tests using different formats (e.g., multiple choice, open questions, cloze procedures or story recall protocols).

Earlier, during the development of the theoretical basis for the relationship between reading fluency (readability) and reading comprehension,

Jenkins and Jewell (1993) were concerned that the correlation between ORF and performance on direct measures of reading might be spuriously inflated because of the heterogeneity of the students studied. To address this question, they compared the performance of second-, third-, fourth-, fifth-, and sixth-graders on two standardized tests, the Gates-MacGinitie (MacGinitie, Kamons, Kowalski, MacGinitie & McKay, 1978) and the Metropolitan Achievement Test (Prescott, Balow, Hogan & Farr, 1984), with their performance on two informal formative measures of reading, including ORF. Concurrent validation revealed that the correlations between the two standardized tests and ORF varied according to students' grade level, with a declining trend in the coefficients as grade level increased. This negatively accelerating curve was attributed to a difference in the sensitivity of the test to individual differences in reading ability for students in the early grades versus those in the later grades. The second explanation that they proposed related to the variable sensitivity of the test to gains in reading proficiency. The third and most powerful explanation distinguished between two kinds of reading skills: word decoding versus language comprehension and world knowledge. Many L2 readers decode word-by-word so slowly that they cannot retain enough information in their working memories long enough to comprehend connected text (Gorsuch, 2013). As a result, some FL/L2 researchers and educators indicate the need to find effective methods to help learners develop their reading fluency (Day & Bamford, 1998; Grabe, 1991, 2004; Silberstein, 1994). Therefore, authors of research articles in the Ghanaian context need to appreciate the fact that since English is a second language for most in academia in Ghana, there is the likelihood that people will have difficulties in

processing text written in complex grammatical structures. It is therefore advisable that in writing research articles, authors adopt plain language to make the articles readable to members of the academic community. This is especially important since many L2 readers decode word-by-word so slowly that they cannot retain enough information in their working memories long enough to comprehend connected text (Gorsuch, 2013).

CHAPTER FOUR

RESEARCH METHODOLOGY

Introduction

This chapter describes the procedures and techniques used to collect and analyse data for the study. It captures the study area, research design, population, sampling procedure, sample size, research instruments, data collection, data processing and analysis procedures that were employed as well as the rationale behind choosing these techniques for the study.

Study Area

The study area for the present study is the University of Cape Coast which was established in October, 1962 out of a dire need for highly qualified and skilled manpower in education. Its original mandate was to train graduate professional teachers for Ghana's second cycle institutions and the Ministry of Education, in order to meet the manpower needs of the country's accelerated education programme at the time. The University started with two departments, namely Arts and Science. These departments have advanced to become independent Faculties.

The Faculty of Arts is currently made up of 9 departments, one centre and one teaching unit. The focus of the Faculty is to promote teaching, research and extension of various aspects of liberal and creative Arts. The total number of academic staff in the Faculty currently stands at 108. This is made up of 15 Professors, 22 Senior Lecturers, 49 Lecturers and 22 Assistant Lecturers.

Comment [P28]: source

Research Design

The study is mainly quantitative, focusing on a descriptive research design. According to Aliaga and Gunderson (2002), quantitative research methods seek to explain phenomena by collecting numerical data that are analysed, using mathematically based methods (in particular statistics). Descriptive research, according to Blessing and Chakrabarti (2009), answers ‘what’ questions, but of the type ‘how many’ and ‘how much’, because it is aimed at “describing the incidence or prevalence of a phenomenon or to be predictive about certain outcomes”. As a form of quantitative research, the descriptive research design uses deduction by deriving hypotheses from theory and analysing the data collected to statistically test the hypotheses (Blessing & Chakrabarti, 2009). Babbie (2001) also indicated that descriptive research design is useful for generalising from a sample to a population so that inferences can be made about the characteristics, attributes or behaviour of the population. The choice of descriptive research design by the researcher was informed by its relevance to the present study. For example, the design helps ascertain the relationship between readability and text comprehension. Moreover, the design is helpful in determining the level of reading comprehension difficulty of sampled research articles across disciplines in the Arts.

Comment [P29]: Refer to comment in the draft.

Population, Sampling and Sample Size

The target population was the entire number of lecturers in the Faculty of Arts, University of Cape Coast, Ghana, who were active in publishing research findings (N=123).

The faculty runs programmes leading to B.A., M.A., M. Phil. and Ph.D Degrees in various disciplines including English Language, Akan, Ewe and Ga, Classics and Philosophy, History, and Communication studies (Vice Chancellor's Report, 2016).

Non-proportional quota sampling technique was employed in sampling RAs for the study. According to Rukmana (2014) non-proportionate quota sampling is a method of non-probability sampling when the samples are selected based on the probability which is non-proportionate to the distribution of a variable in the population. In other words, you specify the minimum number of sampled units you want in each category. The researcher is not concerned with having numbers that match the proportions in the population. Instead, you simply want to have enough to assure that you will be able to talk about even small groups in the population a nonproportionate quota sampling technique was used to sample authors of RAs, using the Departments of the Faculty as the strata. A nonproportionate quota sampling was employed because readability within each Department was of interest. Yet, since the sample frame for RAs (total number of RAs published by faculty) was not readily known, samples were collected to meet the minimum requirements without consideration to the proportions of RAs written per departments and gender. Moreover, the stratification was non-proportionate because RAs published by authors differed in their frequency of publication and number of publication per Department. Some departments had only few lecturers who published their research findings. Hence, a proportionate quota sampling technique could result in selecting RA by just one individual in the Department (Marchevsky, 2000). A simultaneous stratification using both

Department and gender was not possible because certain Departments had only one gender of lecturers.

Data Collection Instruments (Data Sources)

Both Secondary and Primary data were used for the study. The secondary data were obtained from three (3) sources. First, fifty-nine (59) peer-reviewed online RAs authored by members of the eight (8) Departments of the Faculty of Arts were downloaded and readability scores computed. Second, hard copies of twenty-one (21) RAs from three (3) Journals of the Faculty of Arts were collected and retyped, using Microsoft's® Word processor. The journals used were (1) *Journal of Philosophy and Culture* by the Department of Classics and Philosophy, (2) *Drumspeak* by the Faculty of Arts, and (3) *Abibisem* by the History Department. The *Abibisem* is a multidisciplinary journal committed to publishing well-researched general or technical articles in any of the fields pertaining to African history, philosophy, culture and civilization, relationship with the wider world, etc. The journal was first published in 2008 and has been published annually since then. The *Drumspeak International Journal of Research in the Humanities* is a journal of Faculty of Arts of the University of Cape Coast. The journals publish RAs from scholars in the Humanities on any subject that has been adequately researched. The journal has been published annually since 1996. Similarly, the *Journal of Philosophy and Culture* is published by the Classics Department of the University of Cape Coast. It started publishing in 1996 and is published annually. These journals were selected because they are housed in the Faculty and so faculty members publish in them.

Out of the 21 RAs from the three journals, 10 were from the *Journal of Philosophy*, six from Drumspeak, and 5 from Abibisem. Finally, twenty (20) RAs were requested from individual lecturers. Hence, 100 RAs were selected for the analysis.

Primary data was collected, using text comprehension exercises. In order to ascertain the relationship between readability and text comprehension, a comprehension text was conducted using three (3) different passages selected from the sampled RAs. These passages were the introductory sections of the selected RAs. These three articles selected for the comprehension exercises were highest (Text comprehension A), medium (Text comprehension B) and least (Text comprehension C) of the 100 sampled RAs in terms of their Flesch Kincaid Grade level scores. The text comprehension questions were 8, 10, 10 for comprehension A, B and C respectively. The variation in the number of questions for the text items was because the passages upon which the questions were based were not of the same length.

Validity and Reliability of Comprehension Tests items

Both face validity and content validity of the three (3) comprehension test items were ensured. To ensure face validity, the researcher made sure that the test items were not only based on the individual text but also covered the entire text in each case. Content validity was also ensured by the principal supervisor and two other colleagues in the Faculty. The principal supervisor and two faculty members read, edited and modified the test items and made sure they corresponded to the texts upon which they were based. This helped to ensure content validity.

A total of 45 respondents were used for a pretest of the comprehension texts (fifteen respondents for text A, B and C respectively) to ensure reliability of the test items. Fifteen third-year undergraduate respondents were taken from the University of Cape Coast to take the text comprehension exercise A. Also, 15 respondents each were taken from SHS3 and SHS 2 to respond to text B and C respectively. The choice of respondents was guided by the grade level score of the individual passages used as calculated by the FKGL index. The text comprehension exercises were scored and the results were inputted into IBM's Statistical Products and Services Solutions (SPSS) version 23.0 for analysis.

Comment [P30]: principal

The split-half reliability coefficient was used to determine the reliability of the test items in each comprehension exercise. The split-half reliability coefficient ensures that individual text items in each of the comprehension exercises are of equal level of difficulty. Table 5 shows the reliability coefficient of each of the 3 text comprehensions

Table 5: Split-Half Coefficient showing the Reliability of the Text Items of the three (3) Comprehension Exercises

Comprehension Exercises	No. of text Items	Number of Respondents	Reliability Coefficient
A	8	15	0.62
B	10	15	0.73
C	10	15	0.85

Source: Field Data, 2015

The results showed that the reliability coefficients of the comprehension exercises, A, B and C were 0.62, 0.73 and 0.85 respectively. According to Pallant (2013), reliability coefficient should be 0.6 or above in order to consider items in *n* instrument reliable. However, coefficients of above 0.7 are considered to be ideal. The results from the reliability coefficients showed that the text items in each of the three comprehension exercises were reliable. This implies that the items in each of the comprehension exercises were of equal level of difficulty especially the items in comprehension exercises B and C.

Data Collection Procedures

Hard-copy articles that were collected were retyped into electronic formats, using Microsoft's® Word processor. Afterwards, these electronic forms were edited to ensure that the final documents were exactly as the original hard-copies used. This was to ensure that all punctuations were at exact positions as the original hard-copies. Thereafter, three-hundred (300) sampled sentences were selected from each article such that one-hundred sentences each were taken from the introductory, methodology, and conclusion parts. The selection of 100 sentences was done such that the last sentence ended with a 'full stop'. Hence, some samples had little over 300 sentences in all.

Two (2) readability scores: Flesch Reading Ease (FRE) and Flesch – Kincaid Grade Level (FKGL) were computed from these articles. Readability indices were computed, using average scores from three (3) online readability calculators: www.readabilityformulas.com, www.usingenglish.com, and www.dairyscience.info. Three (3) different readability calculators were used because for each readability index, each of these calculators (including Microsoft's

Word) gave slightly differing figures. Hence, the practice of earlier researchers running the readability analysis using more than one calculator and finding the mean of the scores was adopted (Owu-Ewie, 2014). Microsoft's Word processor was not used for the readability indices because it has a bug which prevents it from calculating FKGL values above 12. Therefore, if Microsoft Word Processor were used for computing FKGL, RAs that scored above FKGL of 12 will automatically rounded up to 12 thereby distorting the findings of the research.

For the primary data collection, comprehension texts were administered to 300 respondents, 100 respondents for each text i.e. comprehensions A, B and C. The FKGL scores for the passages for comprehension A, B and C were 15.0, 11.5 and 11.0 respectively. Participants were told the purpose of the exercise and assured of anonymity. As already indicated, FKGL is the number of years of formal education a person must obtain to comprehend a given text (Cutts, 2013; Dolnicar & Chapple, 2014; Fakhfakh, 2015; J. Hartley & Knapper, 2008). Hence, FKGL of 15 implies that a reader must acquire a minimum of 15 years of formal education to comprehend the text. This corresponds to a minimum of 3-year university education; hence 100 respondents were randomly selected across disciplines in the University of Cape Coast to take the text comprehension text item A. Also 100 respondents each were taken from SHS3 and SHS 2 to respond to text B (FKGL=11.5) and C (FKGL=11) respectively. The text comprehensions were scored and scores were converted to percentages for easy analysis.

Data Analysis

Two (2) readability scores: (Flesch Reading Ease (FRE) and Flesch – Kincaid Grade Level (FKGL)) were computed from the 100 RAs. The Flesch Reading Ease Readability Formula is one of the oldest and it is considered to be the most accurate of all the readability formulas. It was developed in 1948 by Rudolph Flesch who is a writer and a reading consultant. It is a simple approach to evaluating the grade-level of readers. This formula is mostly used for academic text. It is largely used to assess the difficulty of a reading text written in English language. According to Cutts (2013), this formula uses a scale from 0 to 100; where 0 corresponds to a reader with graduate level of education and 100 is also equivalent to a reader with 5th grade level of education (Primary 5). This simply means that the higher the score the easier the passage to be read and the lower the score the more difficult the passage.

Flesch-Kincaid Grade Level Test is a related test which translates the Flesch Reading Ease Test scores to grade level (Table 6). It converts the reading ease score to the number of schooling years required to be able to read the text. The formula was propounded by Peter J. Kincaid and his team in 1975. It is mostly used in pedagogy. This formula is used to determine the readability level of a variety of educational materials especially books. This formula makes it easier for parents, teachers, and librarians to select suitable reading texts for their children/learners (Owu-Ewie, 2014) (See Table 6).

Table 6: Flesch Readability score and Corresponding Interpretation

Flesch Reading Ease Score	Interpretation	Equivalent US grade Level
90 – 100	Very easy	5
81 – 89	Easy	6
70 – 79	Fairly easy	7
60 – 69	Standard	8-9
50 – 59	Fairly difficult	10 – 12 (High Schoolstudent)
30 – 49	Difficult	13 – 16 (College Student)
0 – 29	Very difficult	College Graduate
Less than Zero	Extremely Difficulty	Law School Graduate

Source: Adapted from Flesch (1949), Kincaid (1975) and Cutts (2013)

The resultant FKGL and FRE were analysed with the help of IBM Statistical Products and Services Solutions (SPSS) version 23.0. Frequencies, percentages, means, standard deviations, median and Interquartile range were used to describe readability, number of years to be able to read, and patterns in readability scores according to departments in the objective one (Research question 1) which was to determine the level of readability of research articles in the Arts. Wilcoxon signed ranked test was used to compare whether readability score and the number of schooling years required to read research articles differed from standard values stipulated in literature (Hypothesis 1). Microsoft Excel® was used to plot FRE frequency distributions with normal plots to describe patterns in readability.

Objective two (i.e Hypothesis 2) which was to determine the differences in the readability across the eight (8) disciplines/Departments in

the Arts was analysed using Kruskal – Wallis H test. The Kruskal – Wallis H test which is the non-parametric equivalence of One-Way ANOVA was used because the data did not meet the assumptions of parametric analysis such as normality and homogeneity of variance. Moreover, the type of sampling used was non-probability even though some studies have shown that parametric statistics can be used under non-probability sampling situation. (citation) Eta-squared was estimated to measure the magnitude or the effect size of the significance when there were significant differences. In addition, Dunn’s multiple comparison test was employed to determine where differences existed (Field, 2011; Pallant, 2013).

A Mann-Whitney U test was used to analyse objective three (Hypothesis 3) that compared the difference in the readability of female and male authored research articles. The Mann-Whitney U test was used instead of the Independent-sample t-test since test data failed the Shapiro-Wilk Normality Test.

Objective four (Hypothesis 4) was analyzed using Spearman rho correlation coefficient to explore the relationship between the readability, lexical density. Davis Convention was used to interpret the strength of the correlations (Davis, 1971). In addition, percentages and histograms were used to describe the pattern of distribution of readability scores according to departments.

Finally, for objective 5, 2x2 crosstabulation were used to assess pass rate of respondents on the comprehension tests. In addition, Yate’s continuity correction, Fisher’s exact test, and phi were used to determine difference in test comprehension between males and females. In order to categorise the

Comment [P31]: Refer to comment in the draft.

performance of participants in the comprehension test, Participants' performance in the MCCE has been graded as 'above average' or 'average and below'. Participants whose performance were graded as 'above average' were those who scored more than half (greater than 50%) of the questions on the MCCE. Conversely, participants who scored half (50%) or lesser on the MCCE were categorised as 'average and below'. The rationale guiding this categorization is that majority of participants in the MCCE are expected to score above average (50%) if they indeed understood what they read.

For hypotheses 2 and 3 that compared readability using Kruskal – Wallis H and a Mann-Whitney U test, bootstrapping was performed for samples of 1000 to ensure robust estimates of significant or p-value, standard errors and the confident intervals (IBM, 2013;Tabachnick & Fidell, 2013;Field, 2013). Since, Kruskal–Wallis H test was significant, bootstrapping was performed to ensure robust estimates for the mean differences and significant test in the Post hoc Multiple Comparison. Bias corrected and accelerated (BCa) intervals were used to ensure robust confidence intervals (IBM, 2013). Mersenne Twister Random Number Generator was set to replicate a sequence of random numbers. This helped to preserve the original state of the random number generator and restore that state after the analysis was completed (IBM, 2013).

Table 7 shows the summary of the research objectives and hypotheses and their corresponding statistics used for the data analysis.

Table 7: Summary of Statistics used for the Data Analysis

Objective/Hypothesis	Statistical tool/tools
One	Frequencies, percentages, means, standard deviations, median and Interquartile range and Wilcoxon signed ranked
Two	Kruskal – Wallis H test
Three	Mann-Whitney U test and Shapiro-Wilk Normality Test.
Four	Spearman rho correlation coefficient, histogram, and percentages
Five	2x2 crosstabulation, Yate’s continuity correction, Fisher’s exact test, phi and odds ratio

Source: Authors’ Construct (2016)

Comment [P32]: Refer to comment in the draft.

Ethical Considerations

In collecting data for the study, the researcher made some ethical considerations. First, permission was sought from heads of schools to administer the comprehension tests to students at the SHS level for both the pretest and the actual data collection exercise. Again, the students were assured of anonymity of their responses and were thus not required to write their names on the multiple-choice answer sheets. With respect to the administration of comprehension test at the tertiary level too, prior arrangements were made with lecturers to allow the use of their classes to administer the comprehension test for both the pretest and actual data collection exercise. These were also assured of anonymity of their responses. Another stage where ethical consideration was shown was the gathering of RAs from departments within the Faculty of Arts for the study. In this case, I presented my identification card to the librarians who allowed me to select the

journals I needed for the research. Photocopies of RAs suitable for the study were made from these journals after which they were returned to the libraries.

Chapter Summary

In this chapter, I discussed the quantitative research design which was employed in this study, and a justification of the choice of design was presented. In addition, the sample and sampling procedure adopted for the study were considered, as well as the data collection instruments used. A Split-half analysis was used to achieve reliability and validity of the data collection instruments. Further, data collection procedures related ethical issues were presented. Finally, the various analytical techniques employed to answer each research question indicated in chapter one, were also considered.

CHAPTER FIVE

READABILITY OF RESEARCH ARTICLES IN THE HUMANITIES

Introduction

The purpose of reading in academia is to get information from the text read for the purpose of teaching research (Altman, 2015). Therefore, if a text is incomprehensible, members of the research community derive no benefit from the text. Precious time will also be wasted since readers will have to spend a lot of time in trying to understand the text. Since RAs are the chief scholarly source of information in the research community or academia, it is important that writers of this genre take into consideration the level of reading comprehension difficulty of this all-important document (Priestley, 2015). In this chapter, I give attention to the readability of RAs in the humanities, determining their level of reading comprehension difficulty when compared to recommended readability scores.

Readability of RAs in the Humanities

The objective of this chapter was to determine the level of readability of RAs in the Arts in the University of Cape Coast and to compare the level of their readability with standard readability scores. Frequency distribution (Figure 3) was used to assess the readability of research articles in the Arts. The results are presented in Figure 3.

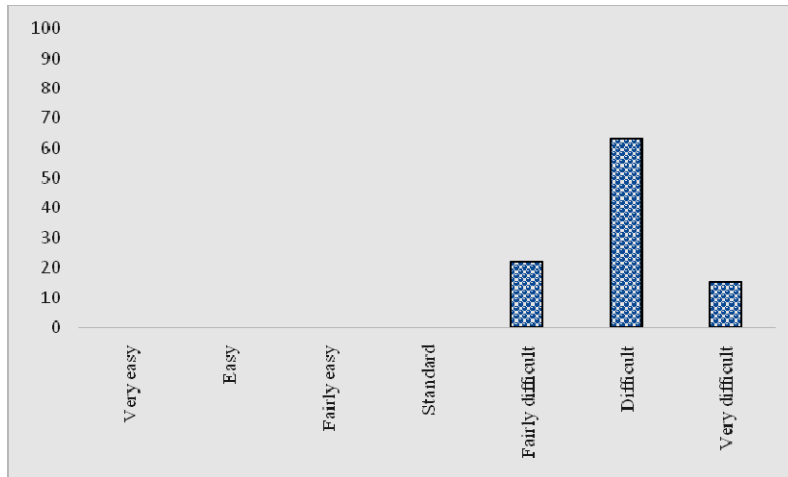


Figure 3- Frequency distribution of readability scores of research articles in the Arts

It is observed that the majority (63%) of the RAs were graded as ‘difficult’ to read, which implies that a reader requires some college level of education to be able to read and comprehend (Cutts, 2013; Dolnicar & Chapple, 2014; Fakhfakh, 2015; J. Hartley & Knapper, 2008). Quite substantial proportions (22% and 15%) of the RA were ‘very difficult’ (college graduate level) and ‘fairly difficult’ (high school level) to read respectively when measured in terms of the FRE formula. None of the research articles was graded as ‘standard’ ‘fairly easy’, ‘easy’ or ‘very easy’ when measured in terms of Flesch reading ease. In other words, all of the research articles were above ‘standard’ readability (Cutts, 2013). The above results agree with the findings of a number of research works both in the Arts and other disciplines. For example, in the Arts (Tourism), over 84% of 493 RA from different institutions that were reviewed by Dolnicar and Chapple (2014) were found to be ‘very difficult’ to read when measured in terms of Flesch reading ease index, requiring college graduate level of education to be

able to read and comprehend these articles. Similarly, Lee and French, (2011) found that 50 per cent and 48 percent of research papers that were published in the *Journal of Property Investments* were in the “very difficult” and “difficult” ranges respectively. In addition, five RAs were found to be in the “fairly difficult” range. Only one research article was in the plain English range suggested by Flesch (1949). Several other authors have found similar proportions of RAs of same readability level (e.g. Garcí'a-Merino & Santos-a'Lvarez, 2009; Lee & French, 2011; Lim, 2010; Otto, Joseph; Otto, Joseph, Parviz Partow-Navid, & Doshi., 2010). This implies that there seem to be general consensus by scholars that RAs are generally difficult to comprehend.

Comment [P33]: Refer to comment

Syntactic features of research articles in the Arts across disciplines

Comment [P34]: Refer to comment.

Table 8 presents syntactic features of the RAs analysed in this study. Here, number of words per sentence and number of characters per words across RAs were analysed. As can be seen from Table 8, the minimum mean length of sentence written by authors from the Faculty of Arts was approximately 21 (\bar{x} = 20.77, SD = 3.25). Cutts (2013) recommends an average sentence length of 15-20 words to achieve a standard readability (FRE = 60-70) of a text.

Comment [P35]: Fix the table

Table 7: Descriptive Statistics of the Syntactic features of Research Articles in the Arts across Disciplines

Department/centre	Words per sentence				Characters per word			
	Min	Max	Mean	SD	Min	Max	Mean	SD
Rel. Human Values	12.50	26.50	20.77	3.25	4.70	5.50	4.96	.22
English	22.20	46.10	27.89	6.68	4.80	5.70	5.25	.276
Gh. Lang. & Linguistics	16.80	31.50	23.39	5.51	4.50	5.60	4.96	.372
History	20.40	39.00	27.51	6.00	4.60	5.50	4.98	.27
Classics and Philosophy	9.80	35.60	26.79	7.39	4.60	5.20	4.85	.20
Communication Studies	19.10	39.60	25.21	5.81	4.80	5.60	5.25	.26
Music and Dance	18.70	41.70	26.52	6.94	4.80	5.70	5.22	.27
African studies	17.10	34.10	24.86	5.13	4.80	5.50	5.10	.21
Faculty	9.80	46.10	25.38	5.84	4.50	5.50	5.07	0.26

Comment [P36]: Supervisors's comment.

Comment [P37]: Supervisors's comment.

Source, Field Data, Gyasi (2015)

The minimum mean length of sentence of approximately 21 words was still above the recommended value by Cutts in order to achieve a standard readability. Cutts (2013) notes the reason why a maximum average number of words per sentence should not be above 20. He indicates that “readers recoil when they see a long sentence slithering across the page towards them like a Burmese python” (Cutts, 2013:21). He further explains that long sentences give the reader too much information to cope with at a time. Again, he notes that unless long sentences are of simple construction, they cause confusion because they demand so much effort and short-term memory. It requires a

reader holding one thought in memory as s/he grapples with another thought when the sentence is long. When one writes long sentences offering so many ideas (points) at once it becomes hard for the reader to read. This is the case even if the script is easy to read at one level (especially if most of the words and ideas are reasonably simple). Thus, the findings as presented in Table 8 indicate that too many ideas are forced in one single sentence by the authors from the Faculty of Arts. This may account for the generally low reading ease (low FRE scores).

Table 4 again indicates that the Department of Religion and Human Values wrote the least average number of words per sentence ($M = 20.77$; $SD = 3.25$). The Department of English on the other hand, recorded the longest average number of words per sentence ($M = 27.89$; $SD = 6.68$). It is apparent, then, that the low FRE scores might be due to the average number of words written per sentence.

The ease of reading RAs according to departments in the Faculty of Arts was measured using FRE, and the results have been presented in Table 9. Generally, the FRE scores, shown in Table 9, indicate low readability of the articles in the departments across Faculty of Arts. This is based on the view that Fleisch Reading Ease score of 60-70 is desirable for English texts (Fakhfakh, 2015). The mean FRE score for all departments fell below the desirable score. The highest maximum FRE score (58.2) which was scored by the Department of Classics and Philosophy, fell below the standard score of 60-70.

Comment [P38]: Introductory comment is needed

Comment [P39]: Refer to comment

Comment [P40]: Introductory comment is needed.

Table 8: Descriptive Statistics of Flesch Reading Ease for Articles from the Faculty of Arts

Department	Descriptive Statistics								Shapiro – Wilk Normality Test		
	Min	Max	Mean	SD	Median	IR	Skewness	Kurtosis	Statistic	df	Sig.
Religion	23.8	57.0	46.34	7.97	48.75	11.47	-1.05	9.62	.927	60	.002
Ghanaian Language	25.9	57.8	44.53	10.96	50.30	19.18	-.511	-1.025	.909	30	.014
Classics & Phil.	30.6	58.2	39.85	9.61	37.55	12.20	.521	-.232	.960	30	.315
African Studies	26.4	45.2	37.17	6.39	40.95	11.34	-1.074	.936	.901	30	.009
History	25.3	47.8	36.93	6.86	38.05	14.07	.418	-.708	.954	30	.215
Comm. Studies	17.8	42.1	33.57	7.48	37.10	9.30	-.718	.222	.947	30	.142
Music and dance	16.6	48.6	30.96	10.74	32.60	20.83	.143	-1.155	.935	30	.068
English	3.2	49.7	29.57	14.42	32.00	21.30	-.454	-.841	.932	60	.003
Faculty of Arts	3.20	59.46	38.66	11.71	40.10	15.76	-.585	.230	.973	300	.000

NB: Means have been arranged in descending rank order.

Source, Field Data, Gyasi (2015)

Therefore, the RAs in the department will generally be classified as ‘difficult’ to read, to the extent that the easiest to read of all the academic papers was still *‘fairly difficult’*. The Department of English recorded the lowest minimum readability score of 3.2, which implied that those texts were ‘very difficult’ to read. The highest mean readability was scored by articles from the Department of Religious and Human Values (\bar{x} =46.340; σ =7.97).

This implies that on the average, the ‘easiest’ articles were scored by the Department of Religious and Human Values. Yet, even the articles from this department were still classified below the ‘standard’ readability level, implying that on the average, the easiest of all the articles was still ‘fairly difficult’ to read.

On the other extreme, the smallest of the mean FRE score was from the Department of English (\bar{X} =29.57; σ =14.42). Hence, the English department produced the lowest FRE score indicating very difficult readability. It is the department with the highest standard deviation (14.42) from least mean index (29.57), at minimum (3.2) and maximum (49.7). The relatively higher standard deviation (compared to the Department of Religion for example) indicates that many of the articles produced by the Department of English fell within the range classified as ‘very difficult’ to read.

These findings parallel several other authors who have found similarly low readability in several fields including banking and finance (e.g. Baker III & Kare, 1992; Bali & Hovakimian, 1999; Lee & French, 2011), Arts (e.g. Dolnicar & Chapple, 2014), internet sites (e.g. AlKhalili, Hubbi, Patel, Sanghvi & Shukla, 2015; Lim, 2010), medical sciences (e.g. Schmitt & Prestigiacomo, 2013), and for pedagogical purposes (e.g. Gupta, 2014;

Comment [P41]: Refer to comment in the draft.

Mavasoglu & Dincer, 2014). The findings of these authors and many others raise serious concerns about the readability of academic journals. This has resulted in readers perceiving these journals as of poor quality. This perception is apparently because, essentially, the purpose of academic research is to develop knowledge which in turn informs policy making (Dolnicar & Chapple, 2014). Writing 'hard' to read texts works much to the detriment of knowledge development, and policy formulation for that matter. Hard-to-read text results in a situation where a given piece of writing fails to reach and affect its audience in the way that the author intends (Tekfi, 1987). Since understanding of a text is negatively related to the "complexity of written material" (Tekfi, 1987), it follows then that articles written by authors from the Faculty of Arts misses this essential intent.

There could be a number of reasons why academic articles in the Faculty of Arts are very hard to read. First, Researchers who are well established and talented enough to have their work published in journals may not necessarily give attention to readability when writing. Though these researchers are highly intelligent and sound in their specialized areas, writing a text with readability in view should always be a matter of consciousness in the case of the writer.

Secondly, an academic article is not intended to be a literary piece, or even a textbook. It is intended to be written by experts, for experts. Moreover, the FRE score used to determine the degree of difficulty of a text does not consider the background of the reader (field of study). This makes inferences made from the FRE score subject to debate. For instance, an article written to communicate newly found research findings mostly target other experts in that

field of study. Such experts are likely to be familiar with the terminologies, jargon, style of writing etc. of that field. For such ones, the degree of difficulty may not be as much as if that same text was to be read by another person from a different field. This gap is irrespective of the person's knowledge level, experience, age, or the years of schooling. Authors do not spend a lot of time providing background information, because the readers are presumed to have an expert background in the same field already. The expected formula is this: concisely introduce your research, explain approach (methodology), and present your findings. This is not to say that a well-written piece will not be preferred over a poorly written piece, but that is not the main emphasis. Those reading journals are reading to be informed about the newest findings, not taught or entertained.

Third, by nature, many journals are hard to read because the subjects are quite advanced. Authors are condensing months' worth of research into a single paper, and, in the scientific fields, their findings are supported by statistics derived from technical experiments. It is not meant to be browsed in the waiting room at the dentist's office.

Comment [P42]: Refer to comment in the draft.

It is true that difficult-to-read texts sometimes indicate poorly written texts. This is so because the logic behind this conclusion is that if a text were well expressed we would understand it; it would be clear and obvious. We would get it straight away. While that might sound logical, it is not always true. Getting into a new area or mode of thinking is like getting to know a new physical location. When you arrive in a new city you don't expect to know how to get around straight away. You don't expect to know a new place in the way you know your own home environment. You understand that you have to

make several trips before you have a sense of what is where, and how to get from one place to another without looking at a map for general directions and/or reassurance. And that is how it often is with new literatures. You have to explore a bit. You have to get a sense of what is where – the histories of debate, the lines of argument, the language used, the kind of questions that are asked, the topics that are pursued, perhaps even the style of writing that is generally used in the field. Hence, a difficult to read text (as has been portrayed by FRE scores of the Faculty of Arts presented above) may be a result of presentation of difficult texts and not poor writing style in itself.

Nevertheless, readability indices anchor on **length of sentences.**

Comment [P43]: Just that ?

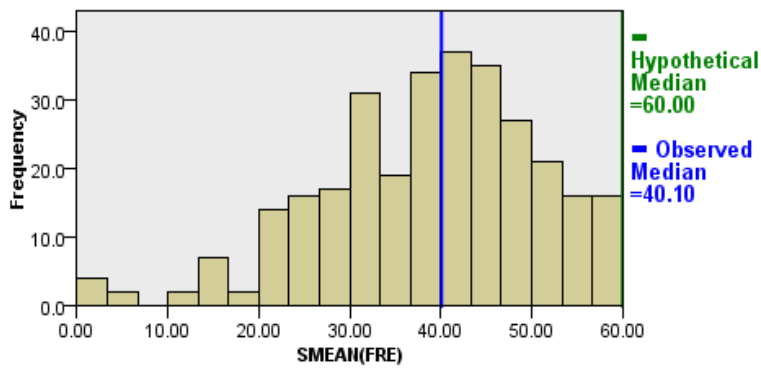
Readability scores predict that longer sentences are difficult to read and understand than shorter sentences (Bailin & Grafstein, 2016). Hence, the findings from this research question suggest that several of the authors were inclined to writing long sentences. The descriptive statistics presented in Table 4 which shows the various features of the sample of texts per departments analysed seem to support the prediction that the authors of the various articles under review, used longer sentences than experts recommend.

Differences in readability of research articles in the Arts and standard readability scores

Comment [P44]: Refer

Several authorities (eg. Cutts, 2013; Fakhfakh, 2015) have suggested an FRE score of 60 to be the standard readability score. A standard FRE score refers to a readability score at which the text is considered generally to be understandable to most readers (Cutts, 2013; Fakhfakh, 2015). Corpora at this FRE score are termed ‘plain language’. Hence, all standard writings must

score an FRE of at least 60. A Wilcoxon signed-rank test was used to test the hypothesis which evaluated whether there were significant differences in readability of research articles in the Arts and standard readability scores. Figure 4 presents the findings from the Wilcoxon signed – ranked test showing the aforesaid difference.



Total N	300
Test Statistic	.000
Standard Error	1,503.742
Standardized Test Statistic	-15.013
Asymptotic Sig. (2-sided test)	.000

Figure 4: One-sample Wilcoxon signed rank test

The results showed that the median FRE score (40.10) for the Faculty of Arts was below the standard readability score of 60 ($z = 15.01$, $p = 0.000$) at the 0.05 alpha level of significance. (Figure 4). This difference (19.90) was statistically different from the standard FRE score of 60. This implies that over half of research articles in the Faculty of Arts are above the 'standard' readability level. Hence, majority of the authors' research articles were relatively difficult to read compared to the 'standard' readability level.

The study therefore rejects the first null hypothesis which stated that there are no significant differences in readability of research articles in the Arts and standard readability score. The alternative hypothesis which states that there are significant differences in readability of research articles in the Faculty of Arts and standard readability scores is therefore accepted.

CHAPTER SIX

RELATIONSHIP BETWEEN READABILITY OF RESEARCH ARTICLES, DISCIPLINE AND GENDER

Introduction

Existing literature on readability studies reveal that several factors influence the readability of a text. Interest of the reader, legibility of the print of the text and motivation of the reader, whether intrinsic or extrinsic, has all been discussed in literature as factors that influence readability (McNamara, 1996). In this chapter, I give attention to two factors that can influence the readability of a text.

Differences in the level of Readability of Research Articles across Disciplines/ Departments in the Arts

The objective of this section was to determine the differences in the readability across the disciplines/ Departments in the Arts. Mean rankings of FRE scores across departments were used to analyse this sub-objective. The mean ranks have been arranged in decreasing order of difficulty level. The mean ranks are such that the higher the mean rank score, the lower its rank position and the easier it is to read. That is, the department with the least means rank score wrote the least readable RAs while the department with the highest mean rank score wrote the most readable research articles. Table 10 shows the mean FRE rank scores across Departments in the Faculty of Arts. Generally, it was observed from Table 10 that the Department of Religion and Human Values (mean FRE rank score = 76.55) wrote RAs which were the

easiest to read while the Department of Music and Dance (mean FRE rank score = 32.70) wrote the most difficult to read research articles.

Table 9: Mean Rank FRE Scores across Departments

Department	N	Mean Rank score	Rank
Music & Dance	10	32.70	8 th
English	20	33.95	7 th
Communication Studies	10	37.10	6 th
Classics & Philosophy	10	44.30	5 th
African Studies	10	46.10	4 th
History	10	53.60	3 rd
Ghanaian Language	10	70.20	2 nd
Religion and Human Values	20	76.55	1 st

n=100

Source: Field Data, Gyasi (2015)

In order to determine if there are any statistically significant difference(s) in FRE scores among the eight (8) departments in the Faculty of Arts, Kruskal – Wallis H (K-W) test was employed. This analytical technique was employed because an important assumption of the analysis of variance (the assumption of normality) was violated (see Table 9). Under such circumstances, the non-parametric equivalent of one one-way, between groups variances, that is Kruskal – Wallis H test, is the appropriate technique to use (Field, 2011; Julie Pallant, 2013).

The results from the K – W test showed that there were significant differences in the FRE scores across the eight (8) departments [H (7, n = 100) = 33.945; p = 0.000] at the 0.05 alpha level (see Table 11). Therefore, the

second null hypothesis that stated that there are no statistically significant differences in the readability of research articles across the disciplines/ departments in the Arts was rejected.

Table 10: K-W H test showing differences in FRE scores across Departments

Test Statistics	FRE
Chi-Square	33.945
Df	7
Asymp. Sig.	.000

n=100; p<0.05

Source: Field Data, Gyasi (2015)

The alternative hypothesis that stated that there are statistically significant differences in the readability of RAs across the disciplines/ Departments in the Arts is therefore accepted. This implies that there were actual differences in readability of research articles across departments in the Faculty of Arts. Therefore, an additional test was needed to determine where differences in readability of research articles across departments existed. Dunn's post hoc multiple comparison test was used to determine where differences in readability of research articles across departments existed. The results of the Dunn's post hoc multiple comparison test is shown in Table 12.

The results show that there are significant differences in FRE scores among the following pairs of Departments: Music & Dance and Religion & Human Values, English Language and Ghanaian Languages & Linguistics, English Language & Religion & Human Values, and Communication Studies and Religion & Human Values.

Table 11: Dunn's Post hoc Multiple Comparison Test of FRE Scores across Departments in the Arts

Sample1	Sample 2	Test statistic	Std. error	Std. test statistic	Sig.	Adj. Sig.
Music & Dance	English	1.25	11.23	.11	.91	1.00
	Comm. Studies	4.40	12.97	.33	.735	1.00
	Classics & Philosophy	11.60	12.97	.89	.37	1.00
	African Studies	-13.40	12.97	-1.03	.30	1.00
	History	20.90	12.97	1.61	.10	1.00
	Gh. Language	37.50	12.97	2.89	.00	.10
	Religion	43.85	11.23	3.90	.000	.00
English	Comm. Studies	-3.15	11.23	-.28	.77	1.00
	Classics & Philosophy	-10.35	11.23	-.92	.35	1.00
	African Studies	-12.15	11.23	-1.08	.280	1.00
	History	-19.65	11.23	-1.75	.08	1.00
	Gh. Language	-36.25	11.23	-3.23	.001	.035
	Religion	42.60	9.17	4.64	.00	.00
Comm. Studies	Classics & Philosophy	7.20	12.97	.55	.57	1.00
	African Studies	-9.00	12.97	-.69	.488	1.00
	History	16.50	12.97	1.27	.203	1.00
	Gh. Language	33.10	12.97	2.55	.01	.30
Classics & Philosophy	Religion	39.45	11.23	3.51	.00	.01
	African Studies	-1.80	12.97	-.13	.89	1.00
	History	-9.30	12.97	-.71	.47	1.00
	Gh. Language	25.90	12.97	1.99	0.04	1.00
African Studies	Religion	32.25	11.23	2.87	.00	.11
	History	7.50	12.97	.57	.56	1.00
	Gh. Language	24.10	12.97	1.85	.06	1.00
History	Religion	30.45	11.23	2.71	.00	.18
	Gh. Language	16.60	12.97	1.27	.20	1.00
Gh. Language	Religion	22.95	11.23	2.04	.04	1.00
	Religion	6.35	11.23	.56	.57	1.00

n=100; p<0.05

Source: Field Data, Gyasi (2015)

The implication is that the Ras from Department of Religion and Human Values (Median FRE score = 48.75) were more readable than those from Department of Music & Dance (Median FRE score =32.60), Department of English Language (Median FRE score = 32.00) and the Department of Communication Studies (Median FRE score =37.10). In other words, research articles from Department of Music & Dance, Department of English Language and the Department of Communication Studies were more difficult to read compared to those from Department of Religion and Human Values. In addition, RAs from Department of English Language (Median FRE score = 32.00) were more difficult to read than those from Department of Ghanaian Languages & Linguistics (Median FRE score =50.30).

Further, there were no statistically significant differences, at an alpha level of 0.05, in readability of RAs among the following pairs of departments: Music & Dance and English Language, Music & Dance and Communication Studies, Music & Dance and Classics & Philosophy, Music & Dance and African Studies, Music & Dance and History, and Music & Dance and Ghanaian Language & Linguistics. In addition, there were no statistically significant differences in readability of research articles among the following pairs of departments: English and Communication Studies, English and Classics & Philosophy, English and African Studies, English and History, Communication Studies and Classics & Philosophy, Communication Studies and African Studies, Communication Studies and History, Communication Studies and Ghanaian Language & Linguistics, Classics & Philosophy and African Studies, Classics & Philosophy and History, and Classics & Philosophy and Ghanaian Language & Linguistics.

Finally, the following pairs of departments also show no statistically significant differences in readability of RAs: Classics & Philosophy and Religion, African Studies and History, African Studies and Ghanaian Language & Linguistics, African Studies and Religion, History and Ghanaian Language & Linguistics, History and Religion, Ghanaian Language & Linguistics and Religion. The implication is that RAs from Departments of Music & Dance were of equal readability with all other departments except Department of Religion & Human Values, where RAs were more readable than those from Department of Music & Dance. Another implication is that RAs from Departments of English were of equal readability with all other departments, except Departments of Ghanaian Language & Linguistics, and Religion & Human Values, where the latter departments (Ghanaian Language & Linguistics and Religion & Human Values) were more readable. The results also imply that research articles from Department of Communication Studies were of equal readability with all other departments except Department of Religion & Human Values. Thus, research articles from Department of Religion & Human Values were easier to read compared to those from Department of Communication Studies.

Apart from departments mentioned above (Classics & Philosophy, African Studies, History, Ghanaian Language and Linguistics Religion & Human Values and History), all other departments wrote RAs whose readability were of equal level of difficulty.

Since readability calculated using Flesch readability index depends on sentence, and word length, it is implied that research articles written by Department of Religion and Human Values were devoid of excessive complex

grammatical structures and polysyllabic words than those written by Departments of Music & Dance, English, and Communication Studies. In addition, research articles from English Department probably contained a higher proportion of polysyllabic words compared to the research articles from Departments of Music & Dance, and Communication Studies.

Similarly, Department of Ghanaian Languages & Linguistics wrote research articles with shorter sentence, word length, and higher proportion of polysyllabic words compared to research articles from Department of English. It is implied also that, comparisons of readability of research articles from all other departments were of equal word and sentence length. A possible reason why Department of English wrote very difficult to read research articles is that, since English Language as a course of study concerns the language itself, there is the likelihood that people associated with the discipline might consciously or otherwise equate linguistic competence to the use of complex grammatical structures and polysyllabic words. This is not surprising because an earlier author mentions that the educated Ghanaian is not only identified with use of learned and complex grammatical forms but is also fond of flamboyance of prose style (Darko, 2003; Mahama, 2012). If what these authors indicate is something worth considering, then the use of learned and complex grammatical forms and fondness for flamboyance of prose style will be more associated with those in the English Language discipline other than others in the other disciplines.

Concerning research articles from Department of Communication Studies, a possible explanation to the reason why they are more difficult in readability compared to research articles from Department of Religion and

Human Values may be as a result of the background of the authors of the research articles. All the authors in the Department of Communication Studies had their Bachelors and Master's degrees in English Language. It is therefore possible that the reason attributed to the difficult nature of research articles published by authors in the Department of English holds here. Thus, these authors are also likely to either consciously or unconsciously equate proficiency in written communication to the use of learned forms of grammatical structures. By so doing, they are clearly defeating the basic fact associated with communication. This fact is that when information is passed on to a person, it is to be done in a way that the receiver of the information will clearly understand it. In that case, effective communication has taken place.

The readability of RAs in Music and Dance reported in this study is similar to earlier findings by Humphreys & Humphreys (2013). After analyzing 22 RAs from the Music Education, the Flesch readability score was found to range between 24 and 52, which is similar to the findings in this study, with a grand mean of 38.68 (SD = 7.11). The study of music and dance cuts across not only those of Ghanaian origins but also those of other cultures across the world. The study therefore makes use of words and expressions in describing all forms of phenomena concerning music and dance across cultures around the globe. It is therefore likely that the study brings to bear all kinds of learned forms and complex grammatical structures borrowed from other cultures. This may be a reason for the complexity of language use in research articles in Music and Dance.

Differences in Readability Scores between Male and Female Authored Articles in the Arts

In this section, I determine the differences in readability of female and male authored research articles in the Arts.

Gender and readability of research articles in the Arts

Table 13 compares the readability of research articles by male and female authors in the Arts. It is observed that majority of research articles (94%) were authored by males. Research articles authored by females accounted for only 6% of the total number of research articles. In addition, the median FRE score for research articles authored by males was 49.87 while that of females was 60.33. This implies that half of research articles authored by males were in the range classified as ‘fairly difficult’ to read. In contrast, half of research articles authored by females were of ‘standard or average’ readability. Therefore, it appears that females seem to write more readable research articles compared to males.

Table 12: Descriptive Statistics of Readability of male and Female Authored Research Articles in the Arts Measured in terms of Flesch Reading Ease

Gender	n	Median	Q1	Q3	IR
Male	94	49.87	32.35	45.08	12.73
Female	6	60.33	37.91	47.50	9.59

n=100; p<0.05

Source: Field Data, Gyasi (2015)

Test of significant differences in the readability of RAs between male and female authors in the Arts

In order to determine if there are any statistically significant difference in readability of research articles written by males and females, Mann – Whitney U test was used instead of independent sample t-test. Mann – Whitney U test was used because FRE scores violated the assumption of normality as described earlier. Findings from the Mann – Whitney U test is presented in Table 14.

Table 13: Results of Mann - Whitney U test of FRE scores between male and Female Authored Research Articles

	Male (n=94)	49.87
Mean FRE Rank	Female (n=6)	60.33
Total <i>n</i>		100
Mann-Whitney U		171.00
Standard Error		68.89
Asymptotic Sig. (2 – sided test)		.107

n=100; p<0.05

Source: Field Data, Gyasi (2015)

A Mann-Whitney U Test revealed no significant difference in the readability of RAs of males (Md = 49.87, n =94) and females (Md = 60.33, n = 6), U = 341.00, z = -2.669, p = 0.39. Therefore, the second null hypothesis that stated that there is no significant difference between the readability of female and male authored research articles in the Arts was accepted. The alternative hypothesis that stated that there is significant difference between the readability of female and male authored research articles in the Arts was rejected. This implies that there was no actual difference in readability of research articles authored by both male and female authors in the Arts.

Therefore, no additional test was needed to determine the magnitude of the difference in readability of research articles between males and females.

The findings from this analysis parallel some earlier works which evaluated gender differences in readability of academic writing. For example, in one complex study which used the Linguistic Inquiry and Word Count (LIWC) software to evaluate readability of research articles in the *Journal of Educational Psychology*, Hartley, Pennebaker, and Fox (2003) only found minor differences when academic articles written by individual and pairs of men were compared with those written by individual and pairs of women. The clearest difference, which these authors could not explain, was that single men, and pairs of women produced texts with higher readability scores than did pairs of men, and single women! Similarly, Hartley and Knapper, (2008) reviewed 19 each of academic papers authored by males and females. These authors found no significant difference between the average scores achieved by men and women in terms of sentence length, percentage of passive words, and FRE scores. Basically, both males and females performed equally well with regard to readability of academic writings.

Contrasting this finding however, Hengel (2015) found that abstracts written by women were 2–6 percent more readable than those by men when she analysed abstracts published in the top four economics journals since 1950. Hengel (2015) found out that while sentence and character counts did not vary significantly by authors' sex, vocabulary and structure did. Women wrote shorter sentences and used fewer total syllables, fewer syllables per word and fewer “harder” words. Tentative explanation to her finding is that, women write more clearly as they publish more papers but men do not. Thus,

the readability gap widens as both gain academic credibility. Hengel (2015) concluded from her findings that (i) Peer review shortens sentences and reduces hard words per sentence: in male-authored papers, sentences are 5 percent shorter and contain 26 percent fewer polysyllabic words; in female-authored papers, they are 7 percent shorter and contain 30 percent fewer polysyllabic words. (ii) As a fraction of total word count, however, syllables, polysyllabic words and difficult words rise. That is, hard word counts and total word count decline, but total word count decline proportionately more; their ratios increase: between 1–3 percent for men and 1–2 percent for women. Therefore, the review process, according to Hengel (2015), influences the readability of RAs.

CHAPTER SEVEN

RELATIONSHIP BETWEEN READABILITY, LEXICAL DENSITY AND TEXT COMPREHENSION

Introduction

The general principle that underscores readability analysis of texts is that when a text is written with the number of years of education of the target readers in mind, the readers tend to derive maximum benefit from the text (Flesch, 1949). In other words, they are able to comprehend the text because the text was written with them in mind. Therefore, as comprehension resides with the reader, readability resides with the text. However, there is a supposed relationship between readability and text comprehension because as already stated, when text is readable, comprehension is appreciable (To, 2013). In this chapter, I explore whether there is indeed a relationship between readability and text comprehension. The chapter also explores the relationship that supposedly exists between readability and lexical density. Thus, the researcher sought to determine whether indeed when a text has a high density, text comprehension is reduced or vice versa.

Relationship between Readability and Lexical Density of RAs in the Humanities

Lexical density (LD) refers to the quantity of content vocabulary (expressed in percentages) present in a text. Content Vocabulary are words which have semantic relevance, both in isolation and within the context of the sentence; these words are normally identified as nouns, adjectives, non-auxiliary verbs, and adverbs. The other parts of speech, such as articles,

pronouns, conjunctions, auxiliaries, etc., are categorized as functional words whose roles are linked to grammar and the production of formal text (Read, 2000). As a rule, texts with a lower LD are more easily understood compared to those with higher LD (Ure, 1971; Halliday, 1985). Research has found that high lexical density makes greater demands on memory processes. Sentences with high density (LD towards 100%) are more difficult to reproduce by recall than sentences with low density. Perfetti (1969) confirmed this inverse relation between LD and recall.

Lexical density across departments in the Arts

Table 15 shows a description of lexical density of research articles across departments in the Faculty of Arts.

Table 14: Descriptive statistics of lexical densities of articles from the Faculty of Arts

DEPARTMENTS	Lexical Density				
	n	Mean	SD	Min	Max
Religion and Human Values	20	55.18	4.21	40.70	59.80
English Language	20	53.68	3.18	48.00	59.90
Ghanaian Languages and linguistics	10	55.58	3.61	48.30	60.80
History	10	56.71	3.80	50.90	62.40
Classics and Philosophy	10	55.61	3.87	49.50	62.50
Communication Studies	10	55.30	5.92	47.20	65.50
Music and Dance	10	54.78	4.72	47.50	61.60
African Studies	10	59.19	2.73	55.50	63.30
Overall	100	55.49	4.17	40.7	65.50

n=100; p<0.05

Source: Field Data, Gyasi (2015)

From Table 15, the overall mean LD was 55.49%. The minimum LD was 40.70% while the maximum LD was 65.50%. Department of African Studies recorded the highest mean LD (59%) while Department of English Language recorded the least mean LD (53.6%). The results agree with the findings of a number of research works both in the Arts and other disciplines. For example, Vera, Sotomayor, Bedwell, Dominguez, and Jeldrez (2016) analysed lexical quality and its relation to writing quality for 4th grade, primary school students in Chile. These authors found that close to half of the words were content words (LD), while the other half was functional words. Similarly, Lu (2012) examined LD as a component of lexical richness among Chinese college students who offered English Language as second language by transcribing narratives of passages by these students at four different times. It was found that the LD of this large data set was between 40% and 42%, which is similar to the findings in this study. Besides, To, Fan, and Thomas (2013) also analysed four English Language textbooks for four different levels of students in Tasmania and concluded that three out of these four books had high LD (about 45%). In contrast, however, Gregori-Signes and Clavel-Arroitia (2015) analysed lexical density in the written production of two groups of first year students at the Universitat de València at the beginning and end of one-semester teaching period. The LD reported by these authors ranged between 28% and 37%, which were quite lower compared to the LD reported in this study, suggesting better readability in the writings reported by Gregori-Signes and Clavel-Arroitia (2015) compared to what is being reported here in this study.

Patterns in lexical density across departments in the Arts

Figure 4 shows patterns in lexical density across departments in the Arts. The overall distribution of lexical density of research articles from the Faculty of Arts was skewed left. This implies that there was a larger concentration of LD of research articles from 53% upwards, and a long tail to the left of the histogram. Hence, majority (95% of all research articles) of the research articles were written at lexical density of 50% or more. Only 5% of all research articles from the Arts were written below an LD of 50%. The histogram is centered on lexical density of 58%, implying that more than half of the research articles written from the Faculty were written around an LD of 58%.

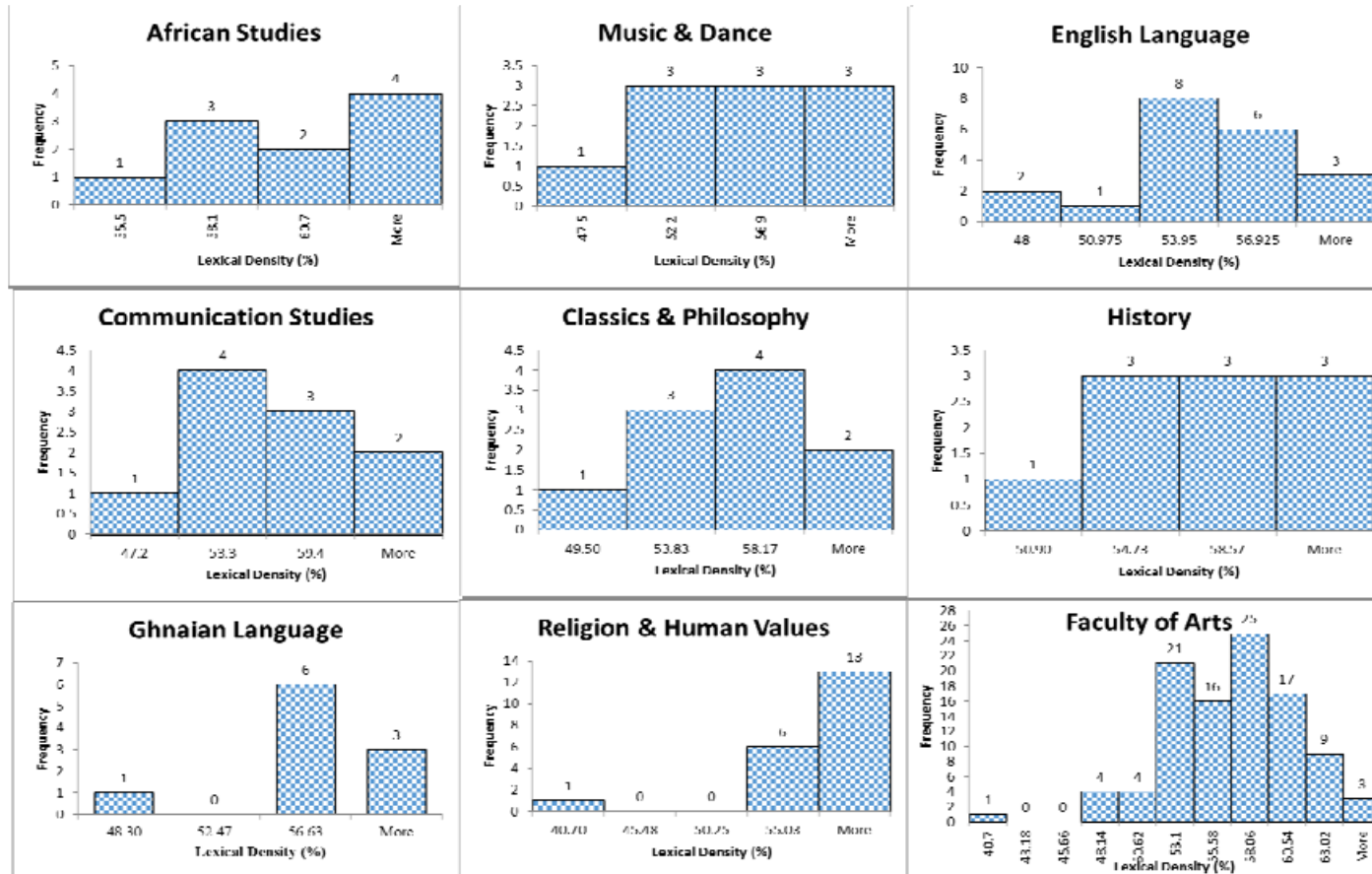


Figure 5: Lexical Density Distribution across the Arts

Again, from Figure 5, majority (60%) of research articles from Department of African Studies produced were written at lexical density of 60% and above. This implies that majority of research articles written scored LD of 60% and above. Similar patterns in LD of research articles were recorded in all the other departments across the Faculty of Arts except Department of Ghanaian Languages and Linguistics, where almost all research articles scored LD of 56%. This means that generally, faculty members in the Arts write research articles that are above the recommended LD. Since lexical density above 50% reduces readability of a text (Tsai, 2010), it can be concluded that faculty members generally produce research articles that are difficult to comprehend when measured in terms of the principle underlying lexical density. Furthermore, that Department of African Studies produced text with a lexical density of 60% and similar trends are followed by the remaining departments implies that departments in the Faculty produced difficult to read research articles when measured from the point of view of LD.

The above findings are similar to what has been reported by Kondal (2015), where 100% of all (10) examination test scripts analysed for tenth class regional medium students in Hyderabad scored LD of 50% and above.

Test of relationship between readability and lexical density

Table 16 shows correlation analysis between readability and lexical density. It is observed from Table 16 that FRE scores, which measures readability, had low and negative correlation with lexical density ($r = -.14$, $p = .159$).

Table 15: Correlation between other Readability indices and Lexical Density

Readability Indices		Lexical Density (%)
FRE	Pearson Correlation	-.142
	Sig. (2-tailed)	.159
	N	100

n=100; p<0.05

Source: Field Data, Gyasi (2015)

The negative correlations between FRE and LD imply that readability decreases with increasing lexical density. In other words, when lexical density of research articles increases, it implies a decreasing of the readability of the research articles. The correlation was however not statistically significant. Therefore, whereas readability index analysis using the Flesch reading ease (FRE) or lexical density analysis can both be used to determine the ease of reading a text (research articles in this case), there is inverse relationship between the resultant scores of the two tools such that the former indicates a low readability when the readability analysis score (FRE) is decreasing, and a higher readability when the readability analysis score is increasing. In the case of lexical density analysis score, however, readability decreases with increasing lexical density score but increase with decreasing lexical density score. The findings from this analysis corroborates what To et al. (2013) reported. After examining the lexical density and readability of four English textbooks from Tasmania, they found no significant and consistent correlation between lexical density and readability, as has been suggested theoretically.

Similarly, Nguyen and Nguyen (2016) investigated lexical density and readability of non-English majored first-year students' writings at a pedagogical university in Vietnam and concluded that the readability score and the lexical density were not really in a corresponding relationship, such that readability decreases with increasing lexical density score and vice versa. The implication here is that the ease of reading research articles published by faculty members in the Arts can be determined through the application of either readability analysis or lexical density analysis. However, the two tools do not necessarily correspond to each other such that the results of applying one of these two tools necessarily correspond to the results of using the other.

Relationship between Readability and Text Comprehension

Readability strongly affects a reader's ability to comprehend what is read (Rello, Baeza-Yates, Dempere-Marco & Saggion, 2013). Aside readability, a reader's interest and learning strategy, among other factors are also known to affect one's comprehension ability (Tin, 2008). Therefore, this objective explores the relationship between readability and text comprehension. The generally accepted hypothesis is that text comprehension increases as readability increases. Therefore, when a text is assigned to an individual or a group on the basis of the text's readability score, it is expected that that individual or group of individuals should be able to comprehend the text with relative ease. In order to determine whether the readability score of a given text really meets the level of reading comprehension of audience who fall within the readability score, a three set of multiple choice comprehension exercises (MCCE) were designed, administered and scored. The analyses of the results are presented in the following table. In addition, in discussing this

objective, the data collected allowed the researcher to compare scores of the MCCE on the basis of gender. Results of this analysis have also been presented below.

The relationship between readability and text comprehension is such that an improved readability of a text is expected to lead to improvement in comprehension. Therefore, the proportion of participants in the MCCE that scored above average was used as cut off point to determine the relationship between readability and text comprehension. Figures 6 – 8 show the pass rate of participants in the MCCE.

It is observed in Figure 6 that majority (66%) of participants scored above average. The text upon which the comprehension exercise was based had a Flesch grade level score of 11 indicating that one needs to acquire 11 years of formal education to comprehend the text. That 66% participants scored above average in the MCCE indicates that the text was to a reasonable extent readable to majority of the participants. Based on this result therefore, it can be concluded that some appreciable relationship exists between readability and text.

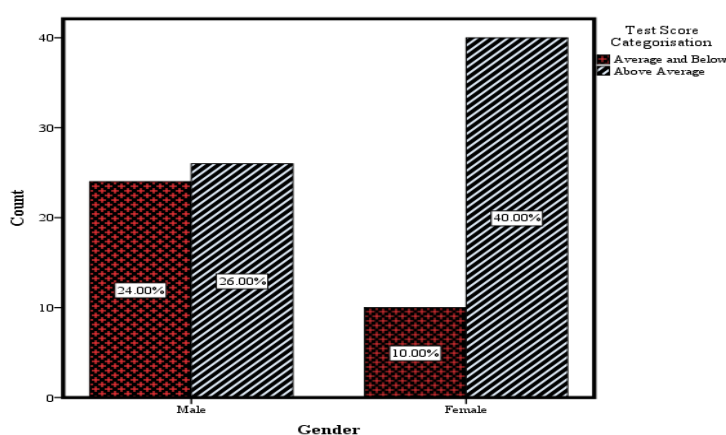


Figure 6: Bar Graph showing the pass rate of participants in MCCE 1

Figure 6 shows that all but one (99%) of participants scored above average in the comprehension exercise. The implication is that almost all the participants demonstrated sound comprehension of the text upon which the exercise was based. It can therefore be concluded that there is a perfect relationship between readability and text comprehension as indicated in MCCE2.

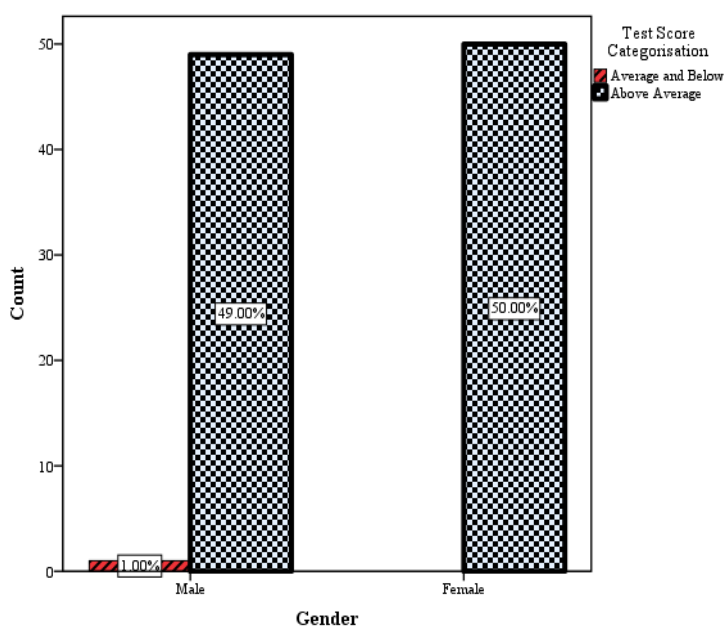
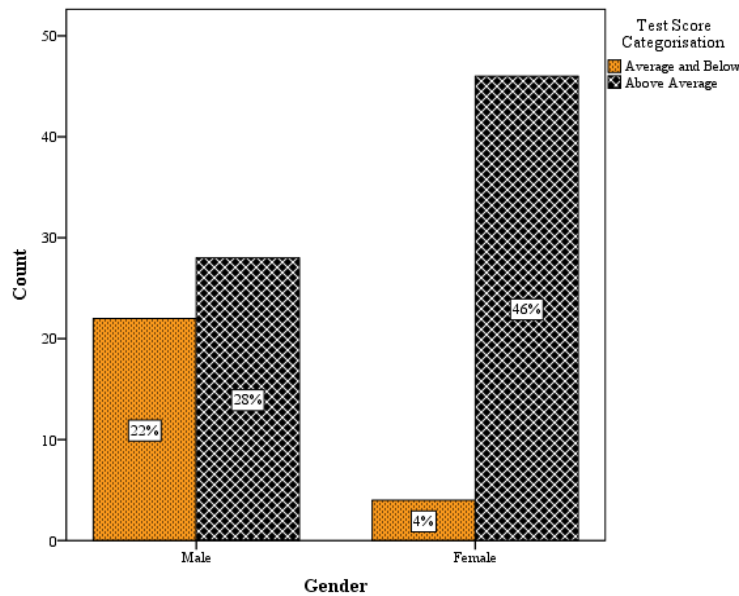


Figure 7: Bar Graph showing the pass rate of participants in MCCE 2

Finally, Figure 7 shows that majority (74%) of participants scored above average in the comprehension exercise. Like the above two results, this result portrays that majority of the participants understood or comprehended the text upon which comprehension exercise 3 was based. This can be seen in the fact that 74% of respondents in the MCCE scored above 50%. Thus, a clear positive relationship between readability and comprehension is evident.

One might be tempted to argue that since results of the comprehension test were high even when respondents were picked from SHS 2, 3, and third year undergraduate students of the university, the RAs are arguably not difficult contrary to the readability test scores. It is important to mention that the comprehension exercises were based on only the introductory sections of the RAs selected and the Flesch Reading Ease scores of these sections corresponded to the number of years of schooling (grade level) of respondents. The readability level of the RAs, on the other hand, was based on three sections of the RAs (introduction, methodology and conclusion). Besides, it can also be argued that the mode of selection of participants for the comprehension exercise could have resulted in selecting only respondents who have high literacy performance since the sampling procedure was convenience (used those that were available at the time of selection). Therefore, if respondents (e.g. third year undergraduate students) were selected from all programs to participate in the exercise, the results could have been different. Additionally, the mode of categorizing the performance of the respondents could have been a factor responsible for the seeming high performance. The researcher used two categories (above 50% and below 50%) to interpret the results. In this case, all respondents who scored 50% and above were considered to have understood the text whereas respondents who scored below 50% were considered to have performed poorly. These two parameters of assessment of performance of respondents (below and above average) were based on the grading system of the University of Cape Coast, the research site. The University of Cape Coast grading system of students in examination is such that those who score 50% and above are considered to have passed

whereas those who score below 50% are considered to have failed. Furthermore, if the comprehension tests were administered to experts for whom the RAs are meant for, they will undoubtedly perform excellently because the FRE scores of the texts upon which the comprehension exercises were based were far below the years of schooling of these experts.



**Figure 8: Bar Graph showing the pass rate of participants in MCCE 3
Gender differences in text comprehension**

Table 17 – 19 presents results of the three MCCES between males and females. The results show proportions of males and females that scored average and below in the test one hand and those who scored above average on the other hand.

Table 16: Gender difference in Text Comprehension for Test 1

			MCCE		Total
			Av. and Below	Above Av.	
Gender	Male	Count	24	26	50
		Expected Count	17.0	33.0	50.0
		% within Gender	48.0%	52.0%	100.0%
		% within MCCE	70.6%	39.4%	50.0%
	Female	Count	10	40	50
		Expected Count	17.0	33.0	50.0
		% within Gender	20.0%	80.0%	100.0%
		% within MCCE	29.4%	60.6%	50.0%
Total		Count	34	66	100
		Expected Count	34.0	66.0	100.0
		% within Gender	34.0%	66.0%	100.0%
		% within MCCE	100.0%	100.0%	100.0%

n=100

Source: Field Data, Gyasi (2015)

It is observed from Table 17 that in total, 34 (34% of the total number of participants who took the test) participants scored average or below in the MCCE. Out of this, 24 (70.6% of those who scored average or below) were males while 10 (29.4% of those who scored average or below) were females. Further, 66 participants scored above average, out of which 26 (39.4% of those who scored above average) were males while 40 (60.6% of those who scored above average) were females. It implies that more females

comprehended the test than males. The difference in proportion of female participants who performed above average in the test was close to double the number of males who performed above average in the test.

Gender differences in text comprehension in the second test are presented in Table 18.

Table 17: Gender difference in Text Comprehension for Test 2

			MCCE		Total
			Av. and Below	Above Av.	
Gender	Male	Count	1	49	50
		Expected Count	.5	49.5	50.0
		% within Gender	2.0%	98.0%	100.0%
		% within MCCE	100.0%	49.5%	50.0%
	Female	Count	0	50	50
		Expected Count	.5	49.5	50.0
		% within Gender	0.0%	100.0%	100.0%
		% within MCCE	0.0%	50.5%	50.0%
Total		Count	1	99	100
		Expected Count	1.0	99.0	100.0
		% within Gender	1.0%	99.0%	100.0%
		% within MCCE	100.0%	100.0%	100.0%

n=100; Source: Field Data, Gyasi (2015)

From Table 18, It is observed that in total, only a single male participant (1% of the total number of students who took the test) scored average or below in the MCCE. Further, 99 participants scored above average, out of which 49 (49.5% of those who scored above average) were males while

40 (50.5% of those who scored above average) were females. It implies that both males and females comprehended the test 2 equally. Gender differences in text comprehension in the third test are presented in Table 19.

Table 18: Gender difference in Text Comprehension for Test 3

	Gender	Male	Count	MCCE		Total
				Av. and Below	Above Av.	
	Male		Count	22	28	50
			Expected Count	13.0	37.0	50.0
			% within Gender	44.0%	56.0%	100.0%
			% within MCCE	84.6%	37.8%	50.0%
	Female		Count	4	46	50
			Expected Count	13.0	37.0	50.0
			% within Gender	8.0%	92.0%	100.0%
			% within MCCE	15.4%	62.2%	50.0%
Total			Count	26	74	100
			Expected Count	26.0	74.0	100.0
			% within Gender	26.0%	74.0%	100.0%
			% within MCCE	100.0%	100.0%	100.0%

n=100; Source: Field Data, Gyasi (2015)

It is observed from Table 19 that in total, 26 (26% of the total number of participants who took the test) participants scored average or below. Out of this, 22 (84.6% of those who scored average or below) were males while 4 (15.4% of those who scored average or below) were females. Further, 74 participants scored above average, out of which 28 (37.8% of those who scored above average) were males while 46 (62.2% of those who scored above

average) were females. It implies that more females comprehended test 3 than males. The difference in proportion of female participants who performed above average on the test was close to two-thirds the number of males who performed above average on the test.

Test of significant differences in text comprehension between Gender

Table 20 presents chi-square analysis which was used to evaluate significant difference(s) in male and female participants' level of comprehension of the texts.

Table 19: Chi-square Analysis of Text Comprehension between Gender

Test	Statistic	Value	Df	Asym. Sig. (2-sided)	Exact Sig. (2-sided)
1	Continuity Correct.	7.53	1	.006	
	Phi	.296		.003	
2	Continuity Correct.	.00	1	1.00	
	Fisher's Exact Test				1.00
	Phi	.101		.315	
3	Continuity Correct.	15.02	1	.000	
	Phi	.410		.000	

n=100; p<0.05
Source: Field Data, Gyasi (2015)

Three statistical tests (Yate's continuity correction, Fisher's exact test, and phi) are reported in Table 20. Both Yate's continuity correction and Fisher's exact test measure the same entity (difference in comprehension, measured by participants score on the MCCE, between males and females). However, different assumptions underline each statistic such that a violation

of assumptions that underpin the use of either of the two statistics renders any inference from it invalid. For example Yate's continuity correction test is the appropriate test for 2x2 crosstabulation since it is designed to compensate for overestimation of the traditional Pearson chi-square value when used with a 2x2 table (Field, 2011; Pallant, 2013). For this reason, Pearson's chi-square test was not used although it is the conventional chi-square analytical technique. Also, another assumption that underpins chi-square analysis is that the lowest expected frequency in any cell should be 5 or more. When this assumption is violated, Fisher's Exact Probability Test is recommended since it compensates for the violation caused by cells that record expected frequency of less than 5. From Table 17, it is observed that the expected frequency of males and females recorded values (0.5) which were less than 5. Hence, Fisher's exact test is reported in the test of comprehension in the second test (Table 20). Phi is used to estimate the effect size of the difference(s) in comprehension levels between males and females, where applicable. This measure indicates the practical importance of differences that are established using Yate's continuity correction or Fisher's exact test, such that the higher (towards 1) the phi score, the greater the practical importance and vice versa.

It is observed from Table 20 that for the first test of comprehension, there was statistically significant association between gender and comprehension level [$\chi^2(1, n=100) = 7.53; p = 0.006, \phi = .296$]. This implies that the proportion of female participants (40%) that comprehended the MCCE was higher compared to males (26%). Hence, females showed better comprehension than males. The effect size ($\phi = .296$) was moderate (Pallant,

2013) indicating that female participants' comprehension of test 1 was moderate in difference compared to that of male participants.

In the second comprehension test, there was no statistically significant association between gender and comprehension level [Fisher's $p = 1.00$, $\phi = .101$]. This implies that the proportion of females (50%) that comprehended the MCCE was not greater compared to males (49%). Hence, neither males nor females showed better comprehension than the other. The effect size ($\phi = .101$) was small and not statistically significant.

In the third comprehension test, there was statistically significant association between gender and comprehension level [$\chi^2(1, n=100) = 15.02$; $p = 0.000$, $\phi = .410$]. This implies that the proportion of females (46%) that comprehended the MCCE was higher compared to males (26%). Hence, females showed better comprehension than males. The effect size ($\phi = .410$) was moderate indicating that female participants' comprehension of test 3 was moderate in difference compared to that of male participants.

For categorical data (such as has been used in this objective), a more useful measure of effect size is the odds ratio. The odds ratio represents 'the change in odds of being in one of the categories of outcome when the value of a predictor increases by one unit' (Tabachnick & Fidell, 2013). Hence, from Table 20, the odds of scoring above average in the MCCE in the first comprehension test was 3.69 times if the participant was female than if the participant was male. Similarly, in the third comprehension test, the odds ratio was found to be 9.03. This implies that the odds of scoring above average in the MCCE in the third comprehension test was 9.03 times if the participant was female than if the participant was male.

Table 20: The odds of scoring above average in the MCCE between males and Females

Test	Odds of male	Odds of female	Odds Ratio Exp. (B)
1	0.92	0.25	3.69
2	0	0	
3	0.79	0.09	9.03

These findings parallel a number of earlier research findings that evaluated gender differences in text comprehension. The overwhelming majority of all empirical research that considered differences in comprehension between males and females have largely favoured females (Logan & Johnston, 2010). For example, Maccoboy and Jacklin (1974) affirmed that females are superior in verbal skills than males. Similarly, previous international studies, which included 35 to 40 countries, examined reading comprehension with 10-year-old children and found gender differences favouring girls in every participating country (Mullis et al. 2003; Mullis et al. 2007). In addition, Chui and McBride-Chang (2006) showed that this gender difference in reading continues into adolescence, as a study with 15-year-old children in 43 countries found that in every country, girls outperformed boys. These gender differences generally appear regardless of the type of reading instruction children have received (Johnston and Watson 2005; Johnston, Watson, and Logan 2009), or the writing system: whether an alphabetic (transparent or opaque) or ideographic orthography (Mullis et al. 2003; Chui and McBride-Chang 2006; Mullis et al. 2007). Earlier works on neuroimaging have suggested that adult males and females display different

patterns of functional activation during reading (Shaywitz et al. 1995; Pugh et al. 1996) and sex differences in children have been found in the localization of brain activation during word reading (Burman, Bitan, and Booth 2008). Aside these differences brain activation between males and females with respect to reading, different reading strategies, motivation difference, between males and females has been suggested to influence comprehension differences between males and females (Thompson 1987). Therefore, for texts of equal readability, it is expected that females will have high verbal efficiency compared to males, in most cases. This high verbal efficiency or word fluency translates into better comprehension in females, than it will be for males.

CHAPTER EIGHT

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Introduction

This chapter presents the summary, conclusions and recommendations of the study. Summary of the results and conclusions have been organized based on the specific objectives and the hypotheses of the study. This section also presents suggested areas for further research.

Summary

Writing for the academic community needs careful planning and consideration if what is written can achieve its intended purpose of transferring knowledge. For instance, time is a very important commodity in the academic community; therefore, members of the community will appreciate it if they do not spend so much time just trying to comprehend a text because its language is difficult.

Recently, there has been agitation among major stakeholders such as parents, students and teachers about the downward trend of the linguistic competence of students, especially in Ghanaian colleges and universities. The implication is that academic communication between teachers and students is likely to be hampered, especially in the case of written communication. This means that any form of written communication targeted at members of the academic community should be written in plain language to enhance comprehension and save time.

However, there are indications that research articles pose problems to readers in terms of readability. There are evidences from several studies that

both students and lecturers in a number of university communities have complained of experiencing difficulty with understanding research articles prepared by lecturers in various institutions. This pattern is exhibited in all RAs regardless of the geographic background of the author. Similar challenge is expected in RAs authored by Ghanaian scholars in all disciplines including the Arts where people belong are generally perceived to have mastery on language and communication.

The objective therefore of this study was to determine the readability of research articles in the Arts of the University of Cape Coast (UCC). Specifically, the study sought to accomplish the following:

1. To determine the level of readability of research articles in the Arts in the University Cape Coast.
2. To determine the differences in the readability across disciplines/ in the Arts.
3. To determine the differences in readability of female and male authored research articles in the Arts.
4. To explore the relationship between readability and lexical density of research articles in the Arts.
5. To explore the relationship between readability and text comprehension.

The study was mainly quantitative as the descriptive research design was employed for the study. The target population was the entire population of lecturers from the Faculty of Arts of the University of Cape Coast. These lecturers were active in publishing research findings (N=123). Of this, 19 were Professors, 23 Senior Lecturers, 60 Lecturers and 21 Assistant Lecturers. A

stratified random sampling technique was used to sample authors of RAs from each of the eight (8) Departments of the Faculty of Arts. Research Articles from twenty authors each from the Departments of Religion & Human Values, and English were selected. In addition, RAs from ten (10) authors each from the Ghanaian Language, Music & Dance, History, Classics & Philosophy, Communication Studies, and African Studies were also randomly sampled. Readability indices were computed using aggregate scores from three online readability calculators www.readabilityformulas.com, www.usingenglish.com, and www.dairyscience.info. Results were analyzed using measures of central tendencies and dispersions, frequencies and percentage distributions, Wilcoxon signed ranked, Kruskal – Wallis H test and Mann-Whitney U test. In addition, Spearman rho correlation coefficient, histogram, and percentages 2×2 crosstabulation, Yate’s continuity correction, Fisher’s exact test, phi and odds ratio were employed. The summary of major findings in relation to the specific objectives of the study is as follows:

How readable are the research articles in the Arts when measured in terms of FRE?

The majority (63%) of the research articles were graded as ‘*difficult*’ to read when measured in terms of FRE. Hence, a reader requires some college level of education to be able to read and comprehend RAs in the Arts. However, since RAs are generally written by experts for experts, this finding can only be considered in terms of the FRE interpretations. Thus, experts who read these RAs may not necessarily find them difficult to comprehend. In terms of syntactic features of RAs in the Arts, the minimum mean length of sentence written by authors from the Faculty of Arts was approximately 21 (\bar{x})

= 20.77, SD = 3.25) which was above the maximum mean sentence length (20 words) recommended for ease of reading. The Department of Religion and Human Values wrote the least mean number of words per sentence (M = 20.77; SD = 3.25) while the Department of English recorded the longest mean number of words per sentence (M = 27.89; SD = 6.68).

Further, Wilcoxon signed-rank test showed that the median FRE score (40.10) for the Faculty of Arts was below the standard readability score of 60 ($z = 15.01$, $p = 0.000$) at the 0.05 alpha level of significance. This difference (19.90) was statistically different from the standard FRE score of 60. This implies that over half of research articles in the Faculty of Arts are above the 'standard' readability level. Hence, the majority of the authors' research articles were relatively difficult to read side by side the 'standard' readability level.

Differences in the Readability of Research Articles across Disciplines/ Departments in the Arts

Of the eight departments (Music & Dance, English, Communication Studies, Classics & Philosophy, African Studies, History, Ghanaian Language Religion and Human Values), the Department of Religion and Human Values (mean FRE rank score = 76.55) produced research articles which were the easiest to read while the Department of Music and Dance (mean FRE rank score = 32.70) wrote the most difficult to read research articles.

In addition, Kruskal-Wallis H test showed that there were significant differences in the FRE scores across the eight (8) Departments [$H(7, n = 100) = 33.945$; $p = 0.000$] at the 0.05 alpha level. Among all the eight departments,

there were significant differences in readability of RAs of Department of Religion and Human Values (Median FRE score = 48.75) and the following departments: Department of Music & Dance (Median FRE score =32.60), Department of English (Median FRE score = 32.00) and the Department of Communication Studies (Median FRE score =37.10).

Gender variation in readability of research articles

The majority of the articles were authored by males (94%). For male authors, half of all RAs were written at a readability level classified as 'fairly difficult' while half of RAs by female authors were graded as standard. A Mann-Whitney U Test revealed no significant difference in the readability of RAs of males (Md = 49.87, n =94) and females (Md = 60.33, n = 6), $U = 171.00$, $z = -1.61$, $p = 0.11$.

Relationship between readability and lexical density of research articles in the Arts

The overall mean LD was found to be 55.49%. The minimum LD was 40.70% while the maximum LD was 65.50%. The Department of African Studies recorded the highest mean LD (59%) while the Department of English recorded the least mean LD (53.6%). The overall distribution of lexical density of research articles from the Faculty of Arts was skewed left. The majority (95% of all research articles) of the research articles were written at lexical density of 50% or more. Only 5% of all research articles from the Arts were written below an LD of 50%. The majority (60%) of research articles from Department of African Studies produced were written at lexical density of 60% and above. Similar pattern in LD of research articles were recorded in

all the other departments across the Faculty of Arts except Department of Ghanaian Languages and Linguistics, where almost all research articles scored LD of 56%.

Further, it was found that FRE scores, which measures readability, had low and negative correlation with lexical density ($r = -.14$, $p = .159$) which was statistically insignificant.

Relationship between readability and text comprehension

In the first multiple choice comprehension exercise, the majority (66%) of participants scored above average. Out of this, 26 (39.4% of those who scored above average) were males while 40 (60.6% of those who scored above average) were females. In the second exercise, 99% scored above average out of which 49 (49.5% of those who scored above average) were males while 40 (50.5% of those who scored above average) were females. In the third exercise, 74% of participants scored above average, out of which 28 (37.8% of those who scored above average) were males while 46 (62.2% of those who scored above average) were females.

In addition, for the first test of comprehension, there was statistically significant association between gender and comprehension level [$\chi^2 (1, n=100) = 7.53$; $p = 0.006$, $\phi = .296$]. In the second comprehension test, there was no statistically significant association between gender and comprehension level [Fisher's $p = 1.00$, $\phi = .101$]. In the third comprehension test, there was statistically significant association between gender and comprehension level [$\chi^2 (1, n=100) = 15.02$; $p = 0.000$, $\phi = .410$].

Conclusions

The following conclusions were drawn from the findings based on the specific objectives: It can be concluded from the findings of objective one that research articles in the Arts are generally '*difficult*' to read when measured in terms of the Flesch Reading Ease such that one required some college level of education to be able to read and understand these articles. The difficulty in reading RAs from the Arts is attributed to longer than recommended sentence length. The Department of Religion and Human Values wrote the least mean number of words per sentence while the Department of English recorded the longest mean number of words per sentence.

From the findings of objective two, it can be concluded that the Department of Religion and Human Values produced research articles which were the easiest to read side by side the other seven departments while the Department of Music and Dance wrote the most difficult to read research articles. The Department of Music and Dance wrote the most difficult to read RAs probably because of the use of complex grammatical structures borrowed from other cultures since the study of music and dance cuts across not only those of Ghanaian origins but also those of other cultures across the world.

It can also be concluded from objective three that RAs in the Arts are of the same readability level regardless of the gender of the author such that neither males nor female authors wrote more difficult RAs. Basically then, both male and female authors performed equally well with regard to readability of academic writings. From objective four, it can be concluded that readability decreased with increasing lexical density. The correlation was however not statistically significant.

Finally, from objective five, it can be concluded that generally, high readability of RAs was related to improved text comprehension. Also, females recorded better comprehension as against males such that the odds of scoring above average on the comprehension exercise were about six times better for females than it was for males.

Recommendations

Based on the conclusions of the study, the following recommendations were made for consideration to improve readability of RAs produced in the Faculty of Arts of the University of Cape Coast. Authors must use plain and concise language when they write. Over the whole document, it is the standard recommendation to authors to write, on the average, sentences of 15 – 20 words. This is so because readers recoil when they see long sentences running across a page. In addition, longer sentences require greater mental effort (working memory) to be able to process and understand compared to shorter sentences. Writing shorter sentences can be achieved by the use of lists to break longer sentences into more visual chunks.

Suggestions for Further Research

1. It has been indicated in literature that authors' sociodemographic background such as ethnicity, place of birth and growth during early years, and level of education affects the individual's writing ability and style. Since how readable a text is largely a reflection of the author's writing style, it is recommended that subsequent research should isolate author's sociodemographic background information to be used as control.

2. Given that text formality is influenced by several factors including the use of polysyllabic words and complex grammatical structures, it will be worthwhile to investigate the relationship between readability and text formality since the use of polysyllabic words and complex grammatical structures also results in low text readability.
3. The sample size for males (94) and females (6) in this study was widely unequal such that findings based on these sample sizes may be biased and possibly not reflect the true relationship between gender of author and the readability of RAs. Hence, it is recommended that subsequent researches should be designed such that the sample size of RAs for males and females will be proportional to the RAs published across gender.
4. This study focused on the readability of research articles in the Arts by using eight disciplines in the Faculty of Arts in the University of Cape Coast. Another study can be conducted to cover other faculties like social science, education, and the natural sciences to determine the nature of the phenomenon in these areas.

REFERENCES

- Adam, E. M., Eltorai, A., Han, A. B., Truntzer, J. M., & Daniels, A. H. (2015). Readability of patient education materials on the American orthopaedic society for sports medicine website. *The Physician and Sportsmedicine*, *42*(4), 125–130. Retrieved from: <https://doi.org/10.3810/psm.2014.11.2099>
- Ahmed, I., Zeeshan Shaukat, M., & Islam, T. (2013). Mission statements readability: An insight into Islamic banks. *Journal of Islamic Accounting and Business Research*, *4*(2), 132–150.
- AlKhalili, R., Shukla, P. A., Patel, R. H., Sanghvi, S., & Hubbi, B. (2015). Readability assessment of internet-based patient education materials related to mammography for breast cancer screening. *Academic Radiology*, *22*(3), 290–295.
- Altman, D. G. (2015). Making research articles fit for purpose : structured reporting of key methods and findings. *Altman Trials*, *16*(53), 15–17.
- Armstrong, J. S. (1988). Readability and prestige in scientific journals. *Journal of Information Science*, *14*(2), 69–75.
- Badley, G. (2013). Academic writing: Contested knowledge in the making? *Quality Assurance in Education*, *17*(2), 104–117.
- Bailin, A., & Grafstein, A. (2001). The linguistic assumptions underlying readability formulae: a critique, *21*, 285–301.
- Bailin, A., & Grafstein, A. (2016). *Readability: Text and context* (Vol. 1). Hampshire: Palgrave Macmillan.

- Baker III, H. E., & Kare, D. D. (1992). Relationship between annual report readability and corporate financial performance. *Management Research News*, 15(1), 1–4.
- Bali, R., & Hovakimian, A. (1999). Financial education association. *Journal of Financial Education*, 26, 91–111.
- Bauerly, R. J., Johnson, D. T., & Singh, M. (2006). Readability and writing well. *Marketing Management Journal*, 16(1), 220–225.
- Benjamin, R. G. (2012). Reconstructing readability: Recent developments and recommendations in the analysis of text difficulty. *Educational Psychology Review*, 24(1), 63–88. Retrieved from <https://doi.org/10.1007/s10648-011-9181-8>
- Blessing, L. T. M., & Chakrabarti, A. (2009). Descriptive study I: Understanding design. *DRM, A Design Research Methodology*.
- Britton, B. K. (1991). Using kintsch's computational model to improve instructional Text: Effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psychology*. 83(3), 329–345.
- Buccini, L. D., Iverson, D., & Caputi, P. (2010). An Australian based study on the readability of HIV / AIDS and type 2 diabetes clinical trial informed consent documents. *Journal of Bioethical Inquiry*. 7(3), 313–319.
- Carol, S. R., & Hosei, B. (2008). Printed health information materials: Evaluation of readability and suitability. *Journal of Community Health Nursing*, 25(2), 73–90.

- Cheng, C., & Dunn, M. (2015). Health literacy and the Internet : A study on the readability of Australian online health information. *Australian and New Zealand Journal of Public Health*, 39(4), 309–314.
- Collins-Thompson, K., & Callan, J. (2005). Predicting reading difficulty with statistical language models. *Journal of the American Society for Information Science and Technology*, 56(13), 1448–1462.
- Corcoran, N., & Ahmad, F. (2016). Patient education and counseling: The readability and suitability of sexual health promotion leaflets. *Patient Education and Counseling*, 99(2), 284–286. Retrieved from: <https://doi.org/10.1016/j.pedc.2015.09.003>
- Cornett, H. E. (2014). Gender differences in syntactic development among English speaking adolescents. *Inquiries Journal*, 6(3). 127-137
- Cortez, S., Milbrandt, M., Kaphingst, K., James, A., & Colditz, G. (2015). The readability of online breast cancer risk assessment tools. *Breast Cancer Research and Treatment*, 154(1), 191–199. Retrieved from: <https://doi.org/10.1007/s10549-015-3601-2>
- Cutts, M. (2013). *Oxford guide to plain English*. Oxford: Oxford University Press. 4(53). Retrieved from: <https://doi.org/10.1017/CBO9781107415324.004>
- Darko, K. (2003). *Ghanaianisms: A glossary*. Accra, Ghana University Press.
- Dolnicar, S., & Chapple, A. (2014). The readability of articles in tourism journals. *Annals of Tourism Research*, 52, 161–179. Retrieved from: <https://doi.org/10.1016/j.annals.2015.03.003>
- Dubay, W. H. (2004). *The principles of readability*. California: Plain-Language Services.

- Edgley, A., Stickley, T., Timmons, S., & Meal, A. (2016). Critical realist review : Exploring the real, beyond the empirical. *Journal of Further and Higher Education, 40*(3), 316–330. Retrieved from: <https://doi.org/10.1080/0309877X.2014.953458>
- Evans, H., Chao, M. G., Leone, C. M., Finney, M., & Fraser, A. (2016). Content analysis of web-based norovirus education materials targeting consumers who handle food: An assessment of alignment and readability. *Food Control, 65*, 32–36. Retrieved from: <https://doi.org/10.1016/j.foodcont.2016.01.003>
- Fakhfakh, M. (2015). The readability of international illustration of auditor's report: An advanced reflection on the compromise between normative principles and linguistic requirements. *Journal of Economics, Finance and Administrative Science, 20*(38), 1–9. Retrieved from: <https://doi.org/10.1016/j.jefas.2015.02.001>
- Field, A. (2011). *Discovering statistics using* (3rd ed.). Washington, DC: Sage Publications, Inc.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics. Statistics* (Vol. 58). Retrieved from: <https://doi.org/10.1016/B978-012691360-6/50012-4>
- Fosu, M. (2014). *The press and political participation: newspapers and the politics of linguistic exclusion and inclusion in Ghana*. University of Leeds.

- García-Merino, T., & Santos-Lvarez, V. (2009). Characterizing the top journals in strategic management: Orientation, style, originality, and readability. *Journal of Scholarly Publishing*, 40(4), 420–439. Retrieved from: <https://doi.org/10.3138/jsp.40.4.420>
- Gazni, A. (2011). Are the abstracts of high impact articles more readable? Investigating the evidence from top research institutions in the world. *Journal of Information Science*, 37(3), 273–281. Retrieved from: <https://doi.org/10.1177/0165551511401658>
- Gizir, S., & Simsek, H. (2005). Communication in an academic context. 197–221. Retrieved from: <https://doi.org/10.1007/s10734-004-6349-x>
- Gorsuch, G. (2013). Evidence from longitudinal student reports longitudinal student reports. Retrieved from: <https://doi.org/10.1177/1362168809346494>
- Grant, M. J., Booth, A., & Centre, S. (2009). A typology of reviews: An analysis of 14 review types and, 91–108. Retrieved from: <https://doi.org/10.1111/j.1471-1842.2009.00848.x>
- Gregori-Signes, C., & Clavel-Arroitia, B. (2015). Analysing lexical density and lexical diversity in university students' written discourse. *Procedia - Social and Behavioral Sciences*, 198(Cilc), 546–556. Retrieved from: <https://doi.org/10.1016/j.sbspro.2015.07.477>
- Gupta, R. (2014). Expectations vs. Reality: The Readability of texts in the primary grades in India. *Procedia - Social and Behavioral Sciences*, 116, 3916–3920. Retrieved from: <https://doi.org/10.1016/j.sbspro.2014.01.866>

- Gyasi, W. K. (2011). An Analysis of the readability of the essays of first year students of Ghanaian universities - The case of University of Cape Coast. *Language in India*, 11, 92–101. Retrieved from [http://ejournal.narotama.ac.id/files/LANGUAGE IN INDIA.doc](http://ejournal.narotama.ac.id/files/LANGUAGE%20IN%20INDIA.doc)
- Gyasi, W. K. (2013a). Readability and academic communication: A comparative study of undergraduate students' and handbook of three Ghanaian universities. *IOSR Journal of Computer Engineering*, 13(6), 41–50. Retrieved from: <https://doi.org/10.9790/0661-1364150>
- Gyasi, W. K. (2013b). Readability and health communication: An analysis of the readability of commonly used malaria drugs information leaflets in Cape Coast, Ghana. *Journal of Research & Method in Education*, 2(4), 17–25. Retrieved from <http://www.iosrjournals.org/iosr-jrme/papers/Vol-2Issue-4/D0241725.pdf?id=3423>
- Gyasi, W. K. (2013c). The Role of readability in science education in Ghana : A readability index analysis of Ghana association of science teachers textbooks for senior high school, 2(1), 9–19.
- Hadden, K., Prince, L. Y., Schnaekel, A., Couch, C. G., Stephenson, J. M., & Wyrick, T. O. (2016). Readability of patient education materials. *Journal of Hand Surgery*, 1–8. Retrieved from: <https://doi.org/10.1016/j.jhsa.2016.05.006>
- Hall, J. C. (2006). The readability of original articles in surgical journals. *Original Article*. Retrieved from: <https://doi:10.1111/j.1445-2197.2006.03651.x>. (76), 68–70
- Hall, S. (2017). Definition of a research article. Retrieved February 10, 2017, from: <http://penandthepad.com/definition-research-article-2711.html>

- Hartley, J. A. H., Pennebaker, J. A. W. P., & Fox, C. L. F. (2003). Abstracts , introductions and discussions: How far do they differ in style? *Scientometrics*, 57(3), 389–398.
- Hartley, J., & Knapper, C. K. (2008). Academics and their writing. *Studies in Higher Education*, 9(2), 161–164. Retrieved from: <https://doi.org/10.1080/03075078412331378814>
- Hartley, J., Pennebaker, J. W., & Fox, C. (2003). Using new technology to assess the academic writing styles of male and female pairs and individuals. *Journal of Technical Writing and Communication*, 33(3), 243–261.
- Hayden, J. D. (2008). Readability in the British journal of surgery. *British Journal of Surgery*, 95(1), 119–124. Retrieved from: <https://doi.org/10.1002/bjs.5994>
- Hengel, E. (2015). Publishing while female: Gender differences in peer review scrutiny (December), 1–64.
- Humphreys, A. H., & Humphreys, J. T. (2013). Reading difficulty levels of selected articles in the journal of research in music education and journal of historical research in music education. *Music Education Research International*, 6, 15–25.
- James, H., Sotto, E., & Fox, C. (2004). Clarity across the disciplines: An analysis of texts in the sciences, social sciences, and arts and humanities. *Science Communication*, 26(2), 188–210. Retrieved from: <https://doi.org/10.1177/1075547004270164>

- Jaroongkhongdach, W., Watson, R., Keyuravong, S., & Hall, D. (2012). Differences in quality between Thai and international research articles in ELT. *Journal of English for Academic Purposes*, 11(3), 194–209. <https://doi.org/10.1016/j.jeap.2012.04.006>
- Johansson, V. (2008). Lexical diversity and lexical density in speech and writing: A developmental perspective. *Working Papers*, 53, 61–79.
- Joubert, P. H., & Rogers, S. M. (2015). *Strategic scientific and medical writing: The road to success*. Written Communication in Academic Settings. In (pp. 13–23). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from: https://doi.org/10.1007/978-3-662-48316-9_3
- Kampourakis, K. (2016). Publish or Perish? *Science & Education*, 25(3), 249–250. Retrieved from: <https://doi.org/10.1007/s11191-016-9828-4>
- Kasabwala, K., Misra, P., Hansberry, D. R., Agarwal, N., Baredes, S., Setzen, M., & Anderson Eloy, J. (2013). Readability assessment of the American rhinologic society patient education materials. *International Forum of Allergy and Rhinology*, 3(4), 325–333. Retrieved from: <https://doi.org/10.1002/alr.21097>
- Kondal, B. (2015). Effects of lexical density and lexical variety in language performance and proficiency. *International Journal of IT, Engineering and Applied Sciences Research*, 4(10), 25–29.
- Lee, S., & French, N. (2011). The readability of academic papers in the journal of property investment & finance. *Journal of Property Investment and Finance*, 29(6), 693–704. Retrieved from: <https://doi.org/10.1080/14635781111171814>

- Li, H., Graesser, A. C., Conley, M., Cai, Z., Jr, P. I. P., & James, W. (2015). A new measure of text formality: An analysis of discourse of Mao Zedong. *Discourse Processes*, 37–41. Retrieved from: <https://doi.org/10.1080/0163853X.2015.1010191>
- Lim, A. (2010). The readability of information literacy content on academic library web sites. *Journal of Academic Librarianship*, 36(4), 296–303. Retrieved from: <https://doi.org/10.1016/j.acalib.2010.05.003>
- Logan, S., & Johnston, R. (2010). Investigating gender differences in reading. *Educational Review*, 62(2), 175–187. Retrieved from: <https://doi.org/10.1080/00131911003637006>
- Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal*, 96, 190–208. Retrieved from: <https://doi.org/10.1111/j.1540-4781.2011.01232.x>
- Majumder, K. (2015). 6 Article types that journals publish: A guide for early career researchers. *Editage Insights (20-02-2015)*.
- Marchevsky, D. (2000). Sampling and design. In *Appraisal of medical literature* (pp. 43–55). New York: Plenum Publishers.
- Mavasoglu, M., & Dincer, S. (2014). Readability and French language teaching texts: An analysis of french language teaching websites and textbooks. *Procedia - Social and Behavioral Sciences*, 116, 256–259. Retrieved from: <https://doi.org/10.1016/j.sbspro.2014.01.204>
- McInnes, N., & Haglund, B. O. J. A. (2011). Readability of online health information: Implications for health literacy, 36, 173–189. Retrieved from: <https://doi.org/10.3109/17538157.2010.542529>

- Morony, S., Flynn, M., Mccaffery, K. J., Jansen, J., & Webster, A. C. (2015). Readability of written materials for CKD patients: A systematic review. *American Journal of Kidney Diseases*, 65(6), 842–850. Retrieved from: <https://doi.org/10.1053/j.ajkd.2014.11.025>
- Nemati, A., & Bayer, J. M. (2007). Gender differences in the use of linguistic forms in the speech of men and women: A comparative study of Persian and English. *Language in Society*, 21, 547–568.
- Nguyen, N., & Nguyen, T. (2016). An investigation into the lexical density and readability of non-English majored first-year students' writings at a pedagogical university in Vietnamese context. In *SEAMEO RETRAC 7th International TESOL Conference* (pp. 1–28).
- Oakland, T., & Lane, H. B. (2009). Validating the beck depression inventory-II for Hong Kong community adolescents. *International Journal of Testing*, 4(3), 199–216. Retrieved from: <https://doi.org/10.1207/s15327574ijt0403>
- Otto, J., Parviz Partow-Navid, & Doshi., M. (2010). Revisiting the readability of management information systems journals again. *Research in Higher Education Journal*, 1–7. Retrieved from: <https://aabri.com/manuscripts/111011.pdf>
- Owu-Ewie, C. (2014). Readability of comprehension passages in junior high school (Jhs) English textbooks in Ghana. *Ghana Journal of Linguistics*, 68, 35–68.
- Pallant, J. (2013). *A step by step guide to data analysis using IBM SPSS. Igarss 2014* (5th ed.). Berkshire: Open University Press. Retrieved from: <https://doi.org/10.1007/s13398-014-0173-7.2>

- Pallant, J. (2013). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*.
- Priestley, J. (2015). Written communication in academic settings. In P. H. Joubert & S. M. Rogers (Eds.), *Strategic scientific and medical writing: the road to success* (pp. 13–24). Heidelberg: Springer.
- Purdue University. (2007). Genre and the research paper. Retrieved February 10, 2017, Retrieved from: <https://owl.english.purdue.edu/owl/resource/658/02/>
- Rakedzon, T., & Baram-tsabari, A. (2016). Assessing and improving L2 graduate students' popular science and academic writing in an academic writing course. *Educational Psychology*, (31)1, 1–19. Retrieved from: <https://doi.org/10.1080/01443410.2016.1192108>
- Rello, L., Baeza-Yates, R., Dempere-Marco, L., & Saggion, H. (2013). Frequent words improve readability and short words improve understandability for people with dyslexia. *Hall Archives-Ouvertes*. 1–18. Retrieved from: <https://doi.org/10.1007/978-3-642-40498-6>
- Richards, G., & Van Staden, C. (2015). The readability impact of international financial reporting standards. *Pacific Accounting Review*, 27(2), 282–303. Retrieved from: <https://doi.org/10.1108/PAR-08-2013-0086>
- Rukmana, D. (2014). Quota sampling. In A. C. Michalos (Ed.), *Encyclopedia of quality of life and well-being research* (pp. 5382–5384). Dordrecht: Springer Netherlands. Retrieved from: https://doi.org/10.1007/978-94-007-0753-5_2393

- Saiegh-Haddad, E. (2003). Bilingual oral reading fluency and reading comprehension: The case of Arabic/Hebrew (L1) – English (L2) readers. *Reading and writing: An Interdisciplinary Journal*, 16, 717–736.
- Salski, Ł. (2013). Demystifying academic writing. In P. Stalmaszczyk, L. Salski, & W. Szubko-Sitarek (Eds.), *Language Learning, Discourse and Communication* (pp. 37–46). Springer International Publishing, Switzerland. Retrieved from: <https://doi.org/10.1007/978-3-319-00419-8>
- Sawyer, A. G., Laran, J., & Xu, J. (2008). The Readability of marketing journals: Are award-winning articles better written? *Journal of Marketing*, 72(January), 108–117. Retrieved from: <https://doi.org/10.1509/jmkg.72.1.108>
- Schmitt, P. J., & Prestigiacomo, C. J. (2013). Readability of neurosurgery-related patient education materials provided by the american association of neurological surgeons and the national library of medicine and national institutes of health. *World Neurosurgery*, 80(5), 33–39. Retrieved from: <https://doi.org/10.1016/j.wneu.2011.09.007>
- Severance, S. J., & Cohen, K. B. (2015). Measuring the readability of medical research journal abstracts. *Proceedings of BioNLP 15*, (BioNLP), 127–133. Retrieved from: <http://www.aclweb.org/anthology/W15-3815>
- Shelley, M. C., & Schuh, J. H. (2001). Are the best higher education journals really the best? A meta-analysis of writing quality and readability. *Journal of Scholarly Publishing*, 33(1), 11–22. Retrieved from <http://search.proquest.com/docview/57553221?accountid=176251>

- Si, L., & Callan, J. (2001). A statistical model for scientific readability, 574–576. Carnegie Mellon University, Pittsburgh, PA
- Soodmand, H., Moradi, M., & Hamzavi, R. (2014). Frequency and type of hedging devices used in the research articles of humanities, basic sciences and agriculture. *Procedia - Social and Behavioral Sciences*, 136, 70–74. Retrieved from: <https://doi.org/10.1016/j.sbspro.2014.05.290>
- Suleiman, A. B., Lin, J. S., Constantine, N. A., Ballonoff, A., Lin, J. S., (2016). Readability of educational materials to support parent sexual communication with their children and adolescents. 730(5). Retrieved from: <https://doi.org/10.1080/10810730.2015.1103334>
- Tabachnick, B., & Fidell, L. (2013). *Using multivariate statistics* (6th). Retrieved from: <https://doi.org/10.1037/022267>
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th). Boston: Pearson Education.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th). New Jersey: Pearson Education Inc.
- Tekfi, C. (1987). Readability formulas: An overview. *Journal of Documentation*, 43(3), 261–273. Retrieved from: <https://doi.org/http://dx.doi.org/10.1108/eb026811>
- Tin, T. B. (2008). Exploring the nature of the relation between interest and comprehension. *Teaching in Higher Education*, 13(5), 525–536. Retrieved from: <https://doi.org/10.1080/13562510802334764>

- To, V., Fan, S., & Thomas, D. (2013). Lexical density and readability: A case study of English textbooks. *Internet Journal of Language, Culture and Society*, (37), 61–71.
- Tsai, Y. (2010). Text analysis of patent abstracts. *The Journal of Specialised Translation*, (13), 61–80. Retrieved from: http://www.jostrans.org/issue13/art_tsai.php
- UNC. (2010). *Citing sources*. The Writing Center at UNC-Chapel Hill.
- University of Toronto. (2017). Types of journal articles. Retrieved February 10, 2017, from <http://guides.library.utoronto.ca/c.php?g=250816&p=1680054>
- Vera, G. G., Sotomayor, C., Bedwell, P., Dominguez, A. M., & Jeldrez, E. (2016). Analysis of lexical quality and its relation to writing quality for 4th grade primary school students in Chile. *Reading and Writing An Interdisciplinary Journal*. Retrieved from: <https://doi.org/10.1007/s11145-016-9637-9>
- Villiers, C. de, & Dumay, J. (2014). Construction of research articles in the leading interdisciplinary accounting journals. *Accounting, Auditing & Accountability Journal*, 26(6), 876–910. Retrieved from: <https://doi.org/10.1108/AAAJ-Apr-2012-01000>
- Wells, J. A. (1994). Readability of HIV / AIDS educational materials: The role of the medium of communication, target audience, and producer characteristics. *Patient Education and Counseling*, 24, 249–259.
- Wennerstrom, A., & Siegel, A. F. (2003). Keeping the floor in multiparty conversations: Intonation, syntax, and pause. *Discourse Processes*, 36(2), 77-107.

- Whatley, A., Ray, B., & Reidy, R. (1973). An analysis of the readability of selected management journals. *The Academy of Management Journal*, 16(3), 522–524.
- Wheeler, G., & Sherman, T. F. (1983). Readability formulas revisited. *Science and Children*, 20(7), 38–40.
- Williams, K. (2005). Lecturer and first year student (mis)understandings of assessment task verbs: Mind the gap. *Teaching in Higher Education*, 10 (2), 157-173.
- White, J. W., & Lowenthal, P.R. (2011). Academic discourse and the formation of an academic identity: Minority college students and the hidden curriculum. *Review of Higher Education*, 34(2).
- Wright, P.M., & Noe, R.A., (1995). *Management of organizations*. Chicago, IL: Irwin.
- Zaidi, F. H., & Jones, C. A. (2009). Informing patients: Oculoplastic surgery and the internet. *Eye*, 23(11), 2090-2093.
- Zhu, M.S. (1979). An investigation of verbal development of infants: A preliminary analysis of affirmative-declarative [in Chinese]. *ACTA Psychologica SINICA*, 3

APPENDICIES

APPENDIX A: MULTIPLE CHOICE TEST 1

UNIVERSITY OF CAPE COAST

FACULTY OF ARTS

DEPARTMENT OF ENGLISH

A test to determine the relationship between readability and text comprehension

INSTRUCTION: Please read the following passage and answer the questions that follow it.

TEACHING THE REPRESENTATIONS

1. What is the existing scope which needs redefinition in this paper?
 - a) Scope of African American Literature
 - b) Scope of Literature
 - c) Purpose of Africa Literature
 - d) All of the above
2. What does “redefinition” as used in the paper mean?
 - a) Broadening
 - b) Narrowing
 - c) Undefined
 - d) Weakening
3. Which of the following does not include the areas that constitute the scope of “redefinition” of African American Literature?
 - a) Caribbean Francophone and Anglophone works
 - b) Continental Anglophone and Francophone African writings
 - c) Canadian Anglophone writings

- d) African American writings
- 4. Mention one thing that had triggered this paper.
 - a) Writer's experience
 - b) Writer's students
 - c) Writer's research
 - d) Curriculum requirement
- 5. According to the paper, what is one result of globalization?
 - a) Experience
 - b) Illinois State University
 - c) Change in demographics
 - d) World view
- 6. How has that consequence affected the present scope 'undefined'?
 - a) Scope remains narrow
 - b) Scope remains undefined
 - c) Scope needs redefinition
 - d) Scope must be analysed

APPENDIX B: MULTIPLE CHOICE TEST 2

UNIVERSITY OF CAPE COAST

FACULTY OF ARTS

DEPARTMENT OF ENGLISH

A test to determine the relationship between readability and text comprehension

INSTRUCTION: Please read the following passage and answer the questions that follow it.

TEACHING THE REPRESENTATIONS

- 1) Which two main religious figures set out the foundation for this study?
 - a) Traditional and Islam
 - b) Islam and Pentecost
 - c) Christianity and Islam
 - d) Catholic and Islam
- 2) State the difference, if any, between fellowship and followership
 - a) They have no difference
 - b) They mean almost the same thing
 - c) The terms have overlapping meanings
 - d) None of the above
- 3) State the similarity between the two terms as defined in the paper
 - a) One cannot be a fellow without following another
 - b) All fellows must learn to do right
 - c) Following means going after someone
 - d) None of the above

- 4) What is the import of Jesus' question "Do you love me" to Peter to this paper?
 - a) Peter become a shepherd
 - b) Peter was called by Jesus
 - c) Jesus loved Peter
 - d) Peter was following Jesus
- 5) What is the similarity between Christianity and Islam as regards fellowship and followership?
 - a) One cannot fellowship in isolation
 - b) Islam is bigger than Christianity
 - c) Religions teach followership
 - d) Mohammed did what Jesus taught
- 6) Why is it impossible for an individual to claim followership to Mohammed?
 - a) Followership is a group thing
 - b) Fellowship is for individuals
 - c) Mohammed forbids followership
 - d) All of the above
- 7) What is the purpose of the paper?
 - a) Maintain balance between Islam and Christianity
 - b) Maintain balance between fellowship and followership
 - c) Explain leadership in Islam and Christianity
 - d) All of the above
- 8) Into what two major varieties has the paper put Christianity as practiced in Ghana?

- a) Mainline churches and Charismatic Churches
 - b) Mainline churches and Christian churches
 - c) Charismatic Churches and Christian Churches
 - d) None of the above
- 9) Into what main varieties is Islam put?
- a) Tijaniyya sufi and Traditional Islamic sect
 - b) Reformist Wahabiyya sect and Traditional Islamic sect
 - c) Tijaniyya sufi and reformist Wahabiyya sect.
 - d) Traditional Islamic Sect and Modern Islamic sect.

APPENDIX C: MULTIPLE CHOICE TEST 3

UNIVERSITY OF CAPE COAST

FACULTY OF ARTS

DEPARTMENT OF ENGLISH

A test to determine the relationship between readability and text comprehension

INSTRUCTION: Please read the following passage and answer the questions that follow it.

TEACHING THE REPRESENTATIONS

1. What major angle can a society be interpreted from?
 - a. Indigenous spirituality
 - b. Rituals
 - c. World view
 - d. The universe
2. What is the difference between worldview and indigenous knowledge systems?
 - a. Indigenous knowledge systems is the broader view
 - b. World view encompasses indigenous knowledge systems
 - c. Indigenous knowledge systems refer to a people's belief
 - d. World view refers to indigenous knowledge systems
3. How related is indigenous knowledge systems to the paper "Indigenous religious environmentalism in Africa"?
 - a. It covers indigenous spirituality
 - b. It relates to the culture of a people
 - c. It is part of a world view

- d. It is difficult to interpret
4. State three things that indigenous knowledge systems encompass
 - a. Practices, skills, thoughts
 - b. Practices, good, speech
 - c. Practices, height, society
 - d. Society, community, practices
5. State one purpose of the paper “Indigenous religious environmentalism in Africa”?
 - a. The relation between indigenous knowledge systems to a society
 - b. The world view of a people
 - c. Usefulness of indigenous knowledge systems to a society’s conservation
 - d. The purpose sampling technique
6. What is the paper’s view regarding Western colonization’s influence on Asante Sekyere?
 - a. Negative
 - b. Positive
 - c. Undefined
 - d. None of the above
7. Justify your position as stated in your response to question 6.
 - a. Sekyere’s affinity has been weakened
 - b. Sekyere’s affinity has improved
 - c. Westernization is for good of Sekyere
 - d. None of the above

8. What is conclusion has the paper drawn as regards Sekyere's affinity to nature without colonial influence?
 - a. Colonization and Westernization weakened Sekyere
 - b. Colonization brought civilization
 - c. Indigenous knowledge systems encompass world view
 - d. None of the above
9. What is the writer's attitude towards Westernization?
 - a. Positive
 - b. Negative
 - c. Undefined
 - d. All of the above
10. Explain your answer in question 9.
 - a. Westernization ruined positive things in Europe
 - b. Westernization is not beneficial for Africa
 - c. Westernization is good
 - d. None of the above

**APPENDIX D: REQUEST FOR PERMISSION FOR DATA
COLLECTION**

Department of Communication Studies
University of Cape Coast
Cape Coast.
4th February, 2015.

The Headmaster
Ghana National College
Cape Coast.

Dear Madam

Request for Permission for Data Collection

I am writing to request for permission to collect data in your highly esteemed institution for my Ph.D thesis. I am a student in Department of English, University of Cape Coast. As part of my thesis, I am required to administer a comprehension text to SHS 2 AND SHS 3 students.

Being aware of ethical considerations in research, I assure you that I will keep responses of my participants confidential and that for the purposes of anonymity, participants will not have to write their names on the answer sheet provided for the exercise.

I am very confident that my request will be granted.

Thank You.

Yours faithfully,



Mr. William Kodom Gyasi